

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université Larbi Tébessi de Tébessa
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

Département de Mathématiques et Informatiques
Laboratoire des Mathématiques, Informatiques et Systèmes



Une Architecture Basée Sur Le Cloud Computing Mobile Pour L'apprentissage Intelligent

Spécialité : Réseaux et systèmes multimédia

Présentée par :

Nadira BENMEDAKHENE

En vue de l'obtention du diplôme de

Doctorat de 3^{ème} cycle en Informatique

Soutenue le 16 décembre 2021 devant le jury

Président :	Pr.HAKIM BENDJENNA	Université de Tébessa
Examineurs :	Pr.DJAMEL BENMERZOUG	Université de Constantine 2
	Dr.AHMED AHMIM	Université de Souk Ahras
	Dr.LAKHDER LAIMECHE	Université de Tébessa
Directeur de thèse :	Pr.MAKHLOUF DERDOUR	Université d'Oum El Bouaghi
Co-Directeur :	Pr.MOHAMED AMROUNE	Université de Tébessa

Année universitaire : 2021-2022

Je dédie ce travail :

À ma mère

qui m'a soutenue et encouragée durant ces années
d'études. Qu'elle trouve ici le témoignage de ma
profonde reconnaissance.

À la mémoire de mon cher père.

À mes soeurs surtout Nounou, mon frère Salah et sa
femme qui ont partagé avec moi tous les moments
d'émotions lors de la réalisation de ce travail.

À mes nièces et mes neveux et leurs pères.

Et tous mes amies et collegus.

REMERCIEMENT

Je remercie tout d'abord DIEU de m'avoir aidé à surmonter tous les obstacles, et qui m'a donné la force et la patience de poursuivre ce parcours de recherche scientifique. Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je voudrais remercier grandement mon directeur de thèse, Pr. Derdour Makhlouf, pour toute son aide. Je suis ravie d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse.

Je remercie également Dr. Mohamed Amroune, mon co-encadreur de la thèse pour ses remarques qui m'ont permis d'envisager mon travail sous un autre angle. Qu'il soit aussi remercié pour sa gentillesse, sa disponibilité permanente et pour les nombreux encouragements et conseils qu'il m'a prodigués.

Je tiens à remercier particulièrement Dr. Mohamed Gasmi, pour son aide précieuse, pour tous ses conseils qui m'ont accompagné. Il a toujours fait tout son possible pour m'aider, il répond avec calme et patience à mes questions et mes préoccupations. Je lui exprime ici mes sincères remerciements et ma gratitude.

Je remercie profondément les membres du jury pour l'honneur qu'ils m'ont fait de bien vouloir évaluer mon travail.

Mes derniers remerciements vont à tous ceux qui ont contribué, de près ou de loin au déroulement de mon travail. Je tiens à vous exprimer ma plus profonde gratitude.

Résumé

L'émergence des cours en ligne introduit un nouveau paradigme d'apprentissage, grâce auquel l'acquisition de la connaissance en dehors des écoles est garantie, mais la plupart des plateformes d'apprentissage en ligne souffrent de plusieurs limites, qui peuvent dans certains cas atteindre un abandon total des cours, car elles se basent sur un seul discours pédagogique ce qui ne pourra jamais être valable pour tout le monde. En effet, ces dernières années, les environnements d'apprentissage en ligne font l'objet de plusieurs recherches scientifiques, en raison de son association avec d'autres technologies qui ont des fonctionnalités robustes d'adaptation et de recommandation et qui répondent aux nouvelles exigences des utilisateurs, permettant ainsi d'adopter une variation dans les méthodes d'apprentissage.

L'objectif de notre travail est de construire une plateforme adaptative d'apprentissage mobile, basé sur le Cloud Computing mobile, robuste, fiable et permet la prise en compte de plusieurs scénarios pédagogiques selon l'apprenant, ses exigences, son contexte et son environnement. L'architecture de la plateforme proposée regroupe quatre modules complémentaires, il s'agit de : un module de gestion de profil, un module d'adaptation du cours, un module de gestion du contexte en temps réel, et un module pour la personnalisation du contenu. Les données sont collectées via les traces de l'apprenant et ses capteurs mobiles. L'architecture a été comparée avec d'autres modèles de système d'apprentissage en ligne, certains d'entre eux ont été implémentés et d'autres sont représentés comme des propositions détaillées, les résultats de comparaison montrent que notre architecture couvre plus d'aspects et prend plus de scénarios en considération.

Nous avons aussi implémenté une partie du module d'adaptation du profil, par un système automatique basé sur les réseaux de neurones convolutifs, pour la classification des états faciaux des apprenants, qui permet par la suite de définir leurs émotions d'apprentissage, en présentant une nouvelle démarche pour la classification de ce types de données. Le système est basé sur deux étapes fondamentales ; la première consiste à une phase de traitement d'images qui donne comme résultats une nouvelle base de données des états faciaux, qui représente un ou plusieurs états d'actions, ces données vont par la suite dans la deuxième étape servir comme entrée de système de classification, le regroupement de différentes unités d'actions permet de définir les émotions d'apprentissage choisi pour la classification. Les résultats d'expérimentations montrent que notre classificateur offre d'excellentes performances et surpasse les autres classificateurs de littérature, avec un taux de reconnaissance de 98.36% sur le même jeu de données.

Mots clés : Les plateformes d'apprentissage en ligne, l'apprentissage mobile, cours en ligne ouverts et massifs, l'informatique en nuage mobile, les réseaux de neurone convolutifs, la reconnaissance des expressions faciales, les systèmes de codage d'action faciale, les unités d'action.

Abstract

The manifestation of online courses introduces a new learning paradigm through which the acquisition of knowledge outside schools is fulfilled, yet most online learning platforms fall into various limitations. The latter can, at some point, cause early drop-outs from the course as they are generally established following a standard teaching method that doesn't comply with everyone's needs. Indeed, in recent years, online learning environments have been the subject of several scientific researches due to their effective role when using other technologies which, in their turn, have unequivocal adaptation and recommendation functionalities that respond to the users' needs. This can be achieved through varying the learning methods regardless of the limited use of technologies hitherto.

The objective of this thesis is to design an architecture for the construction of a new mobile online learning system based on mobile cloud computing which can be described as powerful and reliable. It therefore opens the door for the diversity of educational settings regarding the learner and his requirements in addition to the learning environment and context.

The suggested architecture presents the coordination and the cooperation mechanism of four fundamental complementary modules; Modules for adapting the profile, adapting the course, a real-time context module, and another for the final personalization. During the process, data is collected via the learner's tracked traces and his mobile sensors. The architecture is compared with other M-learning system models where some of them have been successfully implemented, whereas some others were merely introduced in details. The conclusive comparison results revealed that the study's suggested architectures cover more aspects and take more scenarios into consideration.

The study has also implemented part of the profile adaptation module by an automatic system based on convolutional neuron networks aimed at the classification of learners' facial states. It then allows for an accurate definition of their learning emotions by introducing a new approach for the classification of this type of data. The system is based on two fundamental steps; the first consists of an image-processing phase which provides a new database of facial states as an outcome and represents one or more action units. Thus, data will subsequently be used in the second step as the system classification input. The grouping of different action units makes it possible to define the learning emotions selected for the classification. In conclusion, the experimental results show that our classifier offers excellent performance and thereby outperforms other classifiers with a recognition rate of 98.36% on the same dataset.

Keywords : Online learning platforms, mobile learning, massive open online courses, cloud computing mobile, convolutional neural networks, facial expression recognition, facial action coding systems, action unit.

ملخص

تقدم الدورات التدريبية عبر الإنترنت نموذجًا تعليميًا جديدًا، يتم من خلاله ضمان اكتساب المعرفة خارج المدارس، لكن معظم منصات التعلم عبر الإنترنت تعاني من العديد من القيود، والتي يمكن في بعض الحالات أن تؤدي إلى تخلي المستخدم التام عن الدورة، لأنها تستند إلى خطاب تعليمي واحد وذلك لن يكون صالحًا للجميع. في الواقع، في السنوات الأخيرة، كانت بيانات التعلم عبر الإنترنت موضوعًا للعديد من الأبحاث العلمية، نظرًا لارتباطها بالتقنيات الأخرى، التي تتمتع بوظائف تكيف وتوصية قوية والتي تستجيب لمتطلبات المستخدم الجديدة، مما يسمح بتقديم تنوع في طرق التعلم، على الرغم من أنه حتى اليوم لا يزال هناك استخدام محدود لهذه التقنيات.

الهدف من هذه الأطروحة هو بناء بنية لتصميم نظام جديد للتعلم عبر الإنترنت، يعتمد على الحوسبة السحابية المتنقلة، قوي موثوق ويسمح بتنوع السيناريو التعليمي وفقًا للمتعلّم ومتطلباته وسياقه وبيئته. يعرض الهيكل المقترح آلية التنسيق والتعاون بين أربع وحدات تكاملية وهي؛ وحدات لتكييف الملف الشخصي، تكيف الدورة، وحدة السياق في الوقت الفعلي، وأخرى للتخصيص النهائي، يتم جمع البيانات عبر آثار المتعلمين وأجهزة الاستشعار المحمولة. تمت مقارنة هذا النموذج مع نماذج أخرى لنظام التعلم الإلكتروني تم تنفيذ بعضها وتم تمثيل بعضها كمقترحات مفصلة. تظهر نتائج المقارنة أن نموذجنا أعطى المزيد من الجوانب وأخذ المزيد من السيناريوهات بعين الاعتبار.

لقد قمنا أيضًا بتنفيذ جزء من وحدة تكيف الملف الشخصي، بواسطة نظام آلي قائم على شبكات الخلايا العصبية التلافيفية، لتصنيف حالات الوجه للمتعلّمين، والتي تسمح بعد ذلك بتحديد مشاعر التعلم الخاصة بهم، من خلال تقديم نهج جديد لتصنيف هذا نوع من البيانات. يقوم النظام على خطوتين أساسيتين؛ تتكون الأولى من مرحلة معالجة الصور التي تعطي كنتيجة قاعدة بيانات جديدة لحالات الوجه، والتي تمثل حالة واحدة أو أكثر من الإجراءات، حيث يتم استخدام هذه البيانات لاحقًا كمدخل للنظام عند التصنيف، ويسمح بجميع وحدات العمل المختلفة بتحديد عواطف التعلم المختارة للتصنيف. تظهر النتائج التجريبية أن المصنف الخاص بنا قدم أداءً ممتازًا وتفوق على المصنّفات الأخرى في الأدبيات بمعدل التعرف 98.36% على نفس مجموعة البيانات.

الكلمات المفتاحية: منصات التعلم الإلكتروني، التعلم المتنقل، الدورات التدريبية المفتوحة والواسعة عبر الإنترنت، الحوسبة السحابية المتنقلة، الشبكات العصبية التلافيفية، التعرف على تعبيرات الوجه، أنظمة ترميز الوجه، وحدات العمل.

Table des matières

Résumé	iv
Abstract	vi
Table des matières	ix
Liste des Figures	xiii
Liste des tableaux	xvi
Abbreviations	xviii
Introduction	1
Introduction Générale	1
1 Apprentissage à distance et environnements intelligents	8
1.1 Environnements d'apprentissage à distance	9
1.1.1 Apprentissage électronique (E-learning)	9
1.1.2 Apprentissage Mobile (M-learning)	11
1.1.3 Apprentissage Ubiquitaire (U-learning)	13
1.1.4 Apprentissage pervasive (P-learning)	15
1.2 Sensibilité au contexte	17
1.2.1 Définition du contexte	17
1.2.2 Catégorisation de contexte	18

1.2.3	Systèmes sensibles aux contextes	19
1.2.4	Architecture générale des systèmes sensibles aux contextes	21
1.3	Mobile Cloud Computing	22
1.3.1	Définition de Mobile Cloud Computing	23
1.3.2	Caractéristiques de Mobile Cloud Computing	26
1.3.3	Architecture de Mobile Cloud Computing	27
1.3.4	Avantages et limites de Mobile Cloud Computing	28
1.4	Introduction au deep learning	30
1.4.1	Intelligence Artificielle (IA)	31
1.4.2	Machine Learning (ML)	33
1.4.3	Etapes de Machine Learning	35
1.4.4	Deep Learning (DL)	38
1.4.5	Deep Learning pour l'analyse et le traitement d'images	41
1.4.6	Réseaux de neurones convolutifs (CNN)	43
	Conclusion	48
2	Apprentissage intelligent auto-adaptatif	49
2.1	Apprentissage à distance sensible au contexte et Mobile Cloud Computing	50
2.1.1	Apprentissage à distance	50
2.1.2	Historique de l'apprentissage à distance	51
2.1.3	Apprentissage intelligent	54
2.1.4	Systèmes d'apprentissage à distance sensibles aux contextes	55
2.1.5	Travaux connexes sur l'apprentissage auto-adaptatif	56
2.1.6	Apprentissage mobile basé sur le Cloud	61
2.1.7	Systèmes d'apprentissage à distance auto-adaptatif basé MCC	64
2.2	Apprentissage intelligent et émotions	66
2.2.1	Expressions faciales et unités d'action	67
2.2.2	Types des émotions	69
2.2.3	Deep Learning et expressions faciales	73

2.2.4	Etude sur la reconnaissance des émotions	77
2.2.5	Détection des unités d'actions	80
	Conclusion	82
3	Architecture d'un système d'apprentissage mobile sensible au contexte	84
I.	Etude statistique sur l'acceptabilité de M-learning par les lycéens	85
3.1	Contexte du travail	86
3.1.1	Définition des ressources éducatives libres	87
3.1.2	Modèle d'acceptation de la technologie (TAM)	87
3.1.3	Théorie unifiée de l'acceptation et de l'utilisation de la technologie (UTAUT)	88
3.2	Matériel et méthode	90
3.2.1	Participants	90
3.2.2	Constructions	91
3.2.3	Procédure	91
3.2.4	Hypothèses	93
3.2.5	Création du questionnaire et l'analyse des données	94
3.3	Résultats	95
3.3.1	Statistiques descriptives	96
3.3.2	Composants PLS	97
3.4	Discussion	100
II.	Architecture basée sur MCC pour l'apprentissage auto-adaptatif	103
3.5	Relation entre M-learning et MOOC	103
3.6	Architecture CD-Mobile MOOC	105
3.6.1	Processus de conception de l'architecture CD-Mobile MOOC	105
3.6.2	Description de l'architecture CD-mobile MOOC	107
3.6.3	Personnalisation selon CD-Mobile MOOC	116
3.6.4	Avantages de l'utilisation de Mobile Cloud Computing dans l'archi- tecture proposée	118

3.7	Discussion	120
	Conclusion	122
4	Nouvelle approche d'apprentissage sensible aux émotions	123
4.1	Expressions faciales et apprentissage	124
4.2	Emotions et apprenant	125
4.3	Démarche de creation de classifieur	127
4.3.1	Description des émotions d'apprentissage	127
4.3.2	Description de la base de données	129
4.3.3	Prétraitement des images	133
4.3.4	Création de la nouvelle base de données basée unités d'action . . .	137
4.4	Modèle de classification des états	143
4.4.1	Deep Residual Learning (Resnet)	143
4.4.2	Architecture de classification des états	146
4.5	expérimentation : Résultat et discussion	148
4.5.1	Métriques d'évaluation	149
4.5.2	Synthèse et discussion	155
	Conclusion	156
	Conclusion générale	158
	Conclusion	158
	Perspectives	162
	Bibliographie	164

Table des figures

1.1	Application traditionnelle (sans contexte)	17
1.2	Application sensible au contexte	20
1.3	Architecture minimale d'un système sensible au contexte (Soualah-alila, 2016)	22
1.4	Moblie Cloud Computing	23
1.5	Architecture Mobile Cloud Computing (Shamim et al.,2015)	27
1.6	Positionnement de DL par rapport au AI et ML	30
1.7	Le fonctionnement de poids et de bias	36
1.8	Fonctionnement d'un neurone artificiels	37
1.9	La phase d'apprentissage (Yufeng G,2017)	37
1.10	La phase de prédiction	38
1.11	La différence entre le fonctionnement du ML et DL	39
1.12	Processus d'apprentissage (deep Learning) pour identifier un chat	41
1.13	Présentation schématique du modèle ANN profond pour la classification d'image	43
1.14	Révolution d'intérêt pour les réseaux de neurones convolutifs	43
1.15	L'architecture de réseau de neurones convolutifs (Prabhu, 2018)	44
1.16	Volume de matrice d'image, filtre, et la matrice de sortie	45
1.17	Application de padding p=1	45
1.18	Fonctionnement de la fonction RELU (Prabhu, 2018)	46
1.19	Fonctionnement de Max pooling	47
1.20	Fonctionnement d'Average pooling	47
1.21	La couche entièrement connectée	48

2.1	École privée saint-nicolas (paris-normadine, 2019) et Une classe de 1955(canal- blog, 2017)	51
2.2	L'historique de l'apprentissage à distance partie 01	52
2.3	L'historique de l'apprentissage à distance partie 02	53
2.4	L'historique de l'apprentissage à distance partie 03	54
2.5	Un échantillon des émotions universelles depuis la base de données CK	68
2.6	La roue d'émotion de Plutchik traduit en français	71
2.7	Les caractères de la base de données FERG-DB	76
2.8	Les performances des AU individuelles sur la base de données Bosphore (Gehrig et al., 2015)	81
3.1	Le Modèle TAM (Davis et al., 1989)	88
3.2	Le modèle UTAUT (Venkatesh et al., 2003)	89
3.3	L'acceptation du type d'OER choisi	96
3.4	Les étapes de fonctionnements de l'architecture	107
3.5	Distribution géographique	109
3.6	Le niveau éducatif des apprenants	109
3.7	Architecture de CD-Mobile MOOC	111
3.8	Le modèle d'adaptation du profil	116
3.9	CD Mobile-Mooc basée sur le cloud computing	119
4.1	Les muscles avec la liste des unités d'actions associées	132
4.2	Facial Action Coding System (FACS) (Yin et al., 2017)	133
4.3	Les émotions universelles depuis la base de données CK+	133
4.4	Les muscles responsables d'AU4	136
4.5	Les muscles responsables d'AU12 et AU14	136
4.6	Face Landmark 68 points	138
4.7	Landmark et rectangles de division (a) : Les points de repère. (b) : Les rectangles de chaque groupe	139

4.8	Recadrage de l'image sans la valeur α (a). Recadrage de l'image avec la valeur α (b)	140
4.9	Ensemble d'images générées par les techniques d'augmentation	143
4.10	Traitement d'images : image en gris, extraction du visage puis la division d'images en groupe et redimensionnement (Exemple 01 permet d'extraire l'état 09, l'exemple 02 permet d'extraire les états E1, E5, E6, E7, E12	144
4.11	Erreur d'entraînement (gauche) et erreur de test (droite) sur CIFAR-10, (He et al., 2015)	145
4.12	Residual learning : exemple d'un bloc, (He et al., 2015)	145
4.13	L'architecture de notre modèle de reconnaissance d'états faciaux adaptés au ResNet. Chaque rectangle est un bloc Conv comprenant une couche convolutionnelle, BatchNormalization et une fonction d'activation (Relu) ces derniers ne sont pas affichés par souci de simplicité, pour le bloc conv en bleu il contient seulement une couche convolutionnelle	147
4.14	La description détaillée de la construction d'un bloc	148
4.15	Les courbes : d'exactitude de la validation (a) et d'erreur de validation (b) . . .	149
4.16	Matrice de confusion	153
4.17	Quelques émotions d'apprentissage, imitées par des lycéens	155

Liste des tableaux

1.1	Résultat d'ImageNet challenge (de 2010 à 2017)	40
2.1	Représentation des quelques études réalisées dans l'apprentissage auto-adaptative	57
2.2	Les points forts des travaux réalisés dans l'apprentissage auto-adaptative .	60
2.3	Les points faibles des travaux réalisés dans l'apprentissage auto-adaptative	61
2.4	Définition des émotions universelles	67
2.5	Intensité de FACS	69
2.6	Les unités d'action de chaque émotion basique	70
2.7	Les émotions utilisées dans quelques recherches	72
2.8	Ensemble de bases de données utilisées pour la reconnaissance des expressions faciale	75
2.9	Comparaison de la précision de la reconnaissance des émotions basiques . .	79
2.10	La détection des AU sur BP4D selon l'approche pEAC-Net avec de grandes poses de tête(%)	82
3.1	Description des participants	90
3.2	Description des constructions	91
3.3	Description des hypothèses	93
3.4	Partie du questionnaire	94
3.5	Pourcentage des réponses de chaque construction	97
3.6	Le modèle de mesure	98
3.7	La valeur de la validité discriminante	98

3.8	Le modèle structurel	99
3.9	L'effet des variables indépendantes	99
3.10	Pertinence prédictive (Q^2)	99
4.1	La définition des émotions d'apprentissage	126
4.2	Les possibilités de présence de chaque unité d'action pour les émotions d'apprentissage choisi	130
4.3	Les unités d'action et leur région	134
4.4	Les états et les groupes selon les régions des unités d'action	135
4.5	Les possibilités de présence de chaque état pour les émotions d'apprentis- sage choisi	137
4.6	Les coordonnées de rectangles de division des images d'état	140
4.7	Comparaison des résultats d'exactitude avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score . . .	150
4.8	Comparaison des résultats d'exactitude avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score . . .	151
4.9	Evaluation et comparaison de f1-score avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score . . .	153
4.10	Matrice de confusion	154

Abbreviations

E-Learning	Apprentissage électronique
M-Learning	Apprentissage Mobile
U-learning	Apprentissage Ubiquitaire
P-learning	Apprentissage pervasive
MCC	Mobile Cloud Computing
DL	Deep Learning
ML	Machine Learning
IA	Intelligence Artificielle
ANI	Intelligence artificielle faible
AGI	Intelligence artificielle forte
ASI	Super Intelligence Artificielle
ANN	Artificial neural network
CNN	Convolutional neural network
MOOC	Massive Open Online Courses
STI	Système Tutoriel Intelligent
AU	Action unit
FACS	facial action coding system
FER	facial expression recognition
OER	Open Educational Resource

Introduction Générale

Contexte

Le 19 mars 2020, L'Unesco déclare dans son site officiel que plus de 850 millions d'apprenants ont dû rester à l'écart des écoles et des universités en raison de la pandémie de COVID-19. Des fermetures à l'échelle nationale dans 102 pays et des fermetures locales dans 11 autres. Un an plus tard la situation s'est améliorée mais elle reste encore dangereuse selon l'Unicef qui déclare dans un article publié en mars 2021, que les écoles de plus de 168 millions d'enfants dans le monde sont complètement fermées depuis près d'un an. Vue cette situation critique les gouvernements partout dans le monde cherchent à assurer la continuité d'apprentissage, ce qui mène vers l'enseignement à distance, et c'est là où le domaine a connu un intérêt mondial, plusieurs recherches et outils technologiques dédiés à l'apprentissage en ligne ont été exploités depuis des décennies, et c'est le fruit de ces recherches qui permet de minimiser les conséquences des fermetures d'écoles causées par la pandémie.

Le domaine d'apprentissage en ligne devient de plus en plus l'une des stratégies fondamentales de l'université algérienne en particulier et des administrations, des collèges en général. Cette stratégie permet aux universités de trouver un moyen d'enrichir l'apprentissage par les technologies pédagogiques, garantissant la compréhension et facilitant l'intégration en classe. Mais La plupart des systèmes d'apprentissage en ligne souffrent de plusieurs lacunes et problèmes, ils s'avèrent inefficaces dans plusieurs cas et peuvent diriger l'apprenant vers l'abandon du cours. Ces systèmes ne sont pas en mesure de traiter les différents modes d'apprentissage, de plus, leur transmission d'information est souvent

statique, pourtant les apprenants sont de natures différentes, un seul discours pédagogique n'est jamais valable pour tout le monde. Toutes ces plateformes traditionnelles ne sont pas suffisamment fiables elles ne garantissent pas la continuité d'apprendre ni de comprendre le besoin adéquat par rapport à la situation de l'apprenant.

Afin d'affronter tous ces problèmes d'efficacité et de diversité, la révolution technologique fait référence à l'utilisation du contexte d'apprenants ou il faut utiliser toutes les ressources possibles pour personnaliser non seulement le cours mais même la plateforme ou le système utilisé. Les systèmes sensibles au contexte permettent de créer un environnement unique pour chaque apprenant, Ils s'adaptent avec ses préférences et son environnement pour lui donner l'impression que cette plateforme est conçue spécialement pour lui, ils facilitent son mode de vie éducatif en délivrant les savoirs selon son propre rythme ce qui le rend plus stable et motivé. Ce principe, en réalité, reflète le type d'apprentissage positif à l'école, où l'enseignant prend en considération le niveau et la situation de chaque élève lors de l'explication du cours, si la technologie d'aujourd'hui peut être correctement utilisée pour permettre à chaque apprenant d'apprendre à un rythme raisonnable en exploitant correctement ses ressources, l'apprentissage en ligne va connaître une évolution très importante qui permettra d'imiter efficacement les avantages de l'apprentissage classique.

Problématique

L'apprentissage en ligne souffre de plusieurs problèmes, dont le plus remarquable et le plus important est l'abandon (Hansch et al., 2015), ce phénomène est défini dans ce domaine comme l'acte de laisser le suivi d'un cours après s'être inscrit, peu importe le taux de parcours atteint, donc la plupart des plateformes déclarent un nombre d'inscrits largement différent du nombre d'apprenants qui sont arrivés à suivre et à terminer les cours. Il existe plusieurs raisons pour cet acte et sa dilatation, il peut s'agir d'un manque d'intérêt (Reich et Ruipérez-Valiente, 2019), manque de motivation ou d'adaptation, comme il peut aller au-delà, et être un problème de différence entre ce que l'utilisateur cherche

et ce qui lui a été présenté. En partant de ces points la personnalisation de l'apprentissage est apparue comme la solution la plus adéquate à ce problème, la personnalisation garantie un processus d'apprentissage conforme au besoin de l'apprenant en temps réel, son environnement éducatif subit un réglage périodique selon les changements du profil et d'objectif, ce qui permet une distribution d'un apprentissage de qualité en fonction des besoins et auto-adaptatif selon le contexte.

L'utilisation des systèmes sensibles au contexte pour les plateformes éducatives a été étudiée par de nombreuses recherches (Monterrat et al., 2015 ; Mehmood et al., 2017 ; Tarus et al., 2018), qui ont été développées et appliquées afin de délivrer des solutions exploitables dans la réalité. L'utilisation des plateformes éducatives sensibles au contexte augmente de plus en plus la satisfaction des apprenants, mais cela introduit une quantité de traitement de données assez importantes.

Lorsque nous observons l'apprentissage mobile (M-learning) en particulier, comme c'est le plus utilisé parmi les lycéens et les universitaires, l'utilisation d'un mobile pour le suivi du cours, la collection de différents changements qui entourent l'apprenant, le traitement de ces variables, la prise de décision en temps réel etc., présentent les principaux défis qui peuvent diminuer considérablement l'efficacité de ces systèmes. Toutefois, il faut combiner des technologies afin d'arriver à la meilleure solution, l'utilisation des services du Cloud Computing mobile pour les M-learning permet aux utilisateurs de recevoir des services plus puissants, surtout au niveau de calcul et de stockage. Elle est considérée aussi comme l'une des techniques qui peut servir à améliorer les architectures et les plateformes d'apprentissage en ligne.

Ce domaine est immense, comme ses exigences, trouver une solution qui couvre tout le domaine est un peu difficile voire impossible, c'est pour cela qu'il faut se concentrer sur un seul type de cours en particulier, ensuite trouver les défis et les solutions possibles, dilater-les enfin, pour inclure et résoudre le plus possible de problèmes similaires.

M-learning présente plusieurs limites, l'autonomie de la batterie (Yousafzai et al, 2016), la taille de l'écran (Davies et al., 2012), la fiabilité du réseau mobile, l'hétérogénéité des systèmes (Li et Gao, 2017), la mémoire limitée, le risque de distraction, etc. Tous ces pro-

blèmes sont censés être résolus pour arriver à un système d'apprentissage qui répond aux attentes et aux demandes des apprenants, l'utilisation de MCC est très bénéfique pour ce type de problèmes, mais il reste à déterminer le type de variables contextuelles à utiliser, les modules du traitement envisagés, leurs coordinations, et la fluidité de coopération. Un autre défi, souvent négligé dans les recherches de ce domaine, est l'interaction positive, ce qui veut dire l'interaction entre l'apprenant et l'équipe pédagogique.

Reconnaître les émotions des apprenants aidera à atteindre un niveau d'interaction très important. L'utilisateur des plateformes éducatives peut suivre le cours sans réellement le suivre pour obtenir le certificat. Mais arriver à détecter ses émotions d'apprentissage aide à diminuer les suivis errants, le domaine de la reconnaissance des expressions faciales (FER) est devenu de nos jours de plus en plus utilisé et impliqué dans de multiples applications, car le visage donne une vaste quantité d'informations sur l'état de l'utilisateur ce qui renforce de plus en plus les systèmes d'interaction homme machine. Les principaux défis qui peuvent être rencontrés lors de l'utilisation de ce domaine sont la quantité et la qualité des données de traitement, comme c'est censé d'être exploité pour les applications d'apprentissage, les émotions à traiter doivent répondre aux besoins d'un M-learning.

Les émotions d'apprentissage ne sont pas très connues dans les recherches FER, car ils sont complexes, rares et visent un seul domaine, contrairement aux émotions universelles qui sont impliquées dans presque toutes les applications de reconnaissances des expressions faciales. On distingue généralement trois phases importantes pour classifier les émotions faciales, d'abord la phase de détection du visage, ensuite la phase d'extraction de caractéristiques pour arriver en fin à la phase de classification, l'utilisation de l'apprentissage en profondeur (Deep Learning), permet de nous libérer de ces tâches, par contre une étape de prétraitement est souvent obligatoire afin d'arriver aux résultats attendus. Une combinaison bien étudiée des trois domaines (systèmes sensibles au contexte, Cloud Computing Mobile, et le Deep Learning) semble prometteuse.

Objectifs et contributions

L'objectif principal de cette thèse est de concevoir une nouvelle architecture pour les plateformes d'apprentissage intelligentes basées sur le Cloud Computing mobile, qui fait face aux limites dévoilées précédemment. Ce travail est présenté dans le cadre de recherche de la coopération la plus idéale entre les technologies de nos jours pour proposer l'architecture la plus appropriée d'un système d'apprentissage intelligent, et qui permet de surmonter les défis observés dans les M-learning et de classifier les émotions des apprenants d'une manière fiable et efficace. Le travail est présenté sur deux contributions complémentaires :

- Dans la première contribution, nous avons proposé un questionnaire, qui permet d'examiner l'acceptation et le comportement des lycéens envers l'utilisation des ressources éducatives libres, présentées sous le cadre d'un M-learning, afin de déterminer l'importance et l'influence de ce domaine sur l'apprenant algérien, de plus cela nous a aidé à extraire les obstacles rencontrés par ces élèves lors d'une éducation ouverte avec des pratiques innovantes dans un processus d'apprentissage digitalisé. Ces obstacles ont servi comme piliers de base lors de la conception des modules de l'architecture finale. Dans la deuxième étape de cette contribution nous avons exploité les services CCM, la technologie M-learning et les MOOC pour aboutir à modéliser une architecture basée Cloud Computing Mobile comme solution des défis extraits dans la première partie. Cette modélisation passe par plusieurs étapes d'analyse et définit des modules d'adaptation du profil, d'adaptation du cours, un module de contexte temps réel, et un autre pour la personnalisation, la coopération entre ces modules permet au système de générer un style d'apprentissage individuel.
- La deuxième contribution aborde la conception et la réalisation de la phase de classification des états faciaux, définie selon les unités d'actions qui présentent les émotions complexes, dont leur regroupement permet au système de classer les émotions d'apprentissage choisies. Ce système est composé de deux phases, la première est celle de prétraitement d'images de la base de données, qui passe par plusieurs étapes afin

de définir les états faciaux nourrissant le modèle d'apprentissage en profondeur, la deuxième phase consiste à créer le modèle d'apprentissage, basé sur l'architecture d'apprentissage en profondeur ResNet, et qui a donné de très bons résultats lors de son utilisation sur les visages.

Plan de la thèse

Cette thèse a été organisée en quatre chapitres. Les deux premiers chapitres sont consacrés à la présentation de la terminologie utilisée et une revue de la littérature liée à nos recherches, alors que les deux derniers chapitres présentent nos contributions. La description de chaque chapitre est la suivante :

— Chapitre 01 :

Ce chapitre consiste à présenter, le cadre théorique de notre recherche qui se base sur quatre axes essentiels. Il commence par les environnements d'apprentissage en ligne, ensuite il présente les systèmes sensibles au contexte en passant par le Cloud computing mobile et se termine par une introduction du Deep Learning.

— Chapitre 02 : Ce chapitre est divisé en deux parties, la première partie consiste à présenter et à synthétiser la littérature dans les systèmes d'apprentissage en ligne sensibles au contexte. La deuxième partie présente un état de l'art des systèmes de reconnaissance des émotions faciales et les unités d'action en utilisant les réseaux de neurones convolutifs. Pour les deux parties nous nous sommes concentrés sur l'avantage d'associer les Cloud Computing Mobile avec ces systèmes.

— Chapitre 03 : Ce chapitre décrit la première contribution de cette thèse, il est divisé en deux parties : la première présente une enquête statistique sur les élèves du lycée, afin d'examiner leurs acceptations et comportements envers l'utilisation des ressources éducatives ouvertes sur des M-learning, cela permet aussi de déterminer les limites du système selon leurs point de vue. La deuxième partie propose une nouvelle architecture basée sur le Cloud Computing mobile afin de surpasser les limites vécues par les élèves et d'offrir un environnement d'apprentissage mobile personnalisé selon les besoins de chaque apprenant.

-
- Chapitre 04 : Ce chapitre décrit la réalisation d'une partie du module d'adaptation du profil de notre plateforme, il s'agit de la classification des états faciaux des apprenants qui permet par la suite de définir leurs émotions d'apprentissage, ce chapitre présente en détails une nouvelle démarche pour la classification de ce type de données.
 - Conclusion et perspective : La conclusion résume les contributions de cette thèse et discute certains travaux futurs potentiels.

Chapitre 1

Apprentissage à distance et environnements intelligents

Contents

1.1 Environnements d'apprentissage à distance	9
1.1.1 Apprentissage électronique (E-learning)	9
1.1.2 Apprentissage Mobile (M-learning)	11
1.1.3 Apprentissage Ubiquitaire (U-learning)	13
1.1.4 Apprentissage pervasive (P-learning)	15
1.2 Sensibilité au contexte	17
1.2.1 Définition du contexte	17
1.2.2 Catégorisation de contexte	18
1.2.3 Systèmes sensibles aux contextes	19
1.2.4 Architecture générale des systèmes sensibles aux contextes . . .	21
1.3 Mobile Cloud Computing	22
1.3.1 Définition de Mobile Cloud Computing	23
1.3.2 Caractéristiques de Mobile Cloud Computing	26
1.3.3 Architecture de Mobile Cloud Computing	27
1.3.4 Avantages et limites de Mobile Cloud Computing	28
1.4 Introduction au deep learning	30
1.4.1 Intelligence Artificielle (IA)	31
1.4.2 Machine Learning (ML)	33
1.4.3 Etapes de Machine Learning	35
1.4.4 Deep Learning (DL)	38
1.4.5 Deep Learning pour l'analyse et le traitement d'images	41
1.4.6 Réseaux de neurones convolutifs (CNN)	43
Conclusion	48

L'acquisition de la connaissance en dehors des écoles a été un problème majeur avant l'émergence des cours en ligne qui ont introduit un nouveau paradigme d'apprentissage. Ces dernières années, les environnements d'apprentissage en ligne font l'objet de plusieurs recherches scientifiques, en raison de la combinaison entre ce domaine et d'autres technologies qui ont des fonctionnalités robustes d'adaptation et de recommandation et qui répondent aux nouvelles exigences des utilisateurs, permettant de délivrer une variation dans les méthodes d'apprentissage.

Dans ce chapitre, nous allons présenter, le cadre théorique de notre recherche qui se base sur quatre axes essentiels. Nous commençons par les environnements d'apprentissage en ligne, ensuite nous verrons les systèmes sensibles au contexte, passant par le Cloud computing mobile et finissons par une introduction au deep Learning.

1.1 Environnements d'apprentissage à distance

En général les apprenants dans les cours en lignes sont de natures différentes : âge, niveau, méthode d'apprentissage, etc. Délivrer un seul scénario pédagogique, créé souvent des encouragements pour certains types d'apprenants. Vu cette nature de changement de l'être humain beaucoup d'environnements d'apprentissage en ligne ont été délivré l'un après l'autre.

1.1.1 Apprentissage électronique (E-learning)

Dans le domaine d'auto apprentissage, les e-learning présentent un matériel très populaire et répandu, son utilisation a prouvé des améliorations aux compétences des apprenants dans des différents domaines, au fur et à mesure au cours des années les e-learning devenus une nécessité dans la plupart des écoles et des universités, surtout lors de ces dernières années, où l'internet les ordinateurs et les appareils mobiles sont partout. Ces cours sont transmis par internet aux étudiants par leurs enseignants ou autres éducateurs. Selon l'union européenne (UE, 2008) « l'e-learning est l'utilisation des nouvelles technologies multimédias de l'internet pour améliorer la qualité de l'apprentissage en facilitant

d'une part l'accès à des ressources et des services, d'autre part les échanges et les collaborations à distance ». La plupart des universités Algériennes utilisent des systèmes e-learning comme celles de Beji mokhtar¹ Annaba ou Larbi Tebessi Tébessa² dont les deux utilisent MOODLE comme plateforme pédagogique.

Les caractéristiques d'e-learning :

D'une part (Julie et al., 2012) ont retiré des caractéristiques importante de E-learning comme suit :

- Dialogue : entre les acteurs d'e-learning (apprenant, enseignant...) qui améliore les résultats et accélère l'apprentissage des personnes.
- Formation auto-dirigée : Utilisation d'outils d'animation et multimédia.

D'autre part (Behera et al., 2013) ajoute de nouvelles caractéristiques résumées par :

- Accessibilité : les cours sont faciles à obtenir, les apprenants peuvent les suivre à tout moment.
- Renforcement par la technologie numérique : l'un des points fort de l'e-learning est l'utilisation des technologies multimédias et numériques.
- Flexibilité : la flexibilité de l'e-learning en termes de diffusion de supports et de type de cours.

Nous avons également extrait de la définition d' e-learning par l'Union européenne (EU, 2008) les caractéristiques suivantes :

- Utilisation de la technologie : en utilisant Internet et des technologies innovantes pour transférer les connaissances.
- Accessibilité : en facilitant l'accès aux ressources et services.
- Collaboration : en utilisant le forum de la plate-forme ou toute autre offre de technologies de communication.

1. <https://elearning.univ-annaba.dz/>

2. <http://e-learning.univ-tebessa.dz/moodle/>

Une autre caractéristique peut être trouvée dans le domaine de l'apprentissage en ligne, comme la conscience du contexte et l'adaptabilité lié davantage à la plateforme d'apprentissage.

Par conséquent, les modèles constructivistes suggèrent que l'apprentissage en ligne peut être caractérisé par un apprentissage autorégulé : cela est également connu sous le nom d'apprentissage autodirigé, les individus apprennent mieux lorsqu'ils découvrent eux-mêmes des choses à leur propre moment et lieu.

1.1.2 Apprentissage Mobile (M-learning)

Comme les E-learning, les M-learning sont des environnements d'apprentissage en ligne. Comme le montre la nomination, ces types d'environnements se basent sur la technologie mobile, les cours sont délivrés aux apprenants dans leurs appareils mobiles et personnels, comme les smartphone, tablettes, et laptop etc. Beaucoup de recherches définissent ce terme, (kristime et al., 2007) déclare que les M-learning sont parfaitement adaptés à la notion de « *just enough, just in time, just for me* », (Behera et al., 2013) résume la définition de M-learning dans une seule déclaration «Livraison d'éducation ou de tout type d'apprentissage via des appareils portables ». Par conséquent, pour résumer, nous pouvons définir le m-learning comme un apprentissage en ligne dispensé à différents apprenants via leurs appareils portables.

Certains systèmes m-learning importants sont fournis par différentes applications mobiles, par exemple, MosaLingua_IOS³ est une application de formation linguistique sur un smartphone, illustre la capacité de conserver des informations pendant une longue période, en utilisant un système de révision intelligent.

Les caractéristiques de M-learning :

Outre la mobilité, qui est une caractéristique majeure des environnements d'apprentissage Mobile, plusieurs recherches présentent divers autres caractéristiques qui sont à la base de ce type de service en ligne apprentissage. (Behera et al., 2013) a retiré certaines caractéristiques importantes de M-learning :

3. <https://www.mosalingua.com/blog/2019/09/04/applications-ios-mosalingua/>

- Accessibilité : Les cours sont disponibles pour les apprenants à tout moment.
- Immédiateté : les apprenants peuvent récupérer des informations immédiatement.
- Interactivité : l'apprenant peut contacter des enseignants, des experts ou un autre apprenant, par forum, médias sociaux ou audioconférence.
- Basé sur les technologies modernes : assurer un plus grand engagement des étudiants parce que le m-learning se base sur leurs technologies habituelles.

(Fezile et al.,2011) constatent que les caractéristiques essentielles de M-learning sont :

- Taille portable des outils mobiles : Smartphones, tablettes... sont petites à tenir et disponibles, l'étudiant peut les utiliser partout.
- Privé : chaque apprenant à son propre outil, il se connecte et télécharge indépendamment des autres apprenants.
- Interactif : la plateforme m-learning utilise les dernières technologies pour intégrer un environnement d'apprentissage interactif dans les activités d'apprentissage et d'enseignement.
- Collaboratif : la communication entre les apprenants et les enseignants est l'une des bases de l'apprentissage mobile.
- Information instantanée : l'immédiateté est garantie grâce aux outils mobiles.

(Seppälä et al.,2003) Affirment que la caractéristique fondamentale de l'apprentissage mobile permet aux apprenants d'être au bon endroit au bon moment, c'est-à-dire « *to be where they are able to experience the authentic joy of learnin* ».

Les limites de M-learning :

Malgré l'importance majeure de l'utilisation de mobile pour apprendre, les apprenants souffrent toujours de certaines limites causées par l'appareil, citons par exemple :

- Connectivité et autonomie de la batterie.
- Taille de l'écran.
- Problèmes de fiabilité du réseau mobile.

- Systèmes d'exploitation multiples.
- Mémoire limitée.
- Vitesse de téléchargement et accès Internet restreint.
- Risque de distraction.
- Problèmes de coût.
- Problèmes d'utilisabilité.

1.1.3 Apprentissage Ubiquitaire (U-learning)

Les environnements d'apprentissage ubiquitaire sont dérivés de l'informatique ubiquitaire défini par (Weiser, 1999) « Les technologies les plus profondes sont celles qui disparaissent. Elles se faufilent dans le tissu de la vie quotidienne jusqu'à ce qu'elles ne puissent plus en être distinguées », ajoute-t-il également, que ce phénomène relatif à la psychologie humaine, "Une telle disparition est une conséquence fondamentale non pas de la technologie mais de la psychologie humaine". Ensuite beaucoup de définitions ont été proposées dans ce sujet (yang et al., 2008) Confirment qu' «*une définition plus stricte de u-learning consiste à apprendre avec la technologie u-computing*». (Hwang et al., 2008) expriment également que «tout environnement d'apprentissage permettant aux apprenants d'accéder au contenu n'importe où et à tout moment, peu importe si des communications sans fil ou des appareils mobiles sont utilisés».

Par conséquent, pour garantir un environnement d'apprentissage ubiquitaire, nous devons intégrer les dispositifs qui garantissent l'adaptation au contexte de nos vies. (Saadiah et al., 2010) proposent une définition pour reprendre tout l'aspect de l'apprentissage en ligne, déclarent-ils, «l'apprentissage en ligne est un paradigme d'apprentissage qui se déroule dans un environnement informatique ubiquitaire qui permet d'apprendre la bonne chose au bon endroit et au bon moment de la meilleure façon". Par suite, en conséquence, ce type d'apprentissage donne aux apprenants la meilleure façon satisfaisante d'étudier dans les meilleures conditions. (Jones et al., 2004) ajoutent : «*Un environnement d'apprentissage ubiquitaire est tout cadre dans lequel les étudiants peuvent s'immerger totalement dans le processus d'apprentissage*». Par conséquent, l'objectif principal de l'apprentissage

ubiquitaire est de propager l'apprentissage et de le rendre énormément disponible d'une façon telle que les étudiants deviennent si profondément impliqués. Duolingo⁴ présente un système d'apprentissage omniprésent, un site Web et une application mobile, pour l'apprentissage des langues à l'aide d'un arbre de compétences que les utilisateurs peuvent progresser à travers le vocabulaire, la lecture, l'écriture et les expressions orales.

Caractéristiques de l'apprentissage Ubiquitaire :

D'une part dans son travail (Gwo-Jen Hwang, 2006), citent quelques exemples critères d'apprentissage ubiquitaire :

- Un environnement d'apprentissage ubiquitaire fournit activement des supports personnalisés.
- Un environnement d'apprentissage ubiquitaire permet un apprentissage transparent partout et à tout moment.
- Un environnement d'apprentissage ubiquitaire est capable d'adapter le contenu du sujet pour répondre aux fonctions de divers appareils mobiles.

D'autre part, (Saadiah et al,2010) Proposent également des caractéristiques à partir de la combinaison des idées de certains chercheurs et prennent en compte les principales différences.

- Permanence : Les informations sont permanentes, sauf pour l'abolition volontaire.
- Accessibilité : les informations sont disponibles chaque fois que les apprenants ont besoin.
- Immédiateté : les apprenants peuvent interagir facilement et efficacement avec les enseignants, les experts et les pairs.
- Sensibilisation au contexte : des informations appropriées sont fournies aux apprenants selon leur environnement.

4. <https://fr.duolingo.com/>

1.1.4 Apprentissage pervasive (P-learning)

Dans l'apprentissage pervasive, les informations sur le contexte de l'apprentissage proviennent de l'environnement d'apprentissage, où des petits appareils tels que les capteurs, les tampons, les badges sont intégrés l'un avec l'autre. Comme U-learning, le P-learning est un domaine de recherche basée sur l'informatique pervasive. D'après (Weiser, 2015) «L'informatique pervasive intègre le calcul dans l'environnement, plutôt que d'avoir des ordinateurs qui sont des objets distincts ». Ainsi, l'apprentissage pervasive utilise la technologie d'informatique omniprésente dans le domaine de l'apprentissage pour renforcer la capacité de collecter des informations à partir de l'environnement de l'apprenant en utilisant différents appareils. Comme les autres champs d'apprentissage p-learning ont également une définition multiple. Dan Pontefract dans son livre Flat Army (Dan, 2013) définit p-learning en tant que «L'apprentissage à la vitesse du besoin », cela donne au domaine d'apprentissage plus de possibilités de créer un moment de besoin d'apprentissage, (Lucke et al., 2013) croient que le contexte et les systèmes sensible au contexte sont le pouvoir majeur de p-learning. Nous pouvons donc dire que l'environnement d'apprentissage pervasive personnalise le parcours pédagogique de chaque apprenant, via son contexte et cela diminuera le plus souvent l'abandon et augmentera l'efficacité.

Caractéristiques du P-learning :

(Siobhan, 2006) présente l'apprentissage pervasive comme l'idée de toujours être actif "always-on", il en présente également quatre (04) principales composantes comme suit :

- Communauté : l'apprenant est au centre de la situation d'apprentissage, il obtient des informations d'une communauté d'enseignants et d'autres apprenants, tout en diffusant les connaissances.
- Autonomie : les apprenants sont libres ; il n'y a pas de structure d'autorité, pas de règles fixes. Ils peuvent être à l'aise avec les connaissances fournies.
- Lieu : l'apprentissage n'est pas seulement en classe, il n'y a pas de limite d'espace.
- Rationalité : l'apprenant peut créer un moyen de relier les connaissances à sa vie parce qu'il apprend dans son environnement.

(Kearney et al., 2012) considèrent les facteurs situationnels et contextuels comme des facteurs influençant l'apprentissage pervasive.

(Hundebol et al., 2006) considèrent l'apprentissage pervasive comme une activité liée au contexte, «*L'environnement d'apprentissage pervasive est le contexte de médiation de l'apprentissage, dans un environnement physique enrichi d'éléments supplémentaires, spécifiques au site et dépendants de la situation*».

Ainsi, l'apprentissage pervasive crée un environnement dans lequel l'apprenant devient plus autonome, (Nguyen, 2010) soutient que cette autonomie est renforcée par la contribution de l'enseignant dont le rôle est de faciliter la continuité entre le contexte d'apprentissage et par la bonne conception pédagogique. Il ajoute également que les activités d'apprentissage les plus externes sont effectuées par l'apprenant seul, l'enseignant intervient à distance en faisant des commentaires ou en les guidant.

Dans cette section, nous avons présenté une définition claire de l'environnement d'apprentissage (E, M, U, P) basée sur leurs caractéristiques. L'utilisation de l'un ou l'autre de ces types d'apprentissage dépend du besoin des apprenants, des caractéristiques des cours et de la vision de l'enseignant. Même s'ils offrent un nouveau paradigme d'apprentissage, ils ont de nombreuses faiblesses qui forment des perspectives et des défis de domaine.

L'utilisation généralisée de cet environnement d'apprentissage est encore étrange dans certaines universités, en particulier en Algérie, même si elle fournit des résultats très importants en termes de progrès. Le suivi des cours dans des plateformes éducatives différentes dépend du type d'apprentissage, du choix de l'équipe éducative et de la nature des apprenants. L'un des inconvénients majeurs pendant le suivi d'un cours en ligne est la perte de motivation et la diversité des apprenants qui génère un grand besoin d'adapter les services offerts par les plateformes éducatives au besoin de l'utilisateur. Parmi les solutions présentées à propos de ce problème l'utilisation de certaines technologies afin d'améliorer les plateformes ou les applications d'apprentissage en ligne. Dans ce qui suit nous allons présenter trois technologies différentes qui vont être utilisées par la suite dans la contribution afin de délivrer une solution à ce problème.

1.2 Sensibilité au contexte

Les systèmes informatiques traditionnels sont conçus comme une ou plusieurs boîtes noires, qui réagissent à des entrées et produisent des sorties Figure 1.1 ce type de communication homme-machine introduit souvent des déceptions du côté utilisateur, ce qui rend cette communication très complexe. Toutefois le comportement prédictif qui prend en charge les particularités des utilisateurs et l'un des plus importantes forces de l'informatique d'aujourd'hui. Au début de l'apparition des interfaces graphiques dès les années



FIGURE 1.1: Application traditionnelle (sans contexte)

80 la prise en compte du contexte d'utilisation a été présentée discrètement dans quelques logiciels de bureautique où l'utilisateur par exemple pouvait enregistrer des préférences ou définir des modèles de documents. Etant donné que Le comportement d'un système doit être prédictif, quel que soit son cadre d'utilisation (Lieberman et al., 2000), ce domaine a connu une innovation rapide coté système et matériel.

L'utilisation du contexte sert à améliorer les interactions homme-machine basant sur les caractéristique particulières de l'utilisateur d'un côté et aussi sur l'environnement géographique, le temps, le comportement et d'autres variables grâce aux différents capteurs des appareils qui viennent compléter le système existant.

1.2.1 Définition du contexte

Il existe plusieurs définitions du terme contexte chaque équipe de recherche propose une définition différente des autres mais en globales la plupart des définitions mettent l'accent sur certains caractéristiques communes. Commenant par la définition de (Weiser,1993) qui présente le contexte comme « l'ensemble des informations à prendre en compte en vue d'une adaptation », (Dey et al., 2000) dans leurs définitions introduisent la notion d'entité

afin d'éviter de citer toutes les variables contextuelles, pour eux « le contexte couvre toutes les informations pouvant être utilisées pour caractériser la situation d'une entité ». L'entité est une personne, un lieu, ou un objet qui peut être pertinent pour l'interaction entre l'utilisateur et l'application, y compris l'utilisateur et l'application eux-mêmes. Une autre définition présentée par (Lemlouma, 2004) base sur la notion de l'environnement « le contexte est l'ensemble de toutes les informations de l'environnement qui peuvent influencer le processus de l'adaptation et de la transmission du contenu vers l'utilisateur final ». D'autre part (Chaari et al., 2005) définissent le contexte comme « l'ensemble des paramètres externes à l'application, pouvant influencer sur le comportement d'une application en définissant de nouvelles vues sur ses données et ses services ». (Miraoui et al., 2008) résumant la définition du contexte comme « toute information qui déclenche un service, ou modifie la qualité d'un service si sa valeur change ».

Les différentes définitions permettent de classer instantanément ce qui s'agit du contexte et ce qui ne s'agit pas. Si un élément est susceptible de modifier le résultat produit par un système alors, cet élément fait partie du contexte.

1.2.2 Catégorisation de contexte

Pour mieux illustrer la définition du contexte il faut mettre l'accent sur la catégorisation des variables de ce dernier, en classifiant par catégorie les différentes informations contextuelles. Dans cette section, nous présentons les classifications proposées par quelques recherches.

(Schilit et Theimer, 1994), proposent une classification en trois sous classes, dans lesquelles chaque variable répond à l'une des questions : "*Où suis-je ?*", "*Avec qui suis-je ?*", et "*Quelles sont les ressources de mon environnement proche ?*". Une catégorisation en quatre classes a été présentée par (Ryan et al., 1997), ils ont défini les catégories comme suit :

- l'identité de l'utilisateur
- ressources de l'environnement proche
- localisation de l'utilisateur

- la période temporelle d'exécution de l'interaction

(Razzaque, Dobson et al., 2006) ont proposé une autre catégorisation qui distingue six classes comme suit :

- Contexte utilisateur : présente les informations des utilisateurs.
- Contexte physique : présente les informations liées à l'environnement physique.
- Contexte du réseau : tous types d'informations concernant le réseau informatique.
- Contexte d'activité : présente la liste des événements et des interactions.
- Contexte matériel : il identifie les équipements et les appareils utiles de l'environnement.
- Contexte de service : présente les informations relatives aux fonctionnalités du système.

Tandis que (Baldauf, Dustdar et al., 2007) ont présenté deux catégories :

- le contexte physique mesuré par tous types de capteurs physiques.
- le contexte logique qui a son tour contient les informations sur l'interaction.

(Ameyed et al., 2015) adoptent pour une approche basée sur le service, et proposent une catégorisation en deux classes :

- Les informations de déclenchement du service.
- Les informations de changement de forme du service.

Les différentes catégorisations présentées dans cette section semblent complètes et couvrent tous les aspects du contexte, mais cela ne reste que des propositions parmi d'autres, les classifications peuvent être renouvelées selon les caractéristiques des informations contextuelles travaillées, mais cette étape reste très importante pour manipuler efficacement les informations contextuelles et aboutir à un service d'adaptation de qualité.

1.2.3 Systèmes sensibles aux contextes

On rencontre beaucoup de synonymes du terme sensible au contexte, citons par exemple les systèmes adaptatifs, réactifs, ou dirigés par l'environnement, toutes ces dénominations

sont voisines de la notion de sensibilité au contexte (Coutaz, 2005). Un système est vu comme sensible au contexte s'il est capable d'adapter ses résultats par rapport au contexte d'utilisation Figure 1.2. La première définition de la sensibilité au contexte proposé par

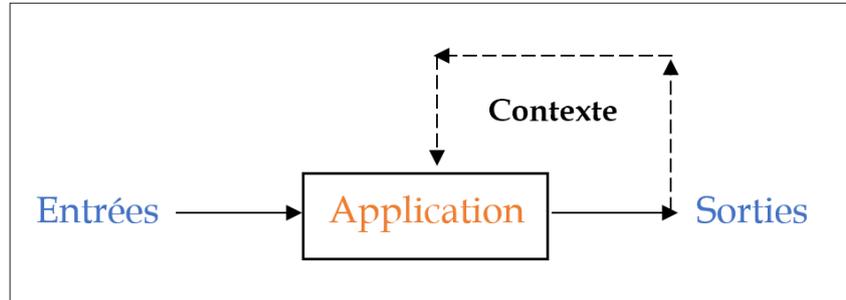


FIGURE 1.2: Application sensible au contexte

(Schilit et Theimer, 1994) définissent cette notion comme étant l'aptitude d'une application à s'ajuster et s'adapter au contexte de son exécution. Depuis, beaucoup d'enrichissements ont été apportés à cette définition. Pour (Salber et al., 1998) l'utilisation du contexte ne signifie pas forcément d'apporter des modifications dans le comportement du système, pour eux la sensibilité au contexte se résume dans la capacité maximale d'agissement en temps réel d'un système avec les différentes données provenant du contexte. (Dey et al., 2000) proposent qu'un système est sensible au contexte s'il délivre des services ou des informations utiles à l'utilisateur selon sa tâche, en utilisant des informations provenant du contexte, pour eux le système peut être désigné comme sensible au contexte même s'il ne subit aucune modification dans son comportement. Pour (Abbas et al., 2007) les systèmes sensibles au contexte fournissent aux utilisateurs des informations adaptées à leurs besoins en suivant leur contexte, ce qui conduit vers des changements dans le système. (Miraoui, 2009) base dans la définition des systèmes sensibles au contexte sur la capacité de déclencher ou de changer un service, pour lui si un système est capable de changer automatiquement les formes de ces services, ou de déclencher un service à cause d'un changement dans les valeurs des informations qui caractérisent le service, alors ce système est dit sensible au contexte.

Donc, un système est sensible au contexte s'il est capable d'interpréter les informations

provenant des différents capteur de l'appareil (le contexte), les présenter dans la mémoire, et les traiter correctement selon le besoin en produisant un changement adéquat qui satisfait l'utilisateur.

1.2.4 Architecture générale des systèmes sensibles aux contextes

Pour bien cerner la fonctionnalité d'un système sensible au contexte (Soualah-alila, 2016) présentent une architecture minimale de ces systèmes. L'architecture proposée par cette recherche, s'articule autour de trois modèles (Acquisition, Agrégation et stockage, et Traitement) comme montre la Figure 1.3

- Modèle d'acquisition du contexte : Le but essentiel de ce module est de capturé les informations disponibles dans le contexte d'un utilisateur particulier, ces informations provenant de deux types de modules l'un physique collecte les informations à partir de capteurs matériels, et l'autre logique fournit les informations contextuelles à partir des différents services logiques.
- Modèle d'agrégation et stockage du contexte : le but de ce modèle est d'interpréter les informations contextuelles fournies par les modules d'acquisition, c'est dans cette étape que ces informations subissent une analyse afin de les transformer dans un format manipulable par le système.
- Modèle de traitement : Au niveau de cette étape le contexte est exploité afin de fournir l'application sensible au contexte, via un ensemble de mécanismes d'adaptation fournis aux utilisateurs des changements adéquats selon le contexte. Pour bien comprendre l'ensemble de mécanismes nous présentons dans le chapitre suivant une étude de quelques architectures dans le domaine des systèmes éducatifs sensibles au contexte.

L'utilisation des plateformes éducatives sensibles au contexte augmente de plus en plus la satisfaction des apprenants, dans cette section nous avons présenté les différentes définitions du contexte et les catégorisations de ce dernier qui le résume comme étant un ensemble d'informations qui concernent l'utilisateur, ses activités et son environnement, l'exploitation du contexte dans ces systèmes sensibles facilite les tâches d'utilisateur.

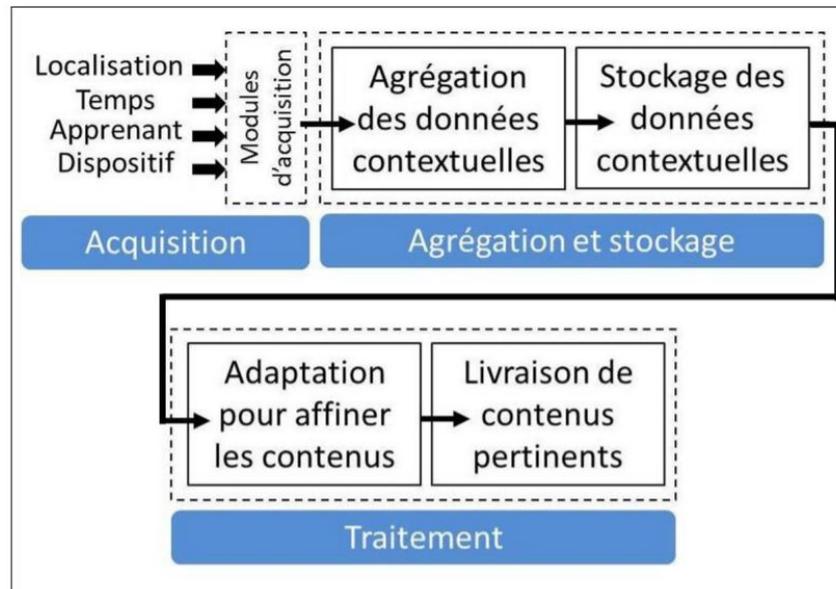


FIGURE 1.3: Architecture minimale d'un système sensible au contexte (Soualah-alila, 2016)

Dans la section suivante nous allons présenter les Cloud computing mobile qui présentent l'une des technologies qui peut servir à améliorer et à développer les architectures et les plateformes d'apprentissage en ligne.

1.3 Mobile Cloud Computing

Cloud computing Mobile (MCC), a été introduit comme une nouvelle technologie puissante pour les services mobile, MCC présente la disponibilité des services Cloud dans l'environnement mobile, de façon que les utilisateurs mobile reçoivent des services plus puissants surtout au niveau de calcul et de stockage, toutes les modules complexes peuvent donc être traitées dans les Cloud ce qui détache les appareils mobiles de la configuration complexe.

D'une part les Cloud computing mobile qui présentent à la base une combinaison de trois domaines (réseau sans fil, informatique mobile et Cloud computing) Figure 1.4, apportent des avantages et des qualités puissants et importants pour les utilisateurs mobile, cette combinaison mène vers une nouvelle infrastructure où les tâches complexes se produisent en dehors de l'appareil mobile (Perez, 2009). D'autre part elle introduit de nombreux problèmes comme la sécurité de communication (Shamim et al., 2015). Nous pouvons

citer plusieurs exemples de Cloud computing mobile qui ont beaucoup amélioré les services fournis aux utilisateurs comme :

- Google Maps : est un service de cartographie en ligne.
- Gmail : est un service de messagerie électronique, gratuit pour les particuliers, proposé par Google.
- Dropbox : est un service de stockage et de partage de copies de fichiers locaux en ligne.

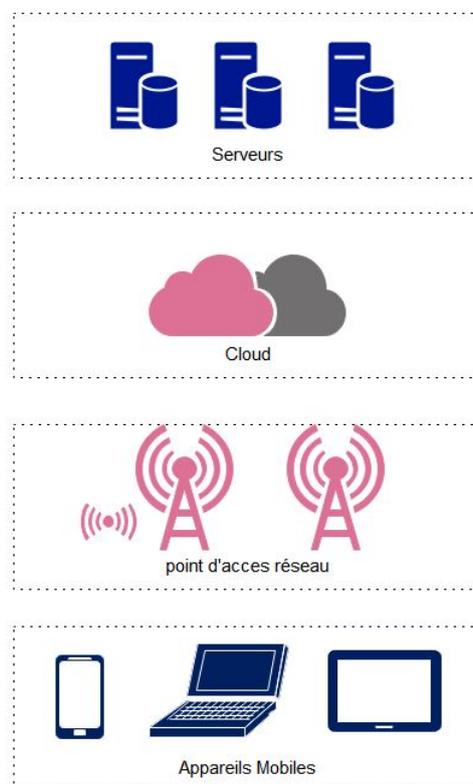


FIGURE 1.4: Mobile Cloud Computing

1.3.1 Définition de Mobile Cloud Computing

Afin de clarifier et d'illustrer la définition de mobile Cloud computing, nous allons d'abord définir les trois concepts qui le constituent.

Informatique Mobile (mobile computing) :

C'est un système dans lequel un ordinateur doit être transporté permet donc la transmission de données, grâce à ces systèmes on peut utiliser l'informatique même en étant

mobile, la portabilité est l'un des aspects les plus importants des systèmes mobile (Bezboruah, 2011). Les avantages des appareils et des applications mobiles sont illimités et les innovations se diffusent très rapidement et efficacement. Dès les années 90 différents types d'appareils mobile ont été introduits ; consoles de jeu portables, assistants personnels (PDA), les lecteurs MP3 , les tablettes, les smartphones etc. Ces appareils apportent beaucoup d'avantages et ajoutent une qualité importante à la vie des utilisateurs, mais ils ont des ressources limitées (la durée de vie de la batterie, le stockage et la bande passante) ce qui affecte le niveau de service (calcul et stockage, etc.), les ressources limitées handicapent considérablement la qualité des services offerts par l'informatique mobiles. Pour surmonter ce problème beaucoup de recherches se dirigent vers l'utilisation des services Cloud.

Cloud Computing :

La technologie de Cloud computing a connu beaucoup de définitions qui se basent sur les services et les caractéristiques de cette dernière. La définition la plus populaire reste celle de NIST (National Institute of Standards and Technology) (Mell et Grance, 2011) « Le Cloud computing est un modèle permettant un accès réseau ubiquitaire et pratique à la demande à un pool partagé de ressources informatiques configurables (par exemple, réseaux, serveurs, stockage, applications et services) qui peuvent être rapidement provisionnées et publiées avec un effort de gestion minimal, ou interaction du fournisseur de service ». Une autre définition proposée par (Paraiso, 2014) « une solution de virtualisation permettant aux entreprises de toutes tailles de se procurer et d'utiliser un large éventail des ressources informatiques à la demande, sur la base du principe payez uniquement ce que vous consommez, accessible de n'importe où, sur différents supports (ordinateurs, smartphones, tablettes, etc.) et à tout moment».

Réseau sans fil :

C'est un réseau dans lequel au moins deux terminaux communiquent sans une liaison filaire, ce qui ajoute la notion de mobilité, les utilisateurs ont la possibilité de se déplacer tout en restant connecté dans un espace plus au moins étendu. Il existe plusieurs défini-

tions pour le Mobile Cloud computing dans la littérature, dans cette section nous allons représenter quelques définitions afin de déterminer une description commune.

- MCC fait référence à une infrastructure où le stockage et le traitement des données se produisent en dehors de l'appareil mobile (dans le Cloud), apportant non seulement aux utilisateurs de smartphone mais à une gamme beaucoup plus large d'abonnés mobiles (MCC-forum, 2011).
- C'est l'intégration de la technologie de Cloud computing avec les appareils mobiles, pour rendre ces derniers pleins de ressources en termes de puissance de calcul, de mémoire, de stockage, d'énergie et de conscience du contexte (Rehman Khan, 2014).
- Aepona définit le MCC comme un nouveau paradigme dans lequel le traitement et le stockage des données sont déplacés de l'appareil mobile, vers des plateformes informatiques situées dans les Cloud, accessibles via la connexion sans fil.
- Pour IBM Le Cloud computing mobile utilise le Cloud computing pour distribuer des applications à des appareils mobiles. Les applications mobiles peuvent être déployées à distance rapidement, de façon souple et avec des outils de développement. La création et la révision des applications de Cloud mobile peuvent se faire rapidement à l'aide de services Cloud. Les applications peuvent être distribuées à de nombreux appareils différents ayant des systèmes d'exploitation, des tâches et des mémoires différents. Les utilisateurs ont ainsi accès à des applications qui ne pourraient pas être prises en charge autrement.

D'après l'organisme (MCC-forums, 2011) qui ont présenté la définition de mobile Cloud computing comme une infrastructure qui déplace la puissance de calcul et de stockage vers le Cloud qui permet de dilater le nombre des abonnés mobile. D'autre auteur comme (rahmen et al., 2014) mettent l'accent sur d'autres ressources mobiles qui vont être développés d'avantage comme l'énergie et la sensibilité au contexte, Aepona considèrent de plus que l'accessibilité au service de MCC se fait via les applications mobiles ou les navigateurs web sur l'appareil.

En tenant compte des caractéristiques communes aux définitions précédentes, nous proposons une définition de mobile Cloud computing comme suit : «Le Mobile Cloud Compu-

ting est une infrastructure qui permet de déplacer les traitements de calcul et le stockage des données de l'appareil mobile vers le Cloud via un réseau sans fil. Ce qui donne aux utilisateurs mobile la possibilité d'utiliser une large gamme de ressources informatiques, accessibles n'importe où à tout moment sur tous les appareils mobiles qui permettent par conséquent de faciliter l'utilisation de l'appareil et dilater le nombre d'abonnés mobile ».

1.3.2 Caractéristiques de Mobile Cloud Computing

Par les définitions et dans la littérature nous pouvons extraire beaucoup de caractéristiques et d'avantages de Mobile Cloud computing cette section présente les caractéristiques et les avantages majeurs des MCC. Selon (Kamalrulnizam et al., 2013; Tawalbeh et al., 2015) et les définitions précédentes, Mobile Cloud computing a plusieurs caractéristiques majeurs citons :

- Service à la demande : l'utilisateur final peut avoir des capacités informatiques selon le besoin.
- Accès au réseau à large bande : les capacités sont disponibles sur le réseau et accessibles via des mécanismes standards par des plates-formes clientes légères ou épaisses. Le temps de réponse de l'utilisateur final est pris en charge.
- Mise en commun des ressources : Afin de servir plusieurs utilisateur les ressources informatiques telle que (mémoire, bande passante, ... etc.), sont mises en commun à l'aide d'un modèle multi-locataire.
- Services mesurés : grâce à la capacité de mesure, l'utilisation des ressources sont contrôlées et optimisées automatiquement par les systèmes Cloud, en tirant parti d'une capacité de mesure à un certain niveau d'abstraction approprié au type de service.
- Évolutivité : Les applications mobiles peuvent être agrandies et réduites pour répondre aux demandes imprévisibles des utilisateurs. En revanche, le fournisseur de services peut facilement faire évoluer et étendre les services.

1.3.3 Architecture de Mobile Cloud Computing

(Suhas, 2015) déclare que la bonne architecture doit représenter souvent la coordination entre l'informatique et le commerce. Beaucoup d'architectures MCC sont présentées dans la littérature et la plupart d'entre elles prennent en considération le coût de déploiement et le marché. D'abord on commence par l'architecture de base présentée par (Shamim et al.,2015), comme montre la Figure 1.5, les utilisateurs via leurs appareils mobiles bénéficient des services de Cloud via internet, premièrement les appareils mobiles (smartphone, tablette . . .), transmettent les informations à l'aide de station de base, point accès ou satellites aux unités centrales de traitement, ici chaque opérateur de réseaux mobile fournit des services précieux à ces utilisateurs (autorisation, authentification et compatibilité etc. . .). Après cela, les demandes des abonnés au Cloud seront transmises via Internet et les contrôleurs du Cloud traitent les demandes pour accéder aux services correspondants au besoin de l'abonné mobile. Dans la même recherche Shamim, Bahar et

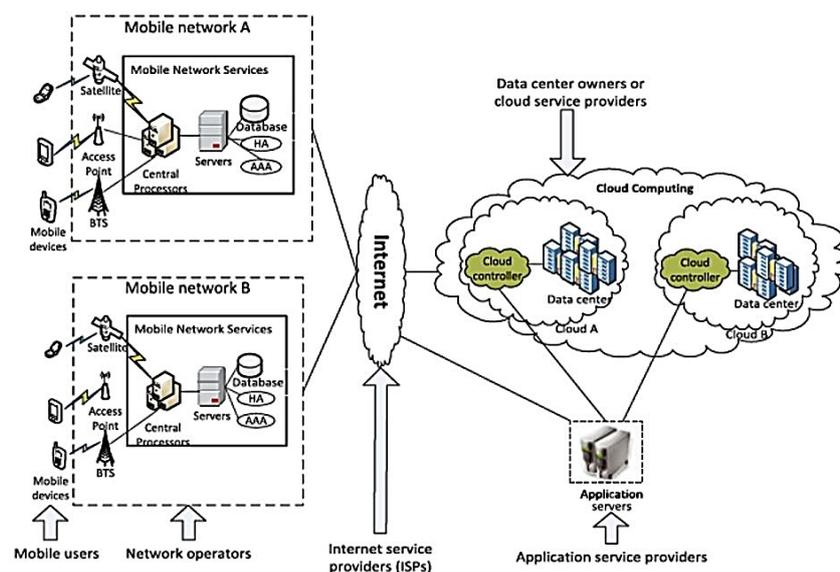


FIGURE 1.5: Architecture Mobile Cloud Computing (Shamim et al.,2015)

Sarker ajoutent qu'il y'a quatre modèles de déploiement identifiés pour les architectures Cloud :

- Cloud privé : l'architecture de Cloud est exploitée pour une organisation privée et gérée par un tiers.

- Cloud communautaire : diverses organisations participent à l'architecture Cloud qui prend en charge une communauté particulière avec des préoccupations communes, différentes organisations ou un tiers peuvent gérer le Cloud communautaire.
- Cloud public : l'architecture Cloud est mise à la disposition du grand public ou d'un groupe industriel.
- Cloud hybride : l'architecture du Cloud hybride est une composition de deux ou plusieurs Cloud qui restent des entités uniques, mais qui sont liés par une technologie standardisée ou propriétaire.

1.3.4 Avantages et limites de Mobile Cloud Computing

Les avantages de Mobile Cloud computing :

Dans la section-1 de ce chapitre nous avons présenté quelques inconvénients de mobile Learning, ces inconvénients ou limites de M-learning trouvent une forte réponse lors d'utilisation de services MCC, dans ce contexte, nous citons par exemple (Tawalbeh et al., 2015) :

- Amélioration de la fiabilité à l'aide de Mobile Cloud Computing : En exécutant les applications sur des Cloud nous obtenons une fiabilité élevée car les données sont sauvegardées sur plusieurs serveurs.
- Réduire la consommation d'énergie et prolonger la durée de vie de la batterie : le principe de MCC est que les calculs complexes passent au niveau des serveurs Cloud non pas aux appareils à source limitée. En conséquence, le temps d'exécution sera raccourci et la consommation d'énergie sera réduite.
- Sécurité et confidentialité : Les problèmes de sécurité de mobile Cloud computing sont soit de l'utilisateur mobile (les différentes menaces de sécurités), soit des données situées au niveau de serveur Cloud, et dans des cas rares nous pouvons faire face à d'autres risques de sécurité dus à des erreurs techniques.
- La mobilité : le Mobile Cloud computing permet aux utilisateurs de se déplacer d'un endroit à un autre, sans faire attention qu'ils ont changé de Cloud lors de changement de zone de couverture, l'appareil mobile doit garder une trace des systèmes Cloud

précédents et actuels qui le gèrent, afin de garantir la communication entre Cloud précédente et actuelle.

Les limites de Mobile Cloud Computing :

Malgré les grands avantages de Mobile Cloud computing quelques limites restent un sujet de challenge dans ce domaine nous citons les limites selon (Miettinen, 2010 ; Khokhar, 2012 ; Buyya, 2014 ; Suhas ,2015) :

- Facturation MCC : Le système de factorisation dans MCC, influencé par la mobilité des utilisateurs, plusieurs paramètres doivent être pris en compte pour créer le système de factorisation approprié pour chaque utilisateur mobile (le délai de rétablissement de session, la capacité de bande passante et la sécurité ...etc).
- Longue latence WAN : La latence a des effets sceptiques sur l'efficacité énergétique et la réponse interactive des applications Cloud mobiles car elle consomme plus de ressources mobiles et retarde les transmissions. Dans la connexion mobile, la distance entre les différentes bandes passantes de la station principale et la vitesse des différentes technologies de réseau affectera l'efficacité énergétique et la convivialité des appareils MCC.
- Connectivité transparente : Les réseaux sans fil sont plus faibles en bande passante par rapport aux réseaux câblés, ce qui mène vers un problème d'installation et de maintenance de connexion transparente entre le mobile Cloud Computing et l'utilisateur mobile, de plus la consommation excessive de ressources mobiles.

L'utilisation des services Cloud Computing mobile dans les environnements d'apprentissage en ligne permet aux utilisateurs de recevoir des services plus puissants surtout au niveau de calcul et de stockage, ce qui aide à augmenter la satisfaction des apprenants, dans cette section nous avons présenté les différentes définitions du MCC, ses caractéristiques, l'architecture minimale et les avantages et les limites de MCC.

Dans la section suivante nous allons présenter une introduction au Deep Learning, qui présente l'une des techniques qui peut servir à l'amélioration des architectures et des plateformes d'apprentissage en ligne.

1.4 Introduction au deep learning

Le Deep Learning est une classe d'algorithmes d'apprentissage automatique, représente un domaine de l'intelligence artificielle, dérivée du Machine Learning Figure 1.6, où la machine est capable d'apprendre par elle-même, le DL a été exposé ces dernières années, il a enrichi progressivement notre quotidien, la reconnaissance vocale de l'assistant Google, la description automatique d'images ou même la conception de nouvelles molécules pharmaceutiques, l'innovation de toutes ces applications est grâce aux techniques de Deep Learning.

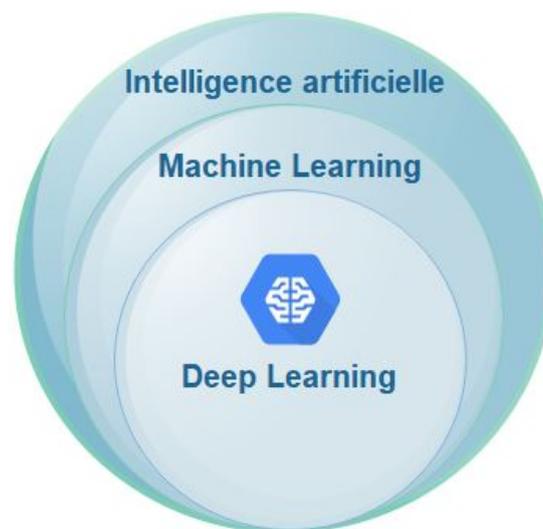


FIGURE 1.6: Positionnement de DL par rapport au AI et ML

Dans cette section nous allons présenter le positionnement de Deep Learning par rapport aux autres champs de l'intelligence artificielle, et ces angles qui nous intéressent dans notre contribution.

1.4.1 Intelligence Artificielle (IA)

L'intelligence artificielle est devenue un sujet très populaire dans les journaux scientifiques les conférences et partout dans les médias, en raison du progrès réalisé par cette dernière. La richesse de ce domaine d'apprentissage automatique attire l'investissement de grandes entreprises dont Microsoft, Google, Facebook et plusieurs autres ce qui ajoute encore un développement impressionnant dans le traitement de la parole, compréhension du langage naturel, la reconnaissance d'expression faciale, traitements d'images etc.

Définition de l'intelligence artificielle

Un mathématicien qui s'appelle Alen Turing a présenté une nouvelle notion « Intelligence Artificielle » dans son livre *Computing Machinery and Intelligence* (Turing, 1950), l'idée c'était de mesurer l'intelligence de la machine via un test « Test de Turing », si les réponses de l'ordinateur sont indifférencies de celles de l'humain l'ordinateur est dit artificiellement intelligent. Donc si une machine peut donner des réponses comme celle de l'humain que ce soit pour la description de l'image la reconnaissance de la parole ou même pour un jeu d'échecs, on parle donc d'une machine doté d'un système intelligent. Une définition claire et nette pour l'intelligence artificielle est difficile, (Kaplan et Haenlein, 2019a) présentent trois raisons pour expliquer l'ambiguïté du terme. Premièrement : il est d'abord difficile de définir l'intelligence humaine, deuxièmement, quand on s'habitue que la machine est capable de faire des tâches complexes on ne peut pas les définir comme une intelligence de la machine, troisièmement, l'effet évolutif de l'IA conduit à mélanger ces différentes étapes et types, ce qui conduit à une confusion sur le terme lui-même. Pourtant plusieurs définitions de l'intelligence artificielle ont été proposées. Dans l'encyclopédie LAROUSSE l'IA est définit comme « Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine ». Selon (Kaplan et Haenlein, 2019b) l'intelligence artificielle est la capacité d'un système à interpréter correctement les données externes, à apprendre ces données et à utiliser ces apprentissages pour atteindre des objectifs et des tâches spécifiques, grâce à une adaptation flexible. France Intelligence artificielle - Rapport de synthèse définit l'IA comme étant

« un usage visant à automatiser les tâches et passant par l'intégration de plusieurs briques technologiques relevant d'un écosystème thématique ». Donc l'intelligence artificielle peut être définie comme une science qui vise à étudier les moyens de créer des programmes et des machines intelligentes, capables de résoudre des problèmes de manière créative comme la compréhension, le raisonnement, le dialogue selon la situation, et l'apprentissage, ce qui a toujours été considéré comme des avantages humains.

Les classes de l'intelligence artificielle

L'IA peut être catégorisée en trois classes, Intelligence artificielle faible (Narrow AI), Intelligence artificielle forte (General AI) et Super Intelligence Artificielle (Super AI) beaucoup de recherches ont entamé ces différentes classes de l'intelligence artificielle telle que (Pennachin Goertzel, 2007; Siau Yang, 2017; D. Jajal, 2018; Kaplan Haenlein, 2019a; Jarek Mazurek, 2019)

a) **Intelligence artificielle faible (ANI) :**

Le rôle de l'IA faible est de traiter et d'identifier les données acquises dans un ou plusieurs domaines pour effectuer certaines tâches, mais elle ne dépassera pas l'homme en dehors de ses capacités prédéfinies, cette classe d'IA est dépourvue de conscience, de motivation et de toutes sortes d'émotions elle fonctionne que dans une plage prédéfinie. Google assistant, Siri, Cortana, et d'autre outils d'intelligence artificielle, représentent une IA faible car ces outils ne peuvent pas penser par eux-mêmes, ce que ces outils font c'est de traiter le langage humain ensuite, envoyer des résultats adéquats après une recherche préalable, ce qui explique que parfois l'assistant Google par exemple nous répond par des liens vers des articles sur internet pour répondre à des questions abstraites.

b) **Intelligence artificielle forte (AGI) :**

La portée de la capacité intellectuelle d'Intelligence artificielle forte est comparable à celle du cerveau humain, donc ce sont des machines intelligentes qui peuvent accomplir toutes les tâches possibles qu'un être humain peut accomplir. L'objectif initial de l'AGI est de créer des machines conscientes, sensibles et motivées par

l'émotion. Une machine dotée de l'intelligence générale artificielle, peut raisonner comme un être humain, se souvient de tous les souvenirs et leurs sentiments associés pour pouvoir créer des décisions se basant sur les expériences du passé, de plus il faut intégrer la conscience pour pouvoir créer des jugements et suggérer des solutions créatives, ce qui est très difficile à produire.

c) **Super Intelligence Artificielle (ASI) :**

La Super Intelligence Artificielle dépasse largement les performances cognitives des humains dans pratiquement tous les domaines, la machine devient Intelligente, sage, créative, dotée de compétences sociales et qui peut résoudre tous les problèmes d'une manière excellente. Bien que l'ASI n'existe pas actuellement, certains chercheurs proposent qu'il puisse être créé dans le futur, d'autres s'inquiètent que ce type de pouvoir peut détruire l'humanité pourtant que son objectif majeur est d'améliorer la vie des humains.

L'intelligence artificielle faible, présente une portée limitée d'imitation d'intelligence humaine, elle offre des bons résultats dans différents domaines comme le traitement de langage humain, la reconnaissance des formes etc. l'intelligence artificielle forte couvre plus de domaines, comme le pouvoir de raisonner et de penser abstraitement. La super intelligence artificielle est une explosion de l'intelligence de la machine qui dépasse l'intelligence humaine dans tous les domaines.

Machine Learning est l'une des technologies de l'IA qui intéresse les scientifiques et les chercheurs actuellement, cette technologie peut aider à créer une intelligence forte et une super intelligence, vu leurs résultats impressionnants.

1.4.2 Machine Learning (ML)

L'observation du timeline du Machine Learning nous prend dans une histoire chronologique assez intéressante. Tout commence en 1959, un informaticien américain appelé Arthur Samuel, considéré comme un pionnier dans le domaine des jeux informatiques et de l'intelligence artificielle utilise pour la première fois le terme «machine learning », pour son programme créé en 1952 capable de jouer aux dames (Moser, 1990). Frank Rosen-

blatt, crée le perceptron qui est un classifieur binaire considéré comme un simple réseau de neurones artificiels, cette invention a connu une énorme couverture par les médias, ce qui a donné l'impression que la communauté scientifique va arriver très vite à créer des innovations assez incroyables. Puis cette communauté a connu une période de reculement à cause du manque de résultats, où ils se sont rendu compte qu'ils ont sous-estimé les difficultés du domaine. Vers les années 90 avec l'augmentation de la puissance de calcul des ordinateurs et l'essor d'internet qui a permis aux chercheurs de travailler plus efficacement et de partager leurs idées et résultats et surtout d'accéder à un nombre important de data, les Machine learning ont connu une forte renaissance. Des innovations et des réussites successives sont apparues jusqu'à nos jours comme : IBM's Deep Blue en 1997 (Campbell et al., 2002) qui ont battu le champion du monde aux jeux d'échecs ; la création d'ImageNet (J. Deng et al., 2009) qui représentent une massive base de données visuelles ; ou encore DeepFace (Taigman et al., 2014), qui représentent un système qui utilise des réseaux de neurones pour identifier les visages avec une précision de 97,35%.

Définition des Machine Learning

Machine Learning est une technologie de l'intelligence artificielle, où les machines ont la capacité d'apprendre et de s'améliorer automatiquement en se basant sur leurs expériences sans être explicitement programmées. Cela permet à la machine de résoudre des problèmes compliqués où c'est très difficile pour l'humain de créer manuellement les algorithmes nécessaires pour les résoudre, dans ce cas-là il est plus pratique que la machine crée ses propres algorithmes par elle-même, en se basant sur les données d'apprentissage fournies au préalable.

(El Naqa et Murphy, 2015) définissent Machine Learning comme une branche évolutive d'algorithmes de calcul, conçue pour émuler l'intelligence humaine en apprenant de l'environnement. Les Machine Learning ont également été définies par (Murdoch et al., 2019) comme une technologie d'intelligence artificielle avec une capacité à prédire avec précision une grande variété de phénomènes complexes. Cependant, on se rend de plus en plus compte que, en plus des prédictions, les modèles ML sont capables de produire des

connaissances sur les relations de domaine contenues dans les données, souvent appelées interprétations.

À partir de ces définitions nous allons présenter dans la section suivante les étapes de machine Learning.

1.4.3 Etapes de Machine Learning

Le but principale de la Machine Learning est de donner un sens au data, afin de répondre aux différentes requêtes de l'utilisateur. Pour arriver à ce stade il faut créer un bon model de machine learning, Yufeng Guo un développeur et avocat pour Google Cloud a proposé les sept étapes permettant généralement de créer une machine learning (Yufeng G,2017) dans ce qui suit nous allons discuter ces étapes :

- Etape 01 la collecte de données : La première étape, dans le processus de la création du modèle Machine Learning, commence par la collecte des données. De nos jours selon le domaine on peut trouver des bases de données massives sur internet pour par exemple la reconnaissance du langage ou des émotions faciales, mais il reste toujours possible de créer les données seules. Il faut surtout noter que la qualité et la quantité des données influencent directement la qualité et la fiabilité du modèle finale.
- Etape 02 préparation de données : Après la phase de collecte de données c'est dans cette étape qu'on les prépare pour une utilisation dans le processus de la machine Learning, cette étape permet surtout de vérifier la balance des données, si l'un des caractéristique est dominant dans notre base de donnée cela risque de données de faux résultats dans le modèle finale. Dans cette étape aussi on divise les données en deux parties l'une pour la phase d'apprentissage et l'autre pour la phase d'évaluation.
- Etape 03 le choix du modèle : cette étape consiste à choisir un modèle adéquat selon les données travaillées, l'apprentissage peut être supervisé ou non supervisé basé sur une classification, une régression ou autre, il existe plusieurs critères pour choisir le modèle approprié comme (la quantité et la qualité de données et sa structure de

représentation etc.). Dans ce contexte il existe de nombreux modèles comme :

- Régression linéaire simple : utilise la forme traditionnelle d'interception de pente, où m et b sont les variables que notre algorithme essaiera d'apprendre pour produire les prédictions les plus précises $y = m * x + b$; tel que : x représente les données d'entrée, y représente la prédiction, m est le poids et b représente le biais (ML Glossary documentation, 2017). Généralement, il y en a beaucoup de m car il peut y avoir de nombreuses caractéristiques. La collection de ces m est formée en une matrice, désignée par w (weights), idem pour b on forme la matrice b (Bias).

L'algorithme essaiera d'ajuster les valeurs pour le poids et le biais, afin de se rapprocher de la ligne du meilleur ajustement Figure 1.7

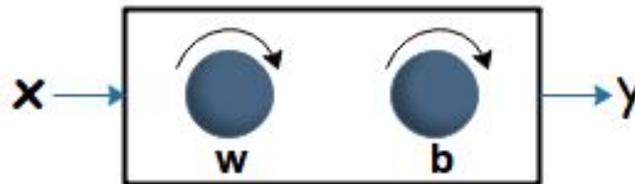


FIGURE 1.7: Le fonctionnement de poids et de biais

Le modèle de régression linéaire simple est souvent utilisé pour des données numériques, mais les machines Learning sont utilisées dans des domaines encore plus compliqués comme la recommandation du fil d'actualité dans les réseaux sociaux, la reconnaissance d'image etc. Et là il faut utiliser des modèles plus complexes comme les réseaux de neurones artificiels.

- Réseaux de neurones artificiels (ANN) : Un ANN est un réseau de neurones artificiels, dont chaque neurone est appelé un nœud. Ces nœuds sont connectés les uns aux autres et pour chaque connexion une valeur (w) est attribué, dans chaque nœud une fonction de transfert est intégrés. Les données circulent dans le réseau de la manière suivante; chaque nœud additionne les valeurs qu'il reçoit des autres nœuds est ensuite ajoute le biais (b) : $\sum_{i=1} (w_i x_i) + b$, ensuite

il modifie la valeur en utilisant sa fonction de transfert : $y=f(I)$, et envoie à son rôle la nouvelle valeur Figure 1.8 jusqu'à ce que le réseau atteigne le nœud final (sortie).

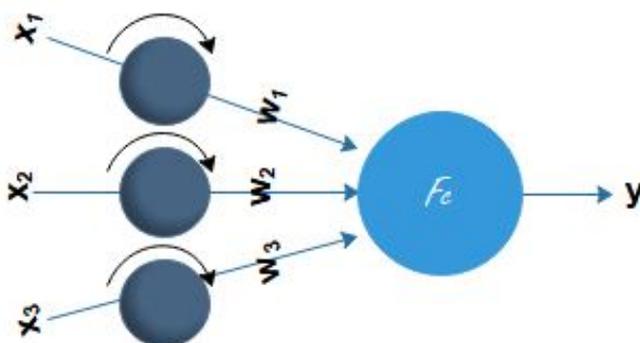


FIGURE 1.8: Fonctionnement d'un neurone artificiels

- Etape 04 la phase d'apprentissage (training) : Cette phase est considérée comme la base du machine Learning, nous utilisons dans cette étape nos données pour améliorer la capacité du modèle ML en ajustant les paramètres W et b jusqu'à trouvé une prédiction plus correcte Figure 1.9.

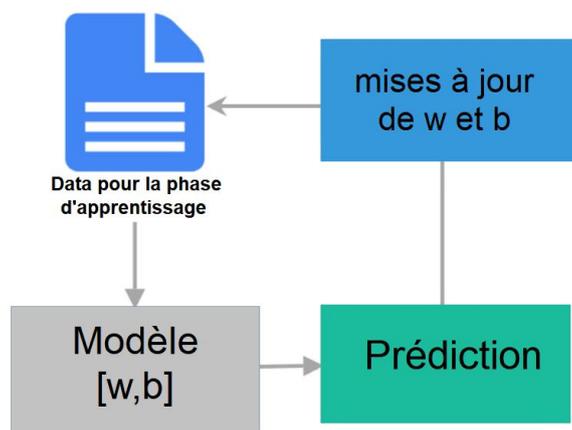


FIGURE 1.9: La phase d'apprentissage (Yufeng G,2017)

- Etape 05 la phase d'évaluation : Dans cette étape on utilise la partie de données réservée pour évaluer la fiabilité du modèle, ce qui nous permet de vérifier la performance du modèle envers les nouvelles données.

- Etape 06 réglage des paramètres : Après la phase d'évaluation il est très important d'ajuster les paramètres pour obtenir des performances optimales (Hutter et al., 2011), c'est dans cette étape que cela se fasse, il existe plusieurs façons de réglages des paramètres selon le modèle choisi et le type des données.
- Etape 07 la prédiction : cette étape et l'étape finale de processus de création du modèle ML, c'est dans cette étape que notre modèle sera capable de répondre aux requêtes des utilisateurs et prédire les résultats Figure 1.10.

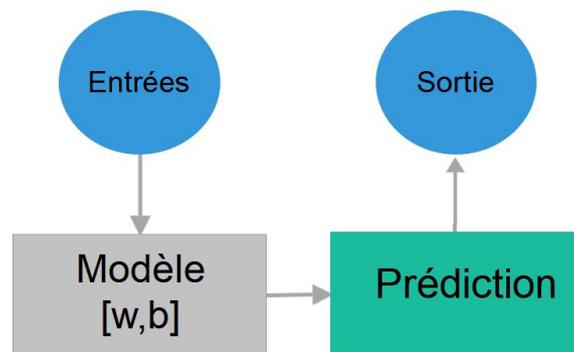


FIGURE 1.10: La phase de prédiction

Le cerveau possède environ 100 milliards de neurones, qui communiquent par des signaux électrochimiques, utiliser le même nombre de neurones artificielles sera très difficile (presque impossible), c'est pour cela que dans les machines Learning on précède l'étape de classification par une étape d'extraction des caractéristiques, qui permet de diminuer d'une manière importante le nombre de nœud d'entrée. Cela représente la différence majeure entre Machine Learning et Deep Learning où dans ce dernier, l'algorithme va découvrir tout seul les caractéristiques, sans aucune intervention, ce qui donne des modèles plus fiables Figure 1.11. Dans la section suivante nous allons présenter le Deep Learning.

1.4.4 Deep Learning (DL)

En 1989 un professeur d'informatique à l'Université de Californie appelé Rina Dechter, a introduit le terme Deep Learning pour la première fois, des années plus tard en 2000, le terme a été utilisé de nouveau par Igor Aizenberg, un professeur en informatique à l'université de Manhattan (Khanna, 2019). Le problème était toujours que ces nouveaux

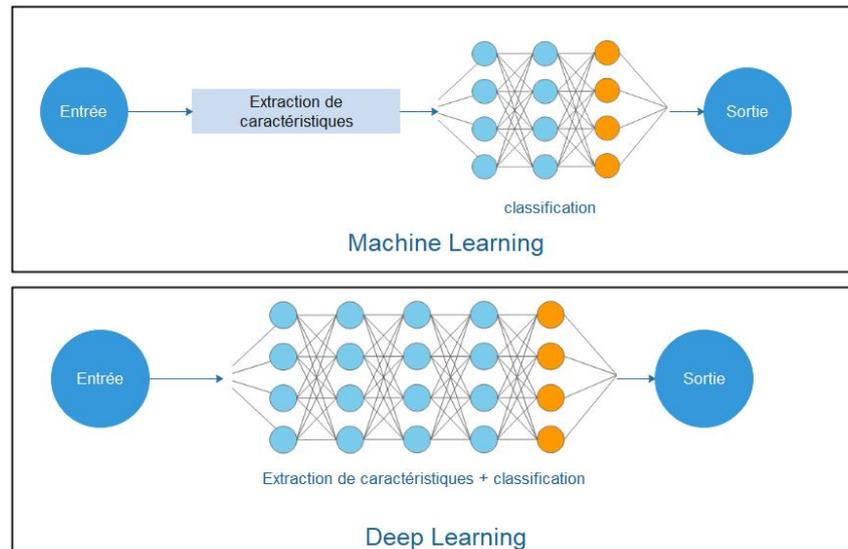


FIGURE 1.11: La différence entre le fonctionnement du ML et DL

algorithmes de réseaux de neurones profonds, nécessitent d'un côté une grande puissance de calcul et de l'autre côté une grande quantité de données. C'est deux obstacles majeurs ralentissent un peu les choses au niveau de la communauté scientifique. Un nombre minime de recherches et de résultats circule dans ce sujet jusqu'à l'année 2012. Lors du challenge d'ImageNet pour la reconnaissance visuelle à grande échelle à l'université de Stanford, le projet vainqueur dans la tâche de classification d'images a battu tout le monde avec un taux d'erreurs de 15.3%, en utilisant un algorithme du deep Learning. Dans les versions suivantes du même challenge la plus part des chercheurs optent pour les algorithmes du deep Learning ce qui affecte d'une manière très positive les résultats obtenus, Table 1.1 présente les résultats de toutes les versions d'ImageNet challenge (première et deuxième place) pour la tâche de classification. Dans cette même année, Google dévoile son projet Google Brain, en utilisant un cluster de 16000 processeurs pour créer un réseau de neurones avec plus d'un milliard de connexions. Ils l'ont ensuite alimenté par des miniatures d'images aléatoires, chacune extraite de millions de vidéos YouTube, la machine s'était entraînée seul avec succès pour reconnaître un chat (Markoff, 2012). Donc, si le réseau est bien entraîné il va créer les caractéristiques par lui-même sans aucune intervention, ce qui permet d'avoir de meilleurs résultats. La progression des algorithmes deep Learning, la puissance de calcul des GPU pour le traitement d'images et surtout la disponibilité

des données, permet au deep Learning d'accélérer brutalement et de dévoiler des résultats extraordinaires dans plusieurs domaines (reconnaissance vocale automatique, reconnaissance d'image, traitement du langage naturel, découverte de médicaments, systèmes de recommandation, voiture autonome, génération d'images etc.).

La compétition	Nom du groupe	taux d'erreur
2010	NEC-UIUC	28%
(ILSVRC,2010)	XRCE	33.6%
2011	XRCE	25.7%
(ILSVRC,2011)	Universities of Amsterdam and Trento	31%
2012	SuperVision	15.3%
(ILSVRC,2012)	ISI	26.1%
2013	Clarifai	11.1%
(ILSVRC,2013)	NUS	12.9%
2014	GoogLeNet	6.6%
(ILSVRC,2014)	VGG	7.3%
2015	MSRAC	3.56%
(ILSVRC,2015)	Reception	3.58%
2016	Trimps-Soushen	2.99%
(ILSVRC,2016)	ResNeXt	3.03%
2017	WMW	2.25%
(ILSVRC,2017)	Trimps-Soushen	2.48%

TABLE 1.1: Résultat d'ImageNet challenge (de 2010 à 2017)

Définition du Deep Learning

Le deep Learning ou l'apprentissage profond est un sous-ensemble particulier des machine Learning, basé sur des réseaux de neurones artificiels avec de nombreuses couches cachées, ce modèle profond a démontré une capacité remarquable par rapport aux autres approches (Antipov, 2017). (Lecun et al., 2015) ajoutent que le deep Learning permet aux modèles de calculs composés de plusieurs couches de traitement d'apprendre des représentations de données avec plusieurs niveaux d'abstraction. Ces méthodes ont considérablement amélioré l'état de l'art pour la reconnaissance vocale, la reconnaissance visuelle d'objets, la détection d'objets, la découverte de médicaments et de nombreux autres domaines.

Les définitions présentées de DL sont pratiquement semblables, nous pouvons les résumer comme suit ; le deep Learning c'est une branche de machine Learning inspirée des neu-

rones du cerveau humain représente une interconnexion de nombreux neurones artificiels dans une structure de centaine de couches, afin d'analyser les différents caractéristiques où chaque couche interprète les données reçues de la couche précédente. Le deep Learning a approuvé des résultats prometteurs dans les différents domaines de l'intelligence artificielle.

Le fonctionnement de deep Learning est similaire au fonctionnement d'ANN, les informations sont transférées d'une couche à une autre via des canaux de connexion. Chaque nœud additionne les valeurs qu'il reçoit multipliées par les poids, puis il active la fonction de transfert, le résultat obtenu détermine si le neurone est activé ou pas. Chaque neurone actif transmet des informations aux couches suivantes. Cela continue jusqu'à l'arrivée à la couche de sortie qui produit un résultat explicite pour l'utilisateur Figure 1.12. Afin de former un deep Learning il faut être doté d'une large quantité de données d'entraînement pour que la solution soit précise.

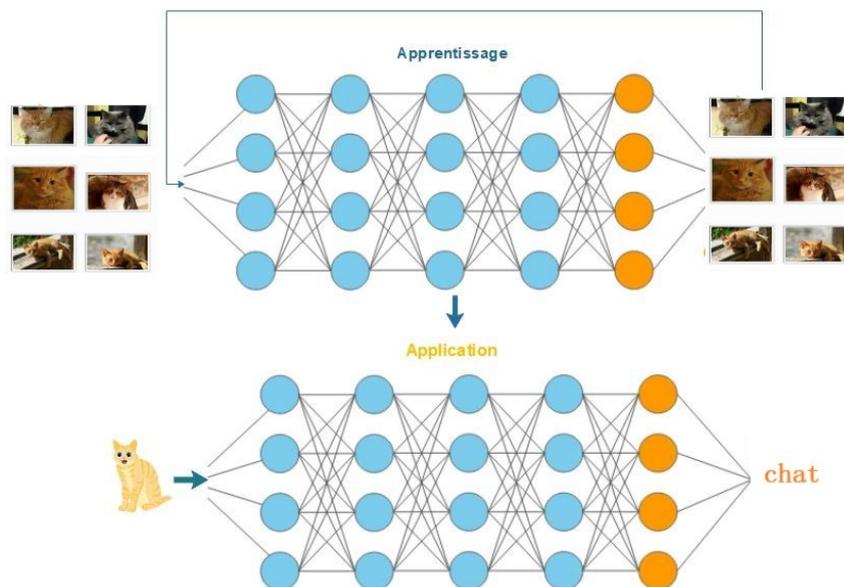


FIGURE 1.12: Processus d'apprentissage (deep Learning) pour identifier un chat

1.4.5 Deep Learning pour l'analyse et le traitement d'images

Le but de l'utilisation de machine Learning c'est de donner aux ordinateurs la capacité d'apprendre sans être programmés auparavant. Les modèles de machine Learning acquièrent des connaissances dans la phase d'apprentissage afin d'effectuer les tâches re-

quises, mais les algorithmes d'apprentissage automatique conventionnel dépendent de la manière dont les caractéristiques sont prétraitées avant d'être fournies à leur entrée. Même pour l'analyse d'images dans le cas de classification ou de détection, les caractéristiques de base sont extraites en utilisant des descripteurs de caractéristiques comme ; Local binary patterns, l'histogramme des gradients orientés, etc. conçue au départ pour un type particulier de problème par des experts du domaine mais par la suite ils ont été utilisés par d'autres domaines. Donc l'utilisation de ce type d'extraction de caractéristiques fortement dépendant de l'expert, est difficile à concevoir d'une part, et n'est pas valide pour tous types de classification d'autre part. Un moyen pour surpasser ce problème consiste à apprendre les représentations des caractéristiques pour chaque problème particulier.

L'utilisation de deep ANN pour la classification d'images, représente une solution très efficace pour ce problème, un ANN profond se compose de plusieurs couches, le système est alimenté en entrée par une image représentée par une matrice de pixels. Ensuite un processus de détection de bords élémentaires (motifs verticaux, horizontaux, coins etc.) est appliqué en utilisant les filtres. Les couches suivantes sont celles responsables d'organisation des bords pour concevoir des parties d'objets, pour qu'à la fin les couches les plus profondes arrivent à déterminer les objets en combinant les parties d'objets reçues des couches précédentes, ce qui permet au modèle de classer l'image (dans notre cas un humain Figure 1.13).

En outre le principe des algorithmes de backpropagation est généré principalement pour les neurones artificiels (Rumelhart et al., 1986), on peut aussi les considérer derrière le succès des ANN. l'un des modèles très efficace dans ce sujet est les réseaux de neurones convolutifs (CNN) principalement conçus pour traiter les images, les vidéos, la reconnaissance d'électromyographie mais ils ont montré de bons résultats même dans d'autres domaines comme le traitement du langage naturel (Grefenstette et al., 2015), la découverte du médicament (Wallach et al., 2015) et autres.

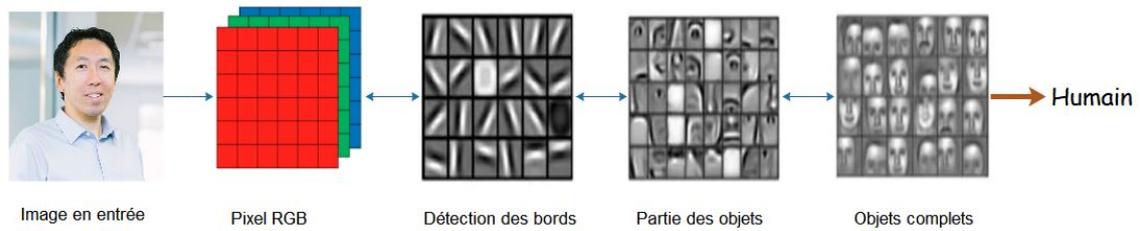


FIGURE 1.13: Présentation schématique du modèle ANN profond pour la classification d'image

1.4.6 Réseaux de neurones convolutifs (CNN)

Les réseaux de neurones convolutifs (CNN), représentent un type de réseau de neurones artificiels inspirés par le cortex visuel humain. (Lecun, Y Bengio,1995) ont utilisé le principe de backpropagation pour créer un apprentissage entièrement automatique où les caractéristiques sont directement extraites à partir des images de nombres écrits à la main par la machine elle-même, ce qui aboutit à un bon fonctionnement permettant de créer une innovation dans le domaine de reconnaissance d'images. Ce qui mène vers une popularité massive de cette approche, juste pendant la période de cette étude de doctorat (2016-2020), la popularité de réseau de neurones convolutifs a augmenté presque par quatre fois (Google Trends), Figure 1.14).



FIGURE 1.14: Révolution d'intérêt pour les réseaux de neurones convolutifs

L'architecture de réseau de neurones convolutifs :

Le principe de CNN qui nous intéresse dans notre contribution c'est la classification des images, chaque image en entrée subit un ensemble de traitements pour déduire sa catégorie approprié. Dans le modèle CNN le traitement d'images d'entrée passe par plusieurs couches (Figure 1.15), on trouve généralement une série de couches de convolution avec des filtres, une couche de regroupement (Pooling), une ou plusieurs couches entièrement connectées fully connected layers (FC) et une couche à la fin pour appliquer la fonction Softmax. Dans ce qui suit nous allons présenter une explication du traitement appliqué dans chacune des couches précédentes.

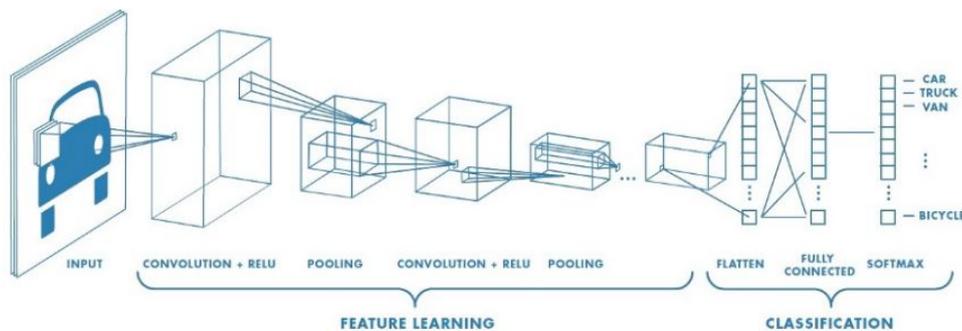


FIGURE 1.15: L'architecture de réseau de neurones convolutifs (Prabhu, 2018)

- La couche de convolution est la couche principale d'un réseau convolutifs, elle fait la plupart des gros travaux de calcul, c'est la première couche à extraire les caractéristiques d'une image d'entrée. Le principe de la couche de convolution s'agit principalement d'une opération mathématique qui prend deux entrées ; la matrice de l'image et le filtre, Une matrice d'image de volume (h, w, d) et un filtre de volume (fh, fw, d) dans comme résultat une matrice de volume $(h-fh+1, w-fw+1, 1)$ Figure 1.16 l'utilisation de plusieurs couches de convolution avec différents filtres dans une image permet de détecter plusieurs types de caractéristiques. En général il existe trois hyper-paramètres qui interviennent dans le traitement de la couche de convolution, ces hyper-paramètres contrôlent fortement la taille du volume de sortie : le nombre de filtres, padding et stride :

- le nombre de filtre : correspond à la profondeur du volume de sortie. Donc

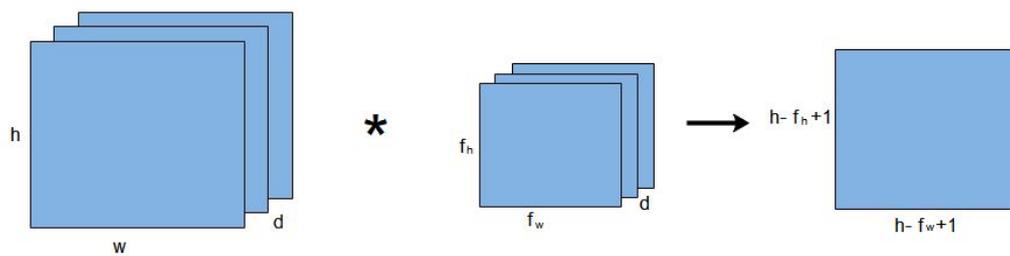


FIGURE 1.16: Volume de matrice d'image, filtre, et la matrice de sortie

lorsque la matrice de l'image est de volume $n \times n$ et le filtre est de volume $f \times f$ et on a utilisé x filtres le volume de la matrice de sortie sera $(n-f+1, n-f+1, x)$

- Padding : lorsque le filtre ne correspond pas parfaitement à l'image d'entrée, nous pouvant contourner l'image avec des zéro ce qu'on appelle (zéro-padding), on peut aussi utiliser cette option lorsque on veut mettre l'accent de plus sur les pixels du coin de l'image (Figure 1.17). Donc lorsque la matrice de l'image est de volume $n \times n$ et le filtre est de volume $f \times f$ et on a utilisé un padding de p (pixels) la matrice de sortie sera du volume $(n+2p-f+1 \times n+2p-f+1)$.

0	0	0	0	0	0
0	35	19	25	6	0
0	13	22	16	53	0
0	4	3	7	10	0
0	9	8	1	3	0
0	0	0	0	0	0

FIGURE 1.17: Application de padding $p=1$

- Stride : Stride ou (foulée) est le nombre de pixels décalés sur la matrice d'entrée. Lorsque le $stride=1$, on déplace les filtres sur 1 pixel à la fois. etc, ce paramètre a pour but de réduire le volume de sortie. Donc lorsque la matrice de l'image est de volume $n \times n$ et le filtre est de volume $f \times f$ et on a utilisé un padding

de p (pixels) et un stride s la matrice de sortie sera du volume :

$$\left\lfloor \frac{n + 2p - f}{s} + 1 \right\rfloor \times \left\lfloor \frac{n + 2p - f}{s} + 1 \right\rfloor$$

La fonction ReLU est généralement appliquée avec la couche de convolution pour introduire la non-linéarité dans le réseau convolutif; $f(x) = \max(0, x)$,

Figure 1.18

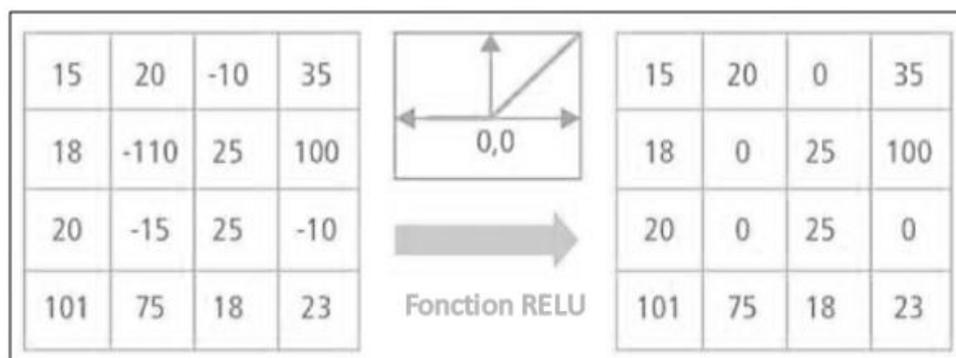


FIGURE 1.18: Fonctionnement de la fonction RELU (Prabhu, 2018)

- Couche de regroupement (Pooling layer) : Les couches de convolution d'un réseau neurones convolutifs, appliquent systématiquement des filtres appris aux images d'entrée afin de créer des matrices de caractéristiques en sortie, ces matrices résumant la présence de ces caractéristiques dans l'image d'entrée, on peut la voir comme une carte d'entité. Mais même avec des petits mouvements dans l'image d'entrée entraîneront une carte d'entités différente. Une des approches courantes pour résoudre ce problème est l'utilisation de la couche pooling (Jason.B, 2019).

la couche pooling permet de réduire le nombre de paramètres lorsque les images sont trop grandes en conservant les informations importantes, y'on a deux types communes de pooling :

- Max pooling : généralement la matrice est décomposée en sous-région ensuite on applique sur chaque sous-région un filtre maximum Figure 1.19.
- Average pooling : généralement la matrice est décomposée en sous-région, ensuite on applique un calcul de la moyenne pour chaque sous-région Figure 1.20.

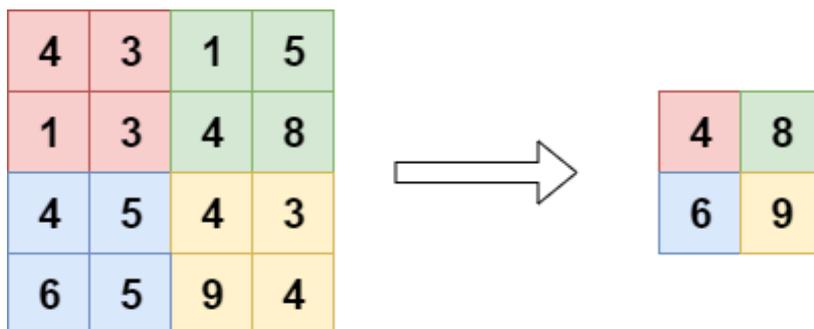


FIGURE 1.19: Fonctionnement de Max pooling

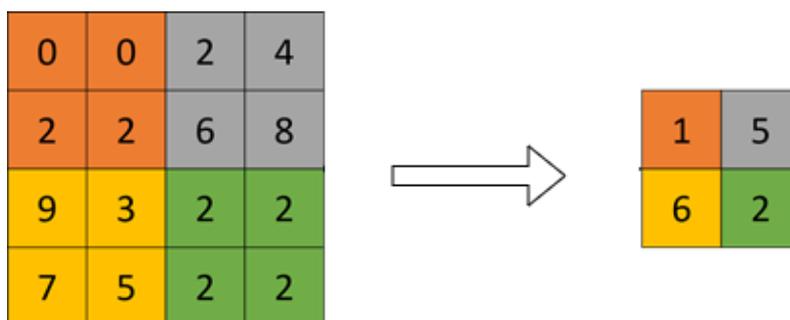


FIGURE 1.20: Fonctionnement d'Average pooling

- Couche entièrement connectée (fully connected layer) : Le processus des réseaux de neurones convolutifs commence par la couche de convolution et la couche pooling, pour extraire les caractéristiques et créer une carte d'entité, le résultat de ce processus va ensuite alimenter la couche entièrement connectée qui a son tour conduit vers la décision de classification finale. Pour aboutir à ce résultat la couche entièrement connectée d'abord transforme la matrice de caractéristiques en un seul vecteur de valeurs, ensuite combine ces caractéristiques pour créer un modèle Figure 2.1, ainsi elle passe par son propre processus de backpropagation pour déterminer les poids les plus précis, enfin elle applique une fonction d'activation tel que softmax ou sigmoïde le réseau fini par classer les sorties en catégories.

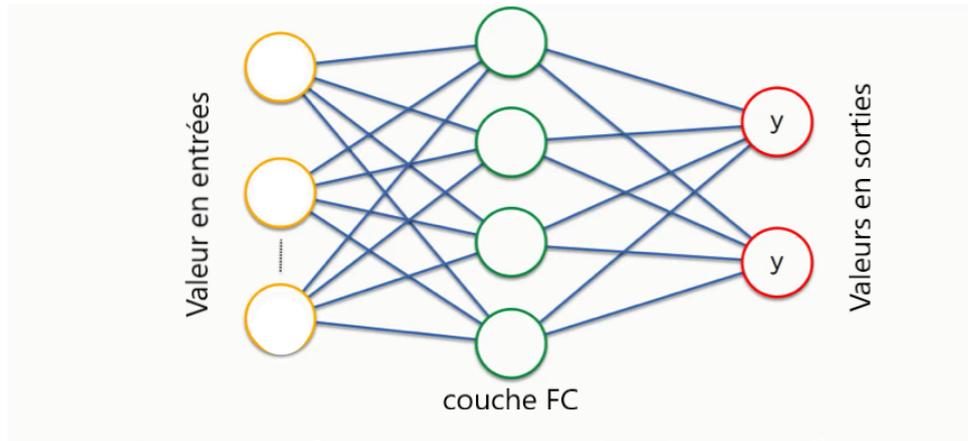


FIGURE 1.21: La couche entièrement connectée

Conclusion

Dans ce chapitre nous avons présenté les types et les caractéristiques des environnements d'apprentissage en ligne, nous avons aussi survolé les technologies principales que nous allons utiliser dans la contribution de l'approche d'adaptation. Dans le chapitre suivant nous présenterons un état d'art sur les travaux déjà présentés dans le domaine d'adaptation des environnements d'apprentissage en ligne et de l'utilisation de réseau de neurones convolutifs pour la classifications des émotions.

Chapitre 2

Apprentissage intelligent auto-adaptatif

Contents

2.1	Apprentissage à distance sensible au contexte et Mobile Cloud Computing	50
2.1.1	Apprentissage à distance	50
2.1.2	Historique de l'apprentissage à distance	51
2.1.3	Apprentissage intelligent	54
2.1.4	Systèmes d'apprentissage à distance sensibles aux contextes . .	55
2.1.5	Travaux connexes sur l'apprentissage auto-adaptatif	56
2.1.6	Apprentissage mobile basé sur le Cloud	61
2.1.7	Systèmes d'apprentissage à distance auto-adaptatif basé MCC .	64
2.2	Apprentissage intelligent et émotions	66
2.2.1	Expressions faciales et unités d'action	67
2.2.2	Types des émotions	69
2.2.3	Deep Learning et expressions faciales	73
2.2.4	Etude sur la reconnaissance des émotions	77
2.2.5	Détection des unités d'actions	80
	Conclusion	82

Ce chapitre va être divisé en deux sections, la première consiste à présenter et à synthétiser la littérature dans les systèmes d'apprentissage en ligne sensibles au contexte. La deuxième présente et analyse les applications et les recherches, qui utilisent les réseaux de neurones convolutifs pour la reconnaissance des émotions faciales, afin d'adapter les systèmes au besoin des utilisateurs et leurs contextes. Pour les deux sections nous allons nous concentrer sur l'avantage d'associer les Cloud computing mobile avec ces systèmes.

2.1 Apprentissage à distance sensible au contexte et Mobile Cloud Computing

Dans les écoles, généralement les élèves se trouvent face à un enseignant et un cours désigné pour tout le monde sans tenir compte de leurs différences, ce qui met les apprenants dans une situation difficile, surtout lorsqu'ils atteignent une phase de rupture d'apprentissage, où la nature des cours imposés la plupart du temps ne peuvent pas les aider. L'apprentissage en ligne est une bonne solution pour ces étudiants, mais que se passe-t-il s'ils font face au même cas lors de leurs études en ligne ? pour éviter ce problème, la meilleure solution est d'adapter les cours et les plateformes d'apprentissage selon les préférences et le contexte de chaque étudiant.

2.1.1 Apprentissage à distance

L'article 26 de la Déclaration universelle des droits de l'homme, stipulait que « Toute personne a droit à l'éducation. L'éducation doit être gratuite, au moins en ce qui concerne l'enseignement élémentaire et fondamental » et que « ... l'éducation doit être également accessible à tous en fonction de son mérite » (Nations Unies, 1948). La plupart des chercheurs dans le domaine de l'éducation estiment que l'apprentissage en ligne ouvre la voie à la mise en œuvre de l'article précédent. Les ressources trouvées sur les plateformes et les applications éducatives, aident beaucoup de gens qui n'ont pas la possibilité d'aller à l'école pour une raison ou une autre. L'apprentissage traditionnel ou "apprentissage à l'école" a été connu depuis des siècles avant J.C, selon l'histoire ils ont trouvé des traces de l'école dans des civilisations comme celle de l'Égypte, la Chine et l'Inde, et depuis l'homme n'a jamais cessé d'apprendre. L'école durant les siècles a connu beaucoup de changements et c'était différent pour chaque pays, mais la plupart sont passés par des phases communes, comme la généralisation de l'école, l'éducation gratuite pour le primaire surtout, l'introduction de nouvelles matières et sciences etc. Mais le principe de l'éducation n'a pas changé, il est resté le même : un enseignant, un ensemble d'élèves et des ressources pédagogiques qui changent selon la période, la Figure 2.1 représente un ensemble d'élèves dans une classe, l'une datée de 1955, et l'autre de 2019 respectivement, pratiquement il n'y a

pas une grande différence, ce qui mène vers la création d'autres systèmes d'apprentissage comme l'apprentissage en ligne. La propagation des technologies web et les technologies d'information et de communication en générale jouent un rôle majeur dans l'innovation de ce type d'apprentissage.



FIGURE 2.1: École privée saint-nicolas (paris-normadine, 2019) et Une classe de 1955(canal-blog, 2017)

2.1.2 Historique de l'apprentissage à distance

L'évolution de l'apprentissage en ligne a connu plusieurs phases, et il a été toujours lié à l'évolution d'autres domaines. Plusieurs écoles et universités même des siècles auparavant, ont créé ce qu'on appelle l'apprentissage à distance, où l'apprenant et l'enseignant ne se situent pas dans la même salle. Les (Figure 2.2, Figure 2.3, Figure 2.4) représentent l'historique de formation en ligne.

Vu que l'apprentissage à distance a connu beaucoup d'innovations nous allons diviser l'historique en trois parties.

Avant 1900

En 1728, un certain Caleb Phillips, enseignant de sténographie a placé une annonce dans le journal de "The Boston Gazette" pour offrir des cours par correspondance en mail, il a mentionné que les cours de sténographie, seront envoyés chaque semaine par mail pour toute personne inscrite et il a garanti la qualité de cours. (Kentnor, 2015 ; Ferrer, 2019) et c'était la première référence connue à l'enseignement par correspondance. En 1840, Isaac Pitman un auteur d'une méthode de sténographie, a commencé l'enseignement

de cette méthode par correspondance en Angleterre. Ils ont utilisé pour transformer les informations les cartes postales (Verduin et Clark, 1991), Pitman est aujourd'hui connu comme le pionnier de l'enseignement à distance (Kentnor, 2015). Quelques années après, l'université de Londres a créé un programme externe et devient la première université au monde à proposer des diplômes complets à distance, comprenant aussi les étudiants à travers le monde et cela a été en 1858 (University of London ; Ferrer, 2019). En 1892, le terme «enseignement à distance» est utilisé pour la première fois par l'Université du Wisconsin – Madison (Ferrer, 2019). Dans la même année, William Harper Rainey un linguiste américain a offert des cours universitaires par correspondance à l'Université de Chicago. C'est ce cours qui a connu un grand succès, environ 3000 étudiants sont inscrits à 350 cours. (Kentnor, 2015).

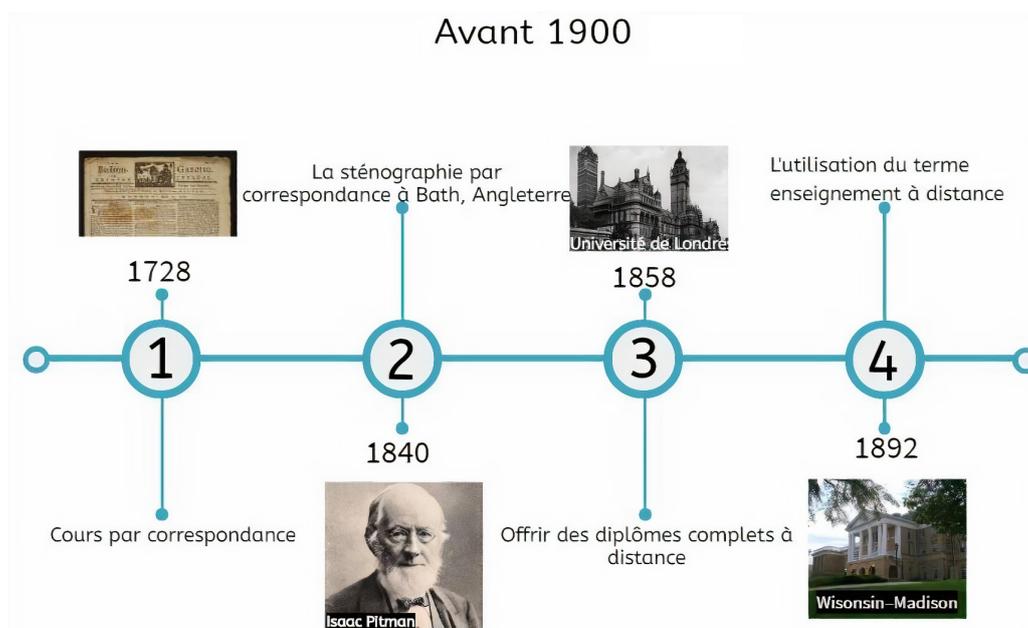


FIGURE 2.2: L'histoire de l'apprentissage à distance partie 01

Période entre 1900 et 2000

Les éducateurs et les chercheurs dans ce domaine commencent à explorer certaines technologies de communication qui vont aider à atteindre beaucoup plus d'apprenants. En 1906, L'Université de Wisconsin Madison utilise les enregistrements phonographiques afin d'envoyer aux apprenants à distance des cours et des conférences, ce qui a préparé le terrain pour l'apprentissage en ligne (Ferrer, 2019). Entre 1932 et 1937, les éducateurs pensent

finalement à utiliser la diffusion télévisuelle, et cela a eu lieu à l'université d'Iowa (Koenig et Hill, 1967). En 1960 l'université de l'Illinois crée des systèmes intranet PLATO signifie « *Programmed Logic for Automatic Teaching Operations* » (White, 2017), qui présentent un système d'apprentissage assisté par ordinateur permettant aux étudiants d'accéder aux supports de cours et aux conférences enregistrées. En 1981, L'institut des sciences du comportement Western B propose le premier programme collégial en ligne, via son école de gestion et d'études stratégiques (Ferrer, 2019). Des années plus tard, en 1992 on a connue l'apparition de l'approche personnalisée assistée par ordinateur (CAPA), développée par l'université de Michigan (Kentnor, 2015). En 1997, Le Journal *d'asynchronous Learning Networks* est créé pour publier et unir les chercheurs dans le domaine de l'apprentissage à distance, encore dans la même année, les réseaux d'apprentissage interactif sont créés et utilisés par plusieurs écoles.

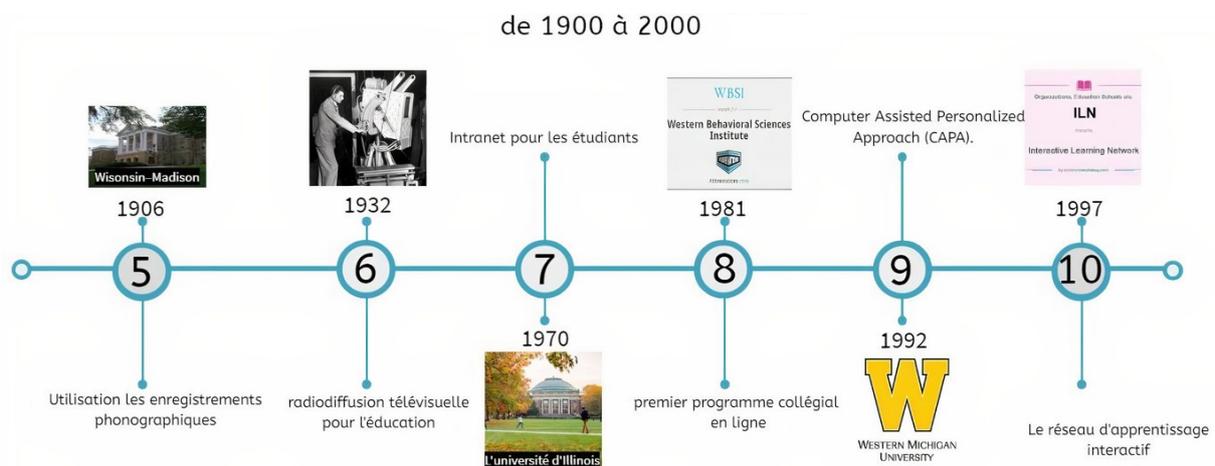


FIGURE 2.3: L'histoire de l'apprentissage à distance partie 02

Après les années 2000

Les années 2000, ont connues une grande évolution technologique qui a touché presque tous les domaines. L'apprentissage en ligne a aussi connu un grand développement dans cette époque, grâce à l'évolution des machines de calcul et l'internet, ainsi que le développement de systèmes de gestion de l'apprentissage (LMS), et la création des académies et des plateformes pédagogiques pour créer et diffuser des cours en ligne comme par exemple Khan Academy, dont ils ont diffusé les premiers cours en 2006 (Baumard, 2020), ensuite

l'académie a été fondée officiellement en 2008. Plusieurs autres plateformes ont été conçues pour diffuser des cours des grandes universités mondiales en plusieurs langues, dans tous les domaines éducatifs basés sur le principe des massive open online courses (MOOC) comme Udacity en 2012, et coursera aussi en 2012 qui est l'une des plateformes éducatives les plus connues à nos jours, ainsi que EdX en 2012, et plusieurs autres. L'an 2013 a connu le lancement de l'université Florida-Online devenant ainsi la première université publique exclusivement en ligne. En 2018 le nombre des universités et collèges publics qui proposent des formations en ligne de tous types a dépassé le 98%. En 2019 l'Université de Pennsylvanie devient la première université à proposer un programme de licence totalement en ligne (Ferrer, 2019).

L'évaluation des cours en ligne et très attachée à celle des technologies, les chercheurs du domaine essaient toujours de trouver des solutions pour faciliter l'accès à des cours de qualité aux apprenants sans tenir compte des frontières.

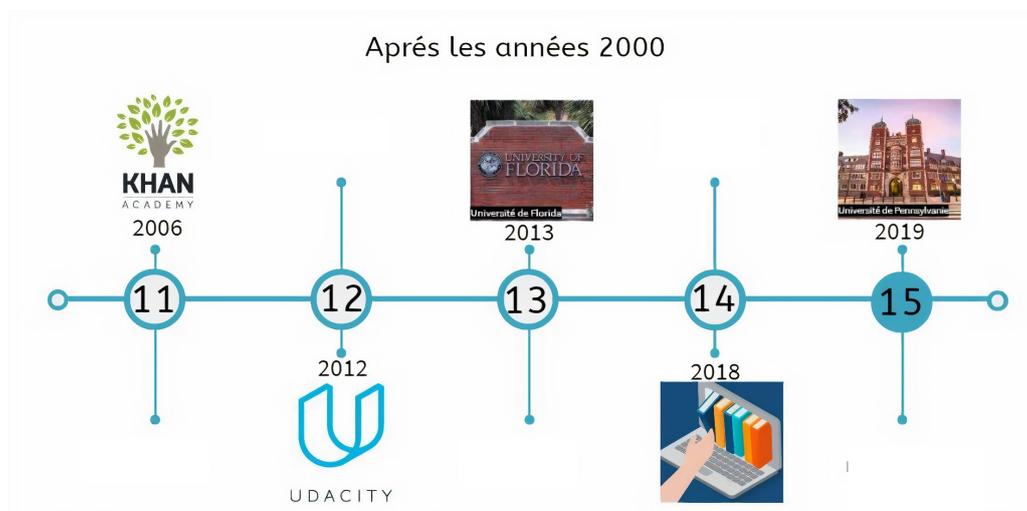


FIGURE 2.4: L'histoire de l'apprentissage à distance partie 03

2.1.3 Apprentissage intelligent

Selon Larousse, L'intelligence est défini comme " l'aptitude d'un être humain à s'adapter à une situation, à choisir des moyens d'actions en fonction des circonstances". Donc on peut dire que l'intelligence reflète l'adaptation dans une situation plus ou moins complexe. Idem pour les systèmes informatiques, un système est dit intelligent s'il peut collecter et

analyser les données des utilisateurs pour arriver à adapter ses services à leurs besoins. Les systèmes d'apprentissages intelligents dits aussi : système tutoriels intelligents (STI), sont des environnements d'apprentissage assistés par ordinateur, qui ont pour but de personnaliser l'apprentissage. Pour (Buche, 2005), ces systèmes visent à diagnostiquer les problèmes des apprenants afin de leur fournir l'aide nécessaire selon leurs besoins spécifiques. (Nyamen et Ange, 2016) ont défini les STI comme les systèmes qui permettent l'automatisation de l'enseignement et permettent aussi de favoriser l'apprentissage sans l'intervention d'un tuteur humain. Les problèmes rencontrés dans les formations d'apprentissage en ligne, favorisent le principe d'adapter ses services aux besoins des apprenants dont la naissance de ce qu'on appelle la personnalisation de l'apprentissage en ligne. Selon (Lefevre et al., 2012), il existe deux types de personnalisations : le premier consiste à adapter les ressources pédagogiques au contexte d'enseignement, ce type s'intéresse au contenu de la formation d'une part et des habitudes des enseignants de l'autre part, le deuxième type consiste à adapter les ressources pédagogiques selon les préférences des apprenants.

La motivation est parmi les raisons pour lesquelles on utilise la personnalisation dans les formations en ligne, lorsque les apprenants contrôlent le temps la façon et les habitudes d'apprentissage, ils vont être plus motivés pour apprendre, de plus la motivation donne envie à accomplir des tâches, et à avancer dans le processus de l'apprentissage. Aussi les différents styles d'apprentissage est l'une des raisons majeurs pour l'utilisation de personnalisation dans les formations en ligne, chaque apprenant a son propre rythme d'apprentissage, l'utilisation de personnalisation basée sur les apprenants va favoriser leurs envies de maintenir la formation en ligne.

Dans la section suivante nous allons discuter des travaux présentés dans la littérature qui favorise les systèmes d'apprentissage en ligne sensibles au contexte.

2.1.4 Systèmes d'apprentissage à distance sensibles aux contextes

Dans cette section nous allons présenter quelques travaux publiés dans les années précédentes, dont les auteurs ont pour but de créer des systèmes, plateformes ou même

de proposer des méthodes ou des architectures pour créer un apprentissage à distance auto-adaptative Table 2.1. Dans cette enquête nous allons présenter la problématique de chaque travail, discuter la contribution et les résultats proposés pour arriver à la fin à déduire les points forts (Table 2.2) et faibles de chacun (Table 2.3).

2.1.5 Travaux connexes sur l'apprentissage auto-adaptatif

Le premier travail qu'on a vu dans la table 2.1, est celui de (Monterrat et al., 2015), qui s'intéresse à la gamification, comme étant un moyen pour motiver les apprenants et améliorer leur participation aux activités d'apprentissage. Ils ont présenté un système de gamification générique et adaptatif qui peut être connecté à plusieurs types d'environnements d'apprentissage web, l'adaptation dans cette approche est basée sur l'analyse des traces d'interaction, afin d'initialiser la liste de types de joueurs représentés dans le modèle d'utilisateur, qui présente l'un des modèles de l'architecture du système et qui permet de prendre en charge l'adaptabilité des éléments du jeu. Le système travaille dans plusieurs étapes successives, premièrement il faut le connecter avec un environnement d'apprentissage web qui existe déjà, ensuite la traçabilité des interactions entre l'apprenant et l'environnement est stocké en permanence. Deuxièmement, les données collectées vont servir à calculer le niveau d'engagement de l'utilisateur dans le système d'analyse de traces, lorsque ce dernier détecte un désengagement de l'utilisateur, il envoie une alerte au moteur d'adaptation qui à son tour met à jour les informations du modèle de joueurs (apprenants) en fonction de l'historique du niveau d'engagement, et sélectionne la fonctionnalité du jeu. Cette approche souligne une caractéristique très importante, l'utilisation des jeux dans le domaine d'apprentissage, en reposant sur le fait qu'un apprenant n'est pas toujours motivé pour jouer. Ils ont donc utilisé la fonctionnalité épiphyte qui permet d'échanger les éléments de jeu au besoin en n'affectant pas l'environnement d'apprentissage une fois intégré dans l'interface utilisateur, par contre elle ajoute une caractéristique d'amusement selon ce que l'apprenant veut recevoir dans un moment particulier, ceci ne transforme pas chaque activité d'apprentissage en un jeu.

Un autre travail est celui de (Drissi et Amirat, 2016) a comme objectif principal de déve-

Auteurs et An- nées	Facteurs d'adaptation	Propositions	Types d'appre- nants
(Montserrat et al., 2015)	- L'analyse des traces d'interaction permet d'initialiser la liste de types de joueurs qui permet de catégoriser les apprenants.	- Un système de gamification générique et adaptatif qui peut être connecté à plusieurs types d'environnements d'apprentissage web.	Tous types d'apprenants.
(Drissi et Amirat, 2016)	- Le style d'apprentissage définit via un questionnaire.	- Proposition d'un système intégrable dans les plateformes web (LS-AEHS).	Des étudiants universitaires
(B. Zhang et al., 2016)	- Catégorisation selon une enquête. - Adaptation selon le travail engagé actuellement, les besoins d'apprentissage des professionnels, les caractéristiques personnelles, et situations environnementales.	- Proposition d'un système mobile contextuel (WoBaLearn).	Des travailleurs
(Mehmood et al., 2017)	- L'environnement (villes intelligente).	- Proposition d'un système ubiquitaire de personnalisation pour la tâche d'enseignement et d'apprentissage dans les villes intelligentes (UTiLearn).	Des étudiants
(Tarus et al., 2018)	- Niveau de connaissances et objectifs d'apprentissage. - L'utilisation de filtrage collaborative.	- Proposition d'un système de recommandation hybride basé sur le contexte et l'exploration séquentielle de modèles.	Des étudiants

TABLE 2.1: Représentation des quelques études réalisées dans l'apprentissage auto-adaptative

opper un style d'apprentissage basé sur un système hypermédia adaptatif d'apprentissage en ligne (LS-AEHS) et d'évaluer l'effet de l'adaptation selon le style d'apprentissage sur

les apprenants. Le système (LS-AEHS) était organisé sous la forme de cinq composants de base, le modèle de l'apprenant, le modèle de domaine, le module d'adaptation, le module d'analyse des traces et le module d'évaluation. Ces composants sont en interaction afin d'adapter les différents matériels d'apprentissage selon les préférences des apprenants, l'adaptation du style d'apprentissage est garantie via un questionnaire délivré aux apprenants dès l'inscription au cours. Cette méthode a prouvé que les apprenants ayant suivi un cours avec l'intégration du système LS-AEHS, ont montré des résultats meilleurs que ceux qui ont suivi le cours sans le système d'adaptation, mais en même temps l'échantillon d'apprenants pour prouver la fiabilité du système était petit d'un côté, et l'utilisation du questionnaire seulement comme facteurs d'adaptation ne peut pas donner des résultats fiables de l'autre côté.

Dans un autre contexte hors l'école, (B. Zhang et al., 2016) proposent un travail dont le type des apprenants visés sont les travailleurs, le but de leur papier est de faciliter l'apprentissage dans l'environnement du travail professionnel. En utilisant un système mobile contextuel appelé WoBaLearn. L'apprentissage en milieu de travail permet l'acquisition de connaissances et de compétences pour résoudre des problèmes afin d'augmenter l'efficacité du travail. Ce système permet d'organiser les travailleurs en utilisant un questionnaire, qui permet de délivrer des résultats sur les préférences professionnelles, les habitudes, et les caractéristiques personnelles de chaque employé. Ensuite, le système propose une série d'activités d'apprentissage personnalisé pour chaque employé afin de l'aider à effectuer son travail ou à renforcer son apprentissage, ces activités d'apprentissage sont conçues sur la base des résultats de l'enquête précédente. Cette méthode a été évaluée sur deux scénarios de travail différents et a prouvé de bons résultats, cependant, l'évaluation d'adaptation est faible à cause du manque de comparaison avec d'autres travaux du même domaine.

(Mehmood et al., 2017) ont comme but de créer un Framework UTiLearn (Ubiquitous eTeaching eLearning), pour la personnalisation du processus d'apprentissage dans un contexte de villes intelligentes en utilisant IoT, big data et le deep Learning afin d'améliorer le développement, la gestion et la prestation de l'enseignement et de l'apprentissage. Le système UTiLearn, est composé de cinq composants ; analyseurs de sentiments, reconnais-

sance d'activité et d'identification des utilisateurs (UARI), conseiller en mode de livraison de contenu adaptatif (ACDMA), composant SNPDP (*System and Network Provisioning Demand Prediction*) et un volet d'analyse et de performance, prévision et intervention (P2AI). Les différentes technologies utilisées dans les composants de ce système sont nécessaires pour améliorer la précision de la formation, du calcul et de la prédiction des caractéristiques du réseau, des modèles de trafic d'application et des modèles de système. Comme résultat le papier montre que les étudiants qui ont interagit avec le système ce sont ceux qui ont obtenu des notes plus élevées, ce qui implique que les notes d'un étudiant sont directement proportionnelles au nombre de fois où l'étudiant interagit avec le système.

À partir du nombre massif de ressources d'apprentissage sur le web et le fait qu'un apprenant peut facilement se noyer dedans, en plus de leurs besoins d'apprentissage différents, Tarus et al., (2018) pensent à créer un système de recommandation hybride. Cette approche se base sur les informations contextuelles de l'apprenant (CA), le filtrage collaboratif (CF) qui présente une méthode de recommandation à l'utilisateur actif que d'autres utilisateurs ayant des goûts similaires, et sur l'algorithme SPM qui a été introduit pour la première fois par (Agrawal et Srikant, 1995), qui permet d'extraire les journaux Web et découvrir les modèles d'accès séquentiel de l'apprenant. L'approche de recommandation hybride proposée dans cette étude combine CF, CA et SPM dans la recommandation de ressources d'apprentissage en ligne passe par trois parties, premièrement, ils ont intégré la connaissance du contexte et les modèles d'accès séquentiel de l'apprenant dans le processus de recommandation afin d'améliorer la personnalisation des recommandations. Deuxièmement, ils ont fait des calculs de similitude afin d'améliorer la précision des prédictions. Enfin, ils ont montré que leur approche fournit des recommandations plus précises que d'autres méthodes de recommandation par expérimentation. Le système proposé est capable de suggérer des ressources d'apprentissage aux apprenants qui répondent à leurs préférences selon leurs informations contextuelles. Afin d'évaluer l'efficacité de ce système ils ont évalué trois autres algorithmes de recommandations sur la même expérimentation. Les trois autres algorithmes de recommandations étaient CF combiné avec CA (CF –

CA), l'algorithme GSP et l'algorithme CF. Les résultats montrent que le système hybride SPM, CA et CF améliore les performances et la qualité du recommandations.

Dans les tableaux suivants nous allons proposer les points forts et faibles pour chaque travail proposé dans le Table 2.1 :

Auteurs et Années	Points fort
(Montserrat et al., 2015)	<ul style="list-style-type: none"> - Cette approche présente un plug-in qui peut être connecté à différents environnements d'apprentissage basés sur le Web. - Ils ont pris en considération le niveau de motivation d'apprenants et leurs types de joueurs. - Ils proposent l'adaptation de fonctionnalités multi-joueurs.
(Drissi et Amirat, 2016)	<ul style="list-style-type: none"> - Système facile à concevoir et efficace comme il a proposé de bons résultats. - La conception des activités et son enchainement ont subi une importance majeure dans le système.
(B. Zhang et al., 2016)	<ul style="list-style-type: none"> - La fourniture d'un support d'apprentissage aux professionnels juste à temps qui est adapté à leurs besoins d'apprentissage, leurs caractéristiques, habitudes et capacités.
(Mehmood et al., 2017)	<ul style="list-style-type: none"> - Enregistrement des données d'activités des utilisateurs et la construction de ses propres ensembles de données selon ses préférences ce qui donne une meilleur personnalisation
(Tarus et al., 2018)	<ul style="list-style-type: none"> -L'approche offre une meilleure performance et une meilleure qualité de recommandations. - Cette approche peut atténuer le problème de la rareté des données.

TABLE 2.2: Les points forts des traveaux réalisés dans l'apprentissage auto-adaptative

L'adaptation de l'apprentissage en général rend ce processus plus facile et fiable, comme le montre les différentes recherches dans ce domaine, plusieurs types d'adaptations peuvent être intégrer en utilisant beaucoup de moyens selon les types d'apprenants et la nature de cours présentés. Cependant ce domaine souffre encore de plusieurs problèmes, comme ; le manque de ressources éducatives organisées pour le test des résultats, le contrôle de grand nombres d'apprenants, le partage efficace et rapide de ressources, etc. Ce qui nécessite souvent l'utilisation de certaines technologies, pour surmonter ces lacunes et avoir de meilleurs résultats. Parmi ces technologies l'utilisation de Cloud computing représente l'une des meilleures solutions.

Auteurs et Années	Points faible
(Montserrat et al., 2015)	- L'approche n'est pas encore testée on ne peut pas être sûr des résultats.
(Drissi et Amirat, 2016)	- La présente étude n'a compris qu'un petit échantillon et cela ne peut pas donner des résultats fiables. - Le seul facteur d'adaptation utilisé dans cette étude est le style d'apprentissage par contre il y en a des facteurs beaucoup plus importante. - L'utilisation de questionnaire pour définir le style d'apprentissage de l'apprenant est une méthode faible et ne reflète pas les vrais besoins de l'apprenant comme les méthodes basées sur d'interaction avec l'interface.
(B. Zhang et al., 2016)	- L'évaluation d'adaptation est faible à cause du manque de comparaison avec d'autres travaux du même domaine. - Se fier aux résultats de l'enquête sans intégrer d'autres capteurs qui vont donner des résultats plus fiables.
(Mehmood et al., 2017)	- Un problème majeur dans la conception d'un vrai système UTiLearn serait la gestion des données et des informations en raison de la quantité de technologies utilisées dans la proposition, qui présente dans le prototype un bon résultat, mais cela peut être une tâche difficile à concevoir dans la vie réelle.
(Tarus et al., 2018)	- L'utilisation de filtrage collaboratif peut enfermer un utilisateur dans une certaine thématique sans la possibilité de l'ouvrir vers d'autres perspectives, et ce n'est pas ce qu'il cherche même en se basant sur son contexte.

TABLE 2.3: Les points faibles des travaux réalisés dans l'apprentissage auto-adaptative

2.1.6 Apprentissage mobile basé sur le Cloud

Dans notre contribution nous allons proposer une architecture pour l'apprentissage basée sur le Cloud mobile, L'utilisation de MCC permet à l'appareil de surpasser ses limites, les différents appareils vont compter sur le Cloud pour le traitement et le stockage. Idem pour l'apprentissage basé sur le Cloud Mobile, l'utilisation de cette technologie permet d'améliorer la qualité d'apprentissage et de faciliter la tâche pour les apprenants et les enseignants, en biais des différents services comme la vitesse de traitement, la capacité de stockage de données, la durée de vie de la batterie, l'offre dynamique de ressources à la demande, etc. (Sarode et Bakal, 2020).

Dans cette section nous allons discuter quelques travaux dans ce contexte, pour arriver à extraire les défis et les problèmes lors de la mise en œuvre d'un système d'apprentissage

basé sur le Cloud mobile.

Les Limites de l'apprentissage mobile

Les utilisateurs de smartphones connaissent une variété d'applications mobiles qui desservent tous les domaines, y compris l'éducation, la communication, la médecine et la publicité, etc. Les utilisateurs plongent de plus en plus sur leurs appareils mobiles, selon Statista (2019), les utilisateurs d'appareils mobiles sont prévus pour atteindre 4,68 milliards d'ici 2019. Grace à cette prolongation du nombre des utilisateurs mobiles, plusieurs écoles et universités optent vers l'utilisation de M-learning dans leur système scolaire pour améliorer les résultats d'enseignement. L'utilisation des appareils mobiles pour améliorer l'apprentissage a connu une grande augmentation ces dernières années (Jones et al., 2013), et même hors des établissements d'apprentissages organisés, plusieurs chercheurs dans le domaines considèrent le M-learning comme un substitut à une éducation de qualité (West et Paine, 2010) et déclarent que le M-learning est bien adapté à la façon dont les personnes ont tendance à travailler. Bien que il soit très répondu dans le domaine de l'apprentissage en ligne en raison de la petite taille et la portabilité du mobile (Domingo et Garganté, 2016), ce type d'apprentissage introduit d'autres limites. La petite taille de l'écran de l'appareil mobile pose un problème majeur dans le M-learning (Davies et al., 2012 ; Fouh et al., 2014) exploitent ce sujet, Ali et Rafie Muhammad,(2017) posent une question à 120 élèves de l'école préparatoire pour savoir s'ils considèrent la taille de l'écran comme un obstacle pendant l'apprentissage, 29 étaient d'accord et 34 étaient tout à fait d'accord ce qui représente un pourcentage de (52.5% en addition). Un autre problème a été mis en évidence dans beaucoup de recherches comme (Chuang, 2015 ; Beniwal et al., 2018) est la durée de vie limitée de la batterie, Ali et Rafie Muhammad (2017) posent encore une autre question dans leurs études sur la batterie, pour vérifier si le chargement régulier représente un obstacle dans l'apprentissage, 24 entre les 120 élèves étaient d'accord et 65 étaient tout à fait d'accord, ce qui représente un taux élevé d'arrangement sur la proposition. Une autre limite du M-learning causée par l'hétérogénéité des périphériques à cause de l'utilisation de différents systèmes d'exploitation, ce qui pose un problème de

communication entre utilisateurs (Sobaih et al., 2016; Yousafzai et al., 2016).

La connexion réseau pose également un problème dans l'utilisation de M-learning, mais ce problème est de nos jours est particulier dans certains endroits du monde, où la connexion internet est très faible. Wu Turner (2020) discutent la relation entre la bande passante, l'interaction et les performances et comment cela affecte les cours en ligne, (Dana et al., 2020) confirment qu'il existe des problèmes de débit pour les communautés rurales où de nombreux étudiants abandonnent les cours en ligne en raison d'un accès internet médiocre, et la même chose s'est produite dans certains pays du tiers monde. Un autre problème présenté dans plusieurs recherches comme celles de (Oz, 2015; Yousafzai et al., 2016) est le stockage limité, les étudiants sont souvent inquiétés de la quantité de données pouvant être associées à un cours en ligne. Dans l'étude précédente de Ali et Rafie Muhammad (2017), ils ont posé une question sur le stockage et est-ce que cela représente un genre d'obstacle pour les élèves, de 120 apprenants plus que 70% ont répondu par d'accord et tout à fait d'accord, ce qui montre la préoccupation de ce sujet pour les apprenants. Beaucoup d'autres limites sont présentées dans la littérature comme la distraction présentée par (Punjabi, 2019) dans une étude sur l'attitude des enseignants du secondaire envers l'apprentissage mobile cette étude tente également d'étudier les problèmes et les menaces associées au M-learning, pour 40 enseignants environ 36 d'entre eux (90%) ont convenu que le M-learning crée une distraction pour les étudiants la plupart du temps. Un tel problème par contre représente une situation beaucoup plus profonde qui nécessite une vue générale sur la nature des apprenants, les objectifs de cours et les unités présentés dans chaque séance. Nous pouvons aussi poser des problèmes comme l'acceptabilité (Mahat et al., 2012; Beniwal et al., 2018), l'auto-efficacité (Shih-hsien Yang, 2012) et l'utilisation correcte de l'appareil, et d'autres qui sont fortement liés à la nature de cours et des apprenants. Bien que nous avons identifié de nombreux obstacles pour le M-learning, ces obstacles ne sont pas insurmontables beaucoup d'entre eux peuvent être dépassés pour changer quelques critères au début de cours, d'autres limites par contre nécessitent des technologies avancées comme le Cloud computing mobile.

2.1.7 Systèmes d'apprentissage à distance auto-adaptatif basé MCC

Le Cloud computing Mobile permet de libérer les appareils mobiles de plusieurs tâches, elles représentent le principe de la fourniture à la demande de différentes ressources informatiques. Dans le domaine de l'apprentissage en ligne beaucoup de recherches sont concentrées sur le MCC, pour submerger les limites de M-learning. Dans cette section nous allons présenter certaines évidences de la littérature, pour surmonter les problèmes de la durée de vie de la batterie, l'hétérogénéité, et le stockage limité.

Yousafzai et al., (2016) confirment que l'une des meilleures solutions pour prolonger la durée de vie de la batterie est l'utilisation du Cloud computing mobile, en déchargeant le traitement vers le Cloud et la présentation sur l'appareil de l'apprenant. Dans leur article qui présente une nouvelle conception pour un M-learning associé avec MCC (Suresh Bhosale et Livingston, 2014) ont déchargé l'appareil mobile des calculs qui présentent la partie la plus gourmande en énergie vers les serveurs Cloud, et cela donne comme résultat la prolongation de la durée de vie de la batterie. Du même pour (Omar Balghosoon et Jameel Qureshi, 2014,) qui présentent dans leur recherche un modèle de m-learning, associé avec MCC et enrichi par une infrastructure de cluster de calcul haute performance (HPC) et d'unité de traitement graphique (GPU). Ils ont essayé de surmonter les obstacles M-learning liés à la performance, à la sécurité et l'environnement inclus l'autonomie de la batterie. Concernant l'hétérogénéité de matériel (Steffenel, 2018) la définit « comme une conséquence de la différente construction des dispositifs qui composent le système, notamment leur composition matérielle (processeurs, mémoire) et leur capacité de calcul », plusieurs chercheurs utilisent les MCC comme une solution pour ce problème et considèrent que cette technologie aide toute l'équipe pédagogique et les apprenants à communiquer sur la même base. Li et Gao (2017), présentent dans leur article un cadre de système mobile qui interconnecte des appareils mobiles hétérogènes vers un cloud mobile personnel, leur idée était de masquer l'hétérogénéité matérielle et logicielle des systèmes mobiles en développant un framework de partage qui exploite les services du système d'exploitation mobile afin de partager les différentes ressources entre les appareils mo-

biles. Pour le problème de Mémoire limitée des appareils mobiles qui gênent beaucoup les apprenants, le MCC clairement présente une solution satisfaisante pour les apprenants et l'équipe pédagogique, il faut juste sauvegarder le tout sur le service Cloud de la plateforme éducatives utiliser, comme dans le cas de l'utilisation de SoundCloud pour la musique ou GoogleDrive pour les différents types de documents. Plusieurs autres recherches présentent des modèles, des architectures et même des propositions pour améliorer le mobile Learning en utilisant les MCC. Comme (Shorfuzzaman et al., 2018) qui présentent un framework M-learning basé sur le Cloud mobile, l'idée était d'utiliser les techniques de BigData pour analyser un volume énorme de donnée extraites des apprenants, pour étudier l'état de préparation des apprenants afin de proposer un modèle hypothétique pour l'adoption de l'apprentissage mobile basé sur un modèle d'acceptation technologique (TAM). (Elkhaateb et al., 2019) développent une application mobile easy-Edu, cette application vise à faciliter le processus d'apprentissage basant sur les besoins de l'apprenant en profitant du service du Cloud mobile, cette application encourage et supporte aussi la collaboration entre les différents acteurs éducatifs.

Chaque recherche présentée dans les sections précédentes, ou étudiée sans être forcément citée dans cette thèse, présente une solution plus au moins pratique qui résume la façon dont les chercheurs voient le processus le plus correcte et fiable pour mettre un cadre de travail sur l'apprentissage en ligne personnalisé. L'utilisation du Cloud mobile a boosté les résultats positifs dans la plupart des cas et limite les difficultés de l'utilisation de ce type d'apprentissage. Dans le cas du papier de (Ahmed et al., 2020), l'utilisation de MCC dans le cadre de l'apprentissage permet d'offrir des solutions pratiques et d'améliorer l'apprentissage centré sur l'étudiant, malgré l'infrastructure informatique limitée de l'université. D'autres recherches utilisent en association avec le Cloud mobile des technologies avancées comme le BigData (Shorfuzzaman et al., 2018) ou HPC (Omar Balghosoon et Jameel Qureshi, 2014) pour prouver leurs solutions, d'autres utilisent la gamification pour créer un environnement plus amusant afin de maintenir la motivation des apprenants (Vinay et al., 2013), etc. Mais tous les travaux dans ce domaine cherchent à participer dans la communauté scientifique avec des solutions qui peuvent aider à délivrer une meilleure

qualité d'éducation, en proposant pour chaque apprenant une expérience éducative adéquate, dont il peut apprendre à son rythme et surtout il peut accéder à toutes sortes de ressources éducatives.

2.2 Apprentissage intelligent et émotions

Niema Moshiri, professeur assistant au département d'informatique et d'ingénierie de l'Université de Californie à San Diego (UCSD), a participé dans une conférence de TEDx Talks¹, par un discours très intéressant sur l'apprentissage en ligne personnalisé intitulé "The Era of Online Learning". Dans son discours qui présente MAIT une solution adaptive de cours en ligne basée sur textes collaboratifs adaptatifs, il a déclaré que nous ne sommes pas au stade où nous pouvons remplacer les cours universitaires par des cours en ligne, à cause de la mauvaise adaptabilité. Pour prouver son point de vue, il s'inscrit à la plateforme éducative coursera et suit un cours de machine Learning², en appliquant le programme de certificat. Son but était de terminer les cours et d'obtenir le certificat sans rien apprendre. Donc il a rapidement passé tous les vidéos d'unités de cours et il a cliqué au hasard sur les questions aux choix multiples, puis pour les défis de programmation, il a simplement utilisé des connaissances de base en algèbre linéaire pour réorganiser les équations et trouver la réponse, il a finalement réussi à obtenir le certificat en complétant les cours de 11 semaines en 5 jours sans rien apprendre (Moshiri,2017). Ce qui prouve que si les apprenants ne sont pas motivés à apprendre, ils peuvent finir le cours en ligne sans avoir aucune des compétences ou d'informations promises dans les objectifs initiaux du cours.

Si on prend la profession de l'enseignant et essayer d'en imiter les caractéristiques de base, on trouve que l'enseignant réellement regarde ses élèves, et depuis ce contact visuel, il peut constater leurs émotions envers le cours et dans les meilleurs des cas il change sa façon pour mieux s'adapter avec la plupart d'entre eux.

Pour imiter cette fonctionnalité on peut utiliser la reconnaissance des émotions faciales qui

1. La conférence TED présente des recherches et des idées importantes de toutes les disciplines. <https://www.ted.com/>

2. <https://www.coursera.org/learn/machine-learning>

représente une tâche assez difficile, mais avec les progrès récents de la vision par ordinateur et du Deep Learning, il y a dans la littérature de nombreux résultats impressionnants.

2.2.1 Expressions faciales et unités d'action

Les émotions sont la méthode de communication non verbale la plus puissante et la plus naturelle entre les humains, tout le monde fait des expressions faciales qui donnent une vue générale sur leur état émotionnel. (Ekman et Friesen, 1971) présentent un ensemble de six émotions acceptées comme émotions universelle Table 2.4, où ils ont confirmé que ces six émotions sont utilisées par tous les êtres humains quelques soient leurs origines, traditions, cultures et civilisations pour définir le même type de sensation, ces émotions sont : la joie, la colère, la peur, la tristesse, la surprise et le dégoût Figure 2.5. Le mépris est ensuite ajouté comme septième émotion (Matsumoto,1992) mais il n'était pas largement utilisé.

NB : Les définitions du Table 2.4 sont extraites du dictionnaire Larousse³ Bien que, les

Emotions	Définition
La joie	Sentiment de plaisir, de bonheur intense, caractérisé par sa plénitude et sa durée limitée.
La colère	sentiment d'une agression, d'un désagrément, traduisant un vif mécontentement et accompagné de réactions brutales
La peur	Sentiment d'angoisse éprouvé en présence ou à la pensée d'un danger, réel ou supposé, d'une menace
La tristesse	État de quelqu'un qui éprouve du chagrin, de la mélancolie ; affliction.
La surprise	État de quelqu'un qui est frappé par quelque chose d'inattendu
Le dégoût	Sentiment d'aversion, de répulsion, provoqué par quelqu'un, quelque chose ; fait d'être dégoûté, de ne plus avoir de l'intérêt, d'attachement ou d'estime pour quelque chose

TABLE 2.4: Définition des émotions universelles

humains font beaucoup plus de six émotions dans leurs conversations quotidiennes, certaines de ces émotions représentent même une combinaison de ces émotions universelles,

3. <https://www.larousse.fr/>



FIGURE 2.5: Un échantillon des émotions universelles depuis la base de données CK

la plupart des recherches dans ce domaine se concentrent que sur ces dernières, même pour les bases de données des expressions faciales. Selon (Martinez et Valstar, 2016) il y en a plus de 7000 différentes expressions faciales, certaines expressions peuvent avoir des interprétations multiples selon le contexte dans lequel elles sont présentées, ce qui nécessite un système pour analyser le signal et un autre pour l'interpréter, une des méthodes pour étudier les émotions faciales est les unités d'action.

En partant de l'hypothèse que l'expression faciale est une combinaison de mouvements musculaires faciaux, on peut travailler avec les unités de codage des actions faciales (AU), ces actions faciales sont l'unité de base de la méthode Facial Action Coding System (FACS), qui représente une méthode de description des mouvements du visage développée par les psychologues Paul Ekman et Wallace Friesen (Ekman et Friesen, 1976 ; Ekman et Friesen, 1978). Elle représente une méthode largement utilisée pour décrire objectivement l'activité du visage (Sayette et al., 2001). Elle est représentée par les unités d'action (AU) afin d'identifier l'expression faciale de chaque mouvement de muscle qui correspond à une AU qui décrit les changements de l'apparence visible (Martinez et Valstar, 2016). Les muscles du visage sont divisés en deux parties l'une correspond à la partie supérieure qui

décrit le mouvement et la forme des yeux, des sourcils et des joues du visage et l'autre s'intéresse à la partie inférieure qui a son tour décrit le mouvement et la forme des lèvres, ainsi que les sillons dans les régions radiculaires nasolabiales et nasales (Nonis et al., 2019). De plus FACS à cinq niveaux d'intensité différentes annotées en ajoutant les lettres A-E (pour une intensité minimale-maximale) au numéro de l'unité d'action (Yan et al, 2014) Table 2.5.

Toutes les émotions et les expressions faciales peuvent être codées en utilisant ce système, en biais de combinaison d'unité d'action la table 2.6 représente le codage par AU des émotions universelles, (Ekman et Friesen, 1976) qui ont définissent 46 unités d'action qui présentent les mouvements musculaires du visage.

A	B	C	D	E
Trace	Léger	Marqué, prononcé	Sévère, extrême	Maximum

TABLE 2.5: Intensité de FACS

On trouve dans la littérature d'autres types d'émotion basique comme celle de (Plutchik, 2001), le professeur et psychologue américain Robert Plutchik propose « the wheel of emotions » ou la roue d'émotions Figure 2.6, Le principe de cette roue et de comprendre les émotions et la relations entre eux, le modèle montre qu'il y a huit émotions de base : la joie, la confiance, la peur, la surprise, la tristesse, le dégoût, la colère et l'anticipation. Chaque émotion de base a une émotion psychologiquement contraire, comme la joie est le contraire de la tristesse, la confiance est le contraire du dégoût, la peur et le contraire de la colère, la surprise est le contraire de l'anticipation, ensuite on trouve sur le modèle les émotions qui représentent une combinaison de deux émotions primaire, par exemple la joie et la confiance représentent l'émotion d'amour, la combinaison de la colère et l'anticipation donnent comme résultats une émotion d'agressivité etc. . .

2.2.2 Types des émotions

Il existe pratiquement trois types d'émotions, le premier type est l'émotion basique comme celle présentée par Ekman et Friesen, le deuxième type est défini par la combinai-

Emotion basique	Unité d'action(Hussain et al., 2017)	Description des AU (Ekman et Friesen, 1976)
La joie	6+12	AU6 : Remontée des joues AU12 : Étirement du coin des lèvres
La colère	4+5+7+23	AU4 : Abaissement et rapprochement des sourcils AU5 : Ouverture entre la paupière supérieure et les sourcils AU7 : Tension de la paupière AU23 : Tension refermant des lèvres
La peur	1+2+4+5+20+26	AU1 : Remontée de la partie interne des sourcils AU2 : Remontée de la partie externe des sourcils AU4 : Abaissement et rapprochement des sourcils AU5 : Ouverture entre la paupière supérieure et les sourcils AU20 : Étirement extrême des lèvres AU26 : Ouverture de la mâchoire
La tristesse	1+4+15	AU1 : Remontée de la partie interne des sourcils AU4 : Abaissement et rapprochement des sourcils AU15 : Abaissement des coins externes des lèvres
La surprise	1+2+5+26	AU1 : Remontée de la partie interne des sourcils AU2 : Remontée de la partie externe des sourcils AU5 : Ouverture entre la paupière supérieure et les sourcils AU26 : Ouverture de la mâchoire
Le dégoût	9+15+16	AU9 : Plissement de la peau du nez vers le haut AU15 : Abaissement des coins externes des lèvres AU16 : Ouverture de la lèvre inférieure.

TABLE 2.6: Les unités d'action de chaque émotion basique

son des émotions basiques comme par exemple la sensation de l'étonnement heureux, et la troisième partie représente les émotions dites complexes comme la confusion, la culpabilité et la jalousie etc...L'évolution du domaine de classification des émotions faciales,

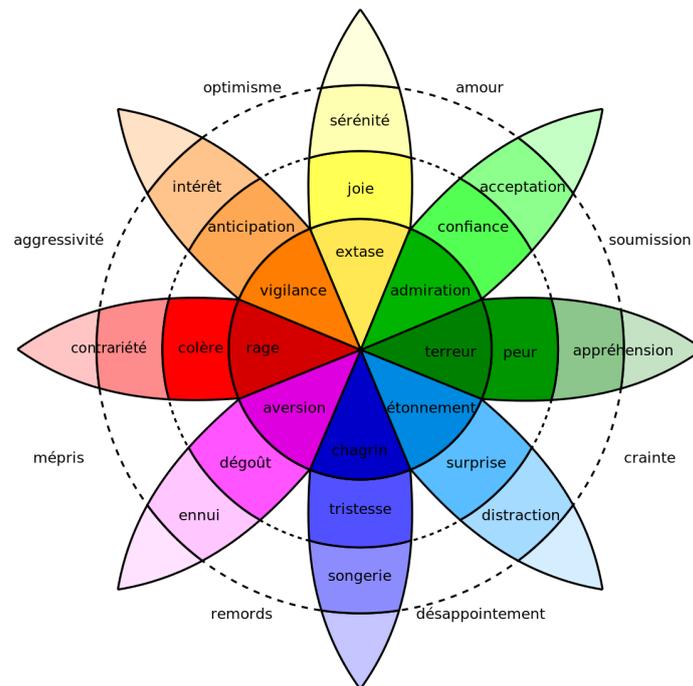


FIGURE 2.6: La roue d'émotion de Plutchik traduit en français ⁴

représente l'une des causes de progression de la reconnaissance des expressions faciales par ordinateur (FER), au niveau du développement de base de données et de description des unités d'actions, mais ce qui nous attire le plus c'est la description des émotions d'apprentissage.

L'utilisation des émotions pour l'apprentissage en ligne augmenterait considérablement les performances des apprenants, si elles sont utilisées pour la personnalisation d'apprentissage (Sarrafzadeh et al., 2008), beaucoup de recherches utilisent les émotions des apprenants pour adapter et ajuster les cours selon leurs besoins, comme un enseignant dans une classe traditionnelle qui peut ajuster sa stratégie d'enseignement en observant les émotions faciales et les mouvements corporels des élèves (Bahreini et al., 2016). Ce que nous avons remarqué dans les travaux présentés dans le domaine de la reconnaissance des émotions faciales des apprenants, est le type des émotions choisies par les chercheurs Table 2.7.

Auteurs et Années	Les émotions utilisées
(Shen et al., 2009)	engagement, confusion, ennui, espoir
(Bahreini et al., 2016)	la joie, la tristesse, la colère, la surprise, la peur, le dégoût et l'état neutre
(Coppens, 2017)	engagement, confusion, frustration, ennui
(Di Leo et al., 2019)	Surprise, plaisir, curiosité, confusion, anxiété, frustration, ennui
(Hossain et Muhammad, 2019)	la joie, la tristesse, la colère, la surprise, la peur, le dégoût
(Mohamad Nezami et al., 2020)	Satisfaction, confusion, ennui
(Zatarain Cabada et al., 2020)	ennuyé, frustré, engagé, excité, détendu, concentré

TABLE 2.7: Les émotions utilisées dans quelques recherches

En observant les émotions utilisées dans les recherches présentés dans le tableau 2.7, on trouve des émotions basiques comme dans le cas de (Bahreini et al., 2016 ; Hossain et Muhammad, 2019) et des émotions complexes réservées pour décrire un état émotionnel d'un apprenant comme le cas des autres recherches. Logiquement une sensation comme la peur, ne présente pas une émotion d'apprentissage, mais cela est pratiquement liée au type du cours présenté, si par exemple un cours qui présente une pandémie mondiale, ou une catastrophe naturelle le sentiment de la peur présente dans ce cas une émotion d'engagement positif de l'apprenant. Généralement pour les émotions d'apprentissage on trouve quatre catégories présentées par (Pekrun, 2006) comme suit :

- Activation positive comme curiosité et espoir.
- Désactivation positive comme soulagement.
- Activation négative comme confusion et frustration.
- Désactivation négative comme ennui, désespoir, tristesse.

Après avoir présenté l'émotion, leurs historiques, leurs états actuels, et leur relation avec le système de codage de l'action faciale et les unités d'actions, dans la section suivante, nous plongerons dans les recherches récentes sur l'extraction, la classification et la reconnaissance des émotions faciales en utilisant les techniques de Deep Learning.

2.2.3 Deep Learning et expressions faciales

Le progrès technologique de ces dernières années a permis au domaine de la vision par ordinateur d'obtenir des meilleures performances dans une variété de défis, dont la reconnaissance des émotions faciales. Parmi les différentes technologies utilisées, l'exploitation du Deep Learning renforce cette innovation dès son utilisation pour le projet SuperVision, lors du challenge ImageNet en 2012 (ILSVRC, 2012), le Deep Learning nécessite une grande quantité de données et une puissance de calcul considérablement importante, c'est pour cela qu'on trouve des recherches qui se concentrent sur la création de bases de données de toutes sortes pour permettre de satisfaire ce besoin, ce qui mène vers de meilleurs résultats pour les différents domaines. Pour la reconnaissance des émotions faciales on trouve dans la littérature plusieurs types de base de données qui servent comme bases de test pour plusieurs recherches, ces dataset contiennent des images de visages étiquetées, des vidéo en 2D et 3D, qui permettent une gestion des variations de poses et de comportements faciaux, et même des unités d'actions, la Table 2.8, présente les bases de données les plus utilisés dans ce domaine.

- JAFFE : japanese female facial expression database (Lyons et al., 1998) c'est une base de données qui comprend 213 images, present de 10 femmes japonaises, la résolution de chaque image est 256x256 pixels, ces femme expriment les six émotions de base plus un visage neutre.
- MMI : La base de données MMI (Pantic et al., 2005 ; Koelstra, 2005), comprend plus de 2900 vidéos et 740 images hautes résolutions de 75 personnes des deux sexes, âgés de 19 à 62 ans. Pour capturer des visages neutres et des expressions de six émotions basiques. La BD est entièrement annotée pour la présence des unités d'action, dans les vidéos.
- BU-3DFE : Binghamton University 3D Facial Expression (Yin et al., 2006), est une base de données constituée de 2500 modèles 3D de 100 personnes (56 femmes et 44 hommes), dont l'âge est entre 18 et 70. Cette base de données contient des expressions de visages avec sept états émotionnels universels : neutre, joie, surprise,

peur, tristesse, dégoût et colère.

- Bosphorus : La base de données Bosphorus (Savran et al., 2008), comprend un totale de 4652 images qui proviennent de 105 personnes des deux sexes 60 hommes et 45 femmes, âgés de 25 à 35 ans et la plupart sont de race blanche. la BD contient diverses poses, expressions et conditions d'occlusion des visages neutres et des expressions de six émotions basiques et 28 unités d'action.
- CMU MultiPIE : La base de données CMU MultiPIE (Gross et al., 2010), comprend plus de 750.000 images d'une taille de 3072 x 2048 qui proviennent de 337 personnes, lors de quatre sessions d'enregistrement, la BD contient 06 expressions faciales y'compris le visages neutre le sourire, la surprise, la strabisme, le dégoût et le cri.
- CK+ : Extended CohnKanade CK+ (Lucey et al., 2010), est une base de données qui comprend un totale de 593 séquences d'images de 123 personnes, âgées de 18 à 50 ans. La BD contient diverses expressions et unités d'action, chaque séquence commence par le visage neutre et finit par l'image qui représente le pic d'émotion.
- DISFA : Denver Intensité des actions faciales spontanées DISFA (Mavadati et al., 2013), est une base de données qui comprend un total de 4845 vidéos, enregistrées depuis 27 personnes, en regardant un clip vidéo de quatre minutes par deux caméras. La BD contient les six émotions de base plus le visage neutre et 12 AU a été codée entre 0-5 selon l'intensité voir Table 2.5.
- BP4D-Spontaneous : La base de donnée BP4D-Spontaneous (X. Zhang et al., 2014), représente une collecte principalement 328 3D vidéo et des images 2D avec une taille de 2.6 TB, enregistrées depuis 41 participants (femmes et hommes), la BD contient huit émotions, joie, colère, peur, tristesse, surprise, dégoût, embarras, douleur physique avec le codage de 27 unités d'actions.

Base de données	L'année	Nombre de données	Nombre de sujets	Taille	Expression Faciale
JAFFE	1999	213 images	10	-	Neutre, joie, colère, peur, tristesse, surprise et le dégoût
MMI	2005	2900 vidéos et 740 Images	75	88.5 GB	Neutre, joie, colère, peur, tristesse, surprise et le dégoût + AU
BU-3DFE	2006	2500 images	100	-	Neutre, joie, colère, peur, tristesse, surprise, dégoût
Bosphorus	2008	4652 Images	105	-	Neutre, joie, colère, peur, tristesse, surprise et le dégoût +28 AU
CMU MultiPIE	2009	750.000 images	337	305 GB	Neutre, Sourire, Surprise, Strabisme, Dégoût, Cri
CK+	2010	593 séquences d'images	123	1.5 GB	Neutre, joie, colère, peur, tristesse, surprise, dégoût, mépris + AU
DISFA	2013	4845 vidéos	27	16.5 GB	Neutre, joie, colère, peur, tristesse, surprise et le dégoût +12 AU (1,2,4,5,6,9,12,15,17,20,25,26)
BP4D-Spontaneous	2013	328 vidéos séquences	41	2.6 TB	joie, colère, peur, tristesse, surprise, dégoût, embarras, douleur physique +27 AU
FERG-DB	2016	55767 images	6	-	Neutre, joie, colère, peur, tristesse, surprise, dégoût
Aff-Wild2	2019	558 vidéos	485	48 GB	Neutre, joie, colère, peur, tristesse, surprise, dégoût, valence-excitation + 08 AU (1, 2, 4, 6, 12, 15, 20, 25)
RML	-	720 Vidéos	8	4.2GB	Joie, colère, peur, tristesse, surprise, dégoût

TABLE 2.8: Ensemble de bases de données utilisées pour la reconnaissance des expressions faciale

- FERG-DB : Facial Expression Research Group 2d Database (Aneja et al., 2017), représente une base de donnée d’images 2D qui contient 55767 images annotées, de caractères stylisés (3 hommes et 3 femmes), les caractères sont modélisés à l’aide du logiciel MAYA Figure 2.7, chacun d’eux présente six émotions de base et un visage neutre.
- Aff-Wild2 : La base de données Aff-Wild2 (Kollias et Zafeiriou, 2019), se compose de 558 vidéos, qui montrent des expressions humaines de l’environnements réel de six émotions basiques plus l’excitation et le visage neutre et 08 unités d’actions. Le nombre de personne dans la BD est de 458, dont 279 hommes et 179 femmes, d’un âge quelconque (des bébés aux personnes âgées) et de toutes races et professions, les vidéos sont composer des extraits de la plateforme YouTube sur 260 vidéos, avec 1 413 000 images et une durée totale de 13 heures et 5 minutes.
- RML⁵ : base de données collectée par Ryerson Multimedia Lab contient 720 vidéos de 08 personnes avec les émotions basiques.



FIGURE 2.7: Les caractères de la base de données FERG-DB

De cette enquête sur les bases de données les plus connues et utilisées pour les expressions faciales et les unités d’action, on peut remarquer que la plupart d’entre eux se concentrent sur les émotions basiques et le visage neutre avec une variation dans les types (des vidéos et des images), la représentation (2D et 3D), les poses (spontanées et structurées), avec ou sans codage d’unités d’actions. Mais ce qui nous a manqué c’est les bases de données des différentes émotions comme (confusion, frustration, ennui, etc...). C’est pour cela que la plupart des travaux sur le domaine de reconnaissance des expressions faciales d’apprentissage, créent leurs propres bases de données qui diminuent la valeur de la

5. Le site officiel de la base de données a été supprimé (<https://www.ryerson.ca/multimedia-research-laboratory/rml-emotion-database>) les données sont collectées de : <http://shachi.org/resources/4965>

recherche en matière de tests, en plus qui donnent des résultats d'un taux d'erreur très faible presque négligeable, comme la base de données est créée pour la tâche proposée dans la recherche.

2.2.4 Etude sur la reconnaissance des émotions

Les émotions faciales jouent un rôle très important dans l'industrie d'aujourd'hui, ce qui guide vers une participation plus large dans la communauté scientifique, contrairement au passé où l'analyse des expressions faciales qui jouent un rôle très important dans la reconnaissance des émotions humaines s'était limitée seulement aux psychologues. Généralement, la reconnaissance d'une émotion passe par trois étapes (Kumar et Gupta, 2015), la détection du visage, l'extraction de caractéristiques et la classification des émotions :

- Détection du visage : le principe de la détection du visage dans une image c'est d'extraire les caractéristiques du visage (Front, yeux, nez, bouche...), s'ils ont présentés selon la structure modèle d'un visage alors on peut classer cette image comme une image qui contient un visage.
- Extraction de caractéristiques : Cette phase consiste à manipuler les pixels de l'image afin que la machine découvre des propriétés sur la forme, la couleur, la texture, des objets présentés dans l'image, cette phases et très importante pour la classification.
- Classification des émotions : cette phase nécessite généralement un classifieur qui consiste d'un modèle qui permet de classer les visages de caractéristiques similaires sous la même classe d'émotion.

L'utilisation du deep learning dans la reconnaissance des émotions faciales aussi passe par trois étapes (Nonis et al., 2019), le prétraitement, l'apprentissage approfondi des fonctionnalités et la classification approfondie.

Le prétraitement présent l'étape juste avant l'alimentation du réseau profond par les données, c'est un ensemble de traitements sur les images comme la normalisation, la rotation ou le recadrage d'images etc. Les deux autres étapes dans les Deep Learning sont inséparables et complémentaires contrairement aux méthodes traditionnelles qui représentent deux phases complètement séparées, et cela représente un point de force du Deep Lear-

ning.

Parmi les techniques les plus utilisées dans la reconnaissance des expressions faciales ou des émotions, on trouve le CNN présenté dans le chapitre 1 1, vu sa force d'extraction de caractéristiques de l'image ou de la vidéo. Dans cette section nous allons présenter les résultats de quelques travaux qui utilisent DL pour la reconnaissance des expressions faciales et des unités d'action dans les images et les vidéos afin de caractériser le meilleur choix entre eux.

Discussion du tableau de la reconnaissance des émotions basiques

En observant les résultats de la Table 2.9, on remarque d'abord que la plupart des recherches visent les images statiques dans leurs tests, pourtant l'utilisation des vidéos dans ce type de détection est plus fiable. Généralement les images captent le moment où l'expression est intensive, l'utilisation d'un modèle entraîné par ce types d'images dans la vraie vie ne donne pas forcément de bons résultats, comme on peut avoir qu'un seul instant de l'expression dans la réalité, c'est pour cela que l'utilisation des vidéos augmente la fiabilité d'une manière remarquable. Parmi les difficultés déclarées par les auteurs des recherches précédentes l'exigence de puissance de calcul lors de l'utilisation des techniques de Deep Learning, souvent hors des capacités des machines des auteurs, ce problème devient encore plus important lorsque la résolution des images augmente. Le prétraitement nécessite aussi un espace de stockage très important. Nous remarquons également que certaines bases de données sont plus populaires que d'autres comme le cas de Ck+, est cela peut être justifié par le fait que cette base de données est gratuite et en ligne mais aussi à cause de sa taille qui ne dépasse pas les 2GB, ce qui permet aux chercheurs d'utiliser les unités d'action et les séquences images , pour entraîner leur réseaux sans ce souci de capacités de calcul. Le meilleur résultat détecté dans le tableau est la reconnaissance de la joie, comme elle est une émotion très simple qui nécessite que deux unités d'action pour se former. Contrairement à la peur qui a été l'émotion la moins précise, comme elle est définie par plusieurs unités d'actions, ce qui est difficile de détecter.

Auteurs et années	DataBase	Technique	Type	Joie%	colère%	Peur%	Tristesse%	surprise%	dégoût%
(Mollahosseini et al., 2016)	MultiPIE, MMI, CK+, DISFA, FERA, SFEW, FER2013	CNN : convolutional neural network	Images	86.6	55	47	56.1	89.3	80.3
(Huynh et al., 2016)	BU-3DFE	CNN : convolutional neural network	Images	100	91.3	86.7	87.5	95.7	95.2
(Lopes et al., 2016)	CK+	Convolutional Neural Networks	Images	98.55	93.33	96	84.52	99.20	100
(Zeng et al., 2017)	CK+	DSAE : the deep sparse auto-encoders	Images	100	90	87	86.61	98.80	99.58
(Al-Darraj et al., 2017)	CK+ et RAF-DB	convolutional neural network	Images	94.1	81	81	90.5	95.2	95.2
(Meng et al., 2017)	MMI	convolutional neural network	Images	92.9	81.8	41.1	73.4	76.9	71.9
(Fan et al., 2018)	RAF-DB	MRE-CNN : Multi-Region Ensemble CNN	Images	88.78	83.95	60.81	79.92	86.02	57.5
(S. Zhang et al., 2019)	MMI	CNN, DBN, SVM	Vidéos et Images	66.67	50	75	100	100	66.67
(Zhao et al., 2020)	RML	RAN Deep Residual Attention Networks	Vidéos	93.33	39.17	49.17	68.33	73.33	85.83

TABLE 2.9: Comparaison de la précision de la reconnaissance des émotions basiques

2.2.5 Détection des unités d'actions

La détection des unités d'action présente une tâche un peu difficile pour plusieurs raisons, d'abord c'est difficile de trouver des bases de tests contenant toutes les unités d'action adéquates à une recherche spécifique, aussi les AU représentent un changement rapide dans les traces de certaines régions du visage ce qui exige un calcul considérablement rapide et précis dans des machines puissantes. De plus la détection des AU est très importante dans plusieurs domaines qui nécessitent des réponses en temps réel, ce qui guide la communauté de recherches vers des solutions qui promettent d'être utiles comme l'utilisation de landmark comme des informations supplémentaires (Li et al., 2017), ou encore de profiter des connaissances d'experts (chen et al., 2019), mais ces solutions n'ont pas été testées à grande échelle. Dans ce qui suit nous allons discuter les solutions de deux recherches sur la détection des AU et de leur intensité.

- Action Unit Intensity Estimation using Hierarchical Partial Least Squares (Gehrig et al., 2015) : cet article traite un sujet très sous-estimé sur l'analyse de l'expression faciale, qui est l'estimation des intensités des unités d'action, même avec les techniques actuelles, ce sujet est considéré comme un problème difficile. Dans cet article, les auteurs proposent une nouvelle approche pour modéliser l'UA individuelle à l'aide d'un modèle de régression hiérarchique, qui peut être vu comme une combinaison de modèles de locally linear Partial Least Squares (PLS) localement linéaires où chacun détermine la relation entre les caractéristiques visuelles et les étiquettes d'intensité de l'UA à son niveau de détails. Cette approche s'adapte automatiquement à la non-linéarité dans le domaine source, en ajustant la structure hiérarchique apprise. L'approche a été évaluée sur la base de données Bosphore Figure 2.8.

Cette approche a des avantages et des inconvénients, premièrement, l'idée est importante et n'est pas vraiment discutée dans la littérature au-delà de son effet sur la classification finale des émotions, deuxièmement, selon l'article, l'approche a donné un bon résultat dans la relation entre les caractéristiques et les intensités de chaque AU. En ce qui concerne les inconvénients, l'étude manque de détails comme l'efficacité de cette approche sur la classification finale des expressions faciales, et comment

a-t-elle utilisé les connaissances sur les participants. L'utilisation de R-CNN donnera également un meilleur résultat en raison de sa robustesse de travailler sur les régions ce qui est parfait pour les AU qui dépendent fortement d'une région spécifique.

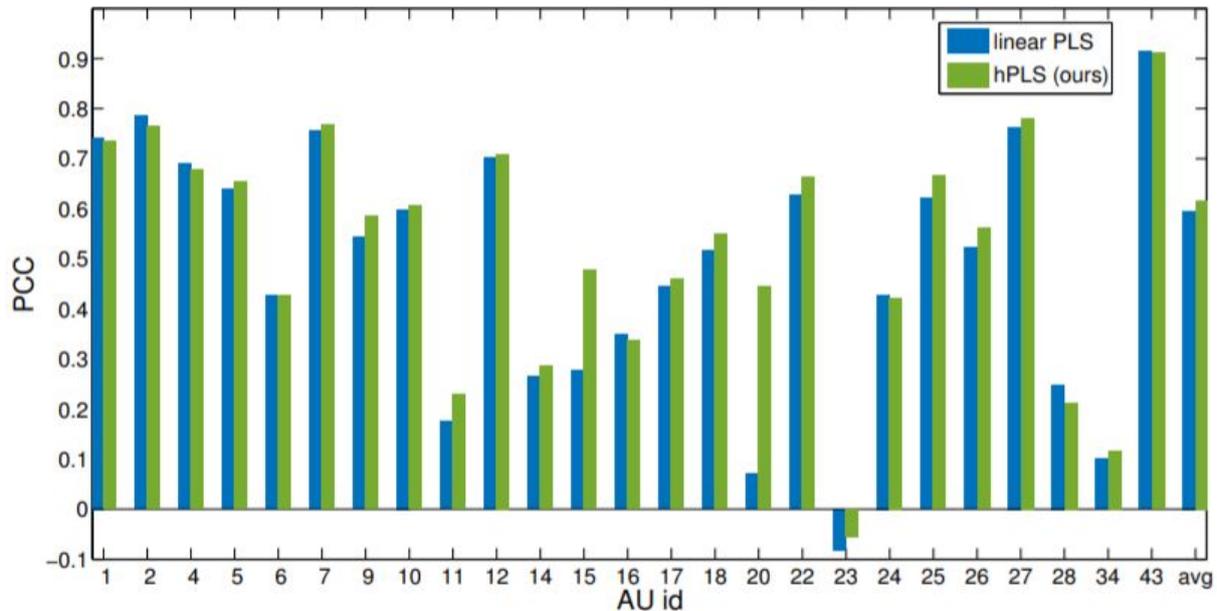


FIGURE 2.8: Les performances des AU individuelles sur la base de données Bosphore (Gehrig et al., 2015)

- EAC-Net : Deep Nets with Enhancing and Cropping for Facial Action Unit Detection (Li et al., 2018) : Dans cet article, les auteurs proposent une approche de détection des unités d'action basées sur le Deep Learning. Ils gèrent la solution en ajoutant deux couches, une couche d'amélioration (enhancing layers) et une couche de recadrage (cropping layers), à un modèle de réseau neuronal convolutif (CNN) pré-entraîné. La couche d'amélioration est chargée d'alimenter le modèle par un ensemble de Landmark faciaux afin d'améliorer l'apprentissage. Pour la couche de recadrage, elle recadre les régions du visage autour des Landmark détectés et conçoit des couches de convolutions individuelles pour apprendre des caractéristiques plus profondes pour chaque région du visage. dans la phase suivante, l'approche combine les deux couches pour construire ce que l'on appelle un réseau d'amélioration et de recadrage (EAC-Net). L'approche a été testée sur la base de données BP4D (Table 2.10, présente le résultat) et DISFA-AU.

Pour les avantages, l'EAC-Net présente des bons résultats de détection, même avec une variation de pose (pEAC-Net), il estime aussi les intensités AU, prédit les unités d'actions même avec seulement la moitié d'un visage visible. Par contre, l'utilisation de Landmark dépend du prétraitement qui peut alimenter le réseau avec des informations imprécises cette méthode a un taux d'erreurs à haut risque, aussi l'utilisation du modèle CNN pour ce type de détection est un peu générale, il est préférable d'utiliser le CNN régional qui fournira un meilleur résultat en moins de temps.

AU	AU1	AU4	AU6	AU7	AU10	AU12	AU14	AU15	AU17	AU23	Moy
pEAC-Net	90.3	88.3	85.3	85.4	88.3	88.2	78.6	90.7	78.2	89.1	86.3

TABLE 2.10: La détection des AU sur BP4D selon l'approche pEAC-Net avec de grandes poses de tête(%)

En se basant sur ce que nous avons cerné dans cette enquête, nous avons remarqué que l'utilisation de Deep Learning dans le domaine de reconnaissance des expressions faciales est très largement répondue, et donne des résultats beaucoup mieux que les modèles classiques. En ce qui concerne les bases de données qui représentent l'un des axes les plus importants dans ce domaine, il existe dans la littérature un manque fatal au niveau des émotions non-basiques, ce qui présente l'un des obstacles de l'innovation de détection des émotions faciales pour l'apprentissage, car comme nous avons présenté, les émotions d'apprentissage sont différentes des émotions basiques. De ce point, plusieurs recherches utilisent la détection des AU pour dépasser ce problème, comme ces 46 unités d'action peuvent former n'importe quelle émotion, d'autres visent la création des bases de données personnalisées pour leurs recherches qui à notre avis diminuent la valeur de leur travail.

Conclusion

Dans ce chapitre nous avons fait deux enquêtes, sur deux axes complémentaires pour adapter l'apprentissage en ligne : le premier concerne les systèmes sensibles aux contextes, et le deuxième s'intéresse à la reconnaissance des émotions faciales.

L'apprentissage adaptatif est une approche très importante qui sert à rendre les plate-

formes éducatives plus intelligentes, pour permettre aux apprenants de créer un environnement permettant de dépasser tous les obstacles de l'école. Le besoin d'adapter les environnements d'apprentissage en ligne augmente de jour en jour, surtout lors de la pandémie du Covid19 qui a montré la nécessité de fonder les bases de ce domaine. Dans le chapitre suivant nous allons discuter la première contribution de cette thèse, qui comprend deux parties permettant de cerner le problème et de présenter l'architecture basée sur le Cloud Computing mobile comme solution.

Chapitre 3

Architecture d'un système d'apprentissage mobile sensible au contexte

Contents

I. Etude statistique sur l'acceptabilité de M-learning par les lycéens	85
3.1 Contexte du travail	86
3.1.1 Définition des ressources éducatives libres	87
3.1.2 Modèle d'acceptation de la technologie (TAM)	87
3.1.3 Théorie unifiée de l'acceptation et de l'utilisation de la technologie (UTAUT)	88
3.2 Matériel et méthode	90
3.2.1 Participants	90
3.2.2 Constructions	91
3.2.3 Procédure	91
3.2.4 Hypothèses	93
3.2.5 Création du questionnaire et l'analyse des données	94
3.3 Résultats	95
3.3.1 Statistiques descriptives	96
3.3.2 Composants PLS	97
3.4 Discussion	100
II. Architecture basée sur MCC pour l'apprentissage auto-adaptatif	103
3.5 Relation entre M-learning et MOOC	103
3.6 Architecture CD-Mobile MOOC	105
3.6.1 Processus de conception de l'architecture CD-Mobile MOOC	105
3.6.2 Description de l'architecture CD-mobile MOOC	107
3.6.3 Personnalisation selon CD-Mobile MOOC	116
3.6.4 Avantages de l'utilisation de Mobile Cloud Computing dans l'architecture proposée	118
3.7 Discussion	120
Conclusion	122

Parmi les problèmes dont souffre l'apprentissage en ligne, le taux d'abandon est considéré comme l'obstacle le plus sévère, ce dernier représente la différence entre le nombre

d'inscrits et le nombre des apprenants qui assistent réellement aux cours, c'est pour cela que cet obstacle intéresse les enseignants, l'équipe pédagogique et la communauté de recherches dans ce domaine.

Souvent ce phénomène d'abandon est interprété par : soit un désintérêt de la part de l'apprenant, soit un abaissement dans sa motivation ou encore une inadaptation dans les objectifs d'activités d'apprentissage. A partir de ces hypothèses la personnalisation de l'apprentissage a été largement utilisée dans les différents environnements en ligne, ce qui permet à un apprenant de prendre en charge son processus d'apprentissage en lui distribuant la qualité d'apprentissage conforme à ses besoins, pour le motiver et l'aider. Mais comment la personnalisation peut-elle résoudre ce problème, et comment l'associer avec d'autres technologies pour garantir des meilleurs résultats ?

Dans ce chapitre nous allons présenter la première contribution de cette thèse, le travail est divisé en deux parties : la première consiste à une enquête sur les élèves du lycée, où nous avons évalué leurs acceptations et comportements envers l'utilisation de la technologie mobile pour l'apprentissage, en utilisant les ressources éducatives ouvertes (OER). Nous avons pu à la fin de cette enquête mesurer leurs aptitudes d'apprendre à distance via leur mobile, et de déterminer les limites du M-learning pour ces élèves.

En ce qui concerne la deuxième partie nous avons utilisé les Cloud Computing Mobile afin de proposer une architecture d'un système d'apprentissage basé sur le contexte de l'apprenant, qui surpasse les limites vécues par les élèves lors du survey précédant, et offre un environnement d'apprentissage mobile personnalisé selon le contexte et les besoins de chaque apprenant.

I. Etude statistique sur l'acceptabilité de M-learning par les lycéens

Les applications mobiles augmentent de jour en jour, l'utilisation du mobile dans l'éducation a également considérablement augmentée, l'objectif de cette contribution est d'examiner l'acceptation et le comportement des élèves du secondaire lors de l'utilisation de la technologie M-learning, et aussi de cerner les obstacles communes. Pour arriver à

ce but, nous avons utilisé des ressources éducatives libres (*Open educational resources* OER), cette fonctionnalité offre des implémentations pédagogiques plus efficaces pour les apprenants et les enseignants. Dans cette veine, l'étude que nous allons présenter adopte une approche quantitative basée sur le modèle d'acceptation de la technologie (TAM) et la théorie unifiée de l'acceptation et de l'utilisation de la technologie (UTAUT), pour étudier l'acceptation et le comportement des apprenants envers la technologie d'apprentissage mobile. Une enquête a été adressée à des lycéens de deux lycées algériens, afin d'examiner leurs besoins et leurs acceptations de la technologie et d'identifier les principaux défis et limites que les élèves rencontrent lors d'une éducation ouverte avec des pratiques innovantes dans un processus d'apprentissage digitalisé.

3.1 Contexte du travail

Les utilisateurs de smartphones connaissent une variété d'applications mobiles dans tous les domaines, y compris l'éducation, la communication, les affaires, la médecine et la publicité, etc. Les utilisateurs se plongent de plus en plus sur leurs appareils mobiles, selon Statista (2020a), les utilisateurs d'appareils mobiles devraient atteindre 3,8 milliards d'ici la fin 2021. Les Algériens ont augmenté le nombre de 0,09 million d'utilisateurs en 2000 à 51,49 millions d'utilisateurs en 2018 Statista (2020b). Selon la même source Statista (2019), dans leur enquête sur l'utilisation d'Internet, les Millennials ont passé 188 minutes sur les appareils mobiles pour des tâches Internet en 2016, par rapport à 2017, ce nombre est passé à 223 minutes par jour. Dans le même temps, Generation_z (selon la même statistique) a passé plus de 250 minutes sur mobile et 229 minutes sur d'autres appareils (pc, ordinateur portable, tablette, etc.), ce qui montre à quel point cette génération est en train de devenir immergée de plus en plus dans la technologie.

Pour cette raison, de nombreux chercheurs souhaitent combiner les technologies mobiles et l'apprentissage, pour obtenir un apprentissage auto-adaptatif, où la satisfaction des étudiants est nourrie par une expérience M-learning plutôt que par l'utilisation des méthodes d'éducation traditionnel (Hwang et Lai, 2015), et vu les avantages des ressources éducatives libres (OER), leur intégration dans ce processus, va ajouter des améliorations

impressionnantes pour la plupart des étudiants.

3.1.1 Définition des ressources éducatives libres

Les ressources éducatives libres ont été adoptées pour la première fois en 2002, lors d'un forum sur l'impact des didacticiels ouverts pour l'enseignement supérieur dans les pays en développement, organisé par (l'Unesco, 2002) «La mise à disposition ouverte de ressources éducatives, grâce aux technologies de l'information et de la communication, pour consultation, utilisation et adaptation par une communauté d'utilisateurs à des fins non commerciales ». Au cours des dernières années, plusieurs explications du terme ont été introduites et développées. La présente recherche fait référence aux OER selon la définition de (Unesco, 2017) « Les ressources éducatives libres, sont des ressources d'enseignement, d'apprentissage et de recherche qui résident dans le domaine public ou qui ont été libérées sous une licence de propriété intellectuelle qui permet leur libre utilisation et réutilisation par d'autres ».

Les OER comprennent des cours, des manuels, des vidéos, des tests, des quiz, des jeux d'apprentissage, des logiciels, des cours en ligne ouverts et massifs (MOOC), ou tout autre outil technique ou matériel utilisé pour soutenir le processus d'enseignement et d'apprentissage. Bien que la plupart des êtres humains aient la même capacité d'apprentissage, les systèmes éducatifs et les ressources disponibles au niveau des écoles, universités ou même les maisons, ont considérablement limitées la productivité des apprenants dans certains pays et l'ont augmentée dans d'autres. C'est pour cette raison que dans cette recherche nous avons utilisé l'avantage significatif des OER, car ils sont disponibles presque gratuitement et partout.

Nous avons associé les ressources éducatives libres avec la technologie mobiles pour créer un environnement adaptatif flexible pour chaque apprenant, afin de vérifier son comportement et acceptation de cette approche nous allons utiliser deux modèles TAM et UTAUT.

3.1.2 Modèle d'acceptation de la technologie (TAM)

Le modèle d'acceptation de la technologie (TAM), est une théorie des systèmes d'information pour l'étude de l'utilisation des technologies de l'information développée par Fred

Davis, Richard Bagozzi et Warshaw (Davis et al., 1989). Il a été largement utilisé dans l'étude du comportement des utilisateurs et de la manière dont ils acceptent et utilisent la technologie. Les constructions de la théorie sont : l'intention comportementale (Behavioral Intention), l'attitude envers l'utilisation (attitude towards use), l'utilité (usefulness), la facilité d'utilisation (ease-of-use) et l'usage effectif. Le modèle présente une suggestion sur les facteurs qui influence l'utilisateur lors de son utilisation d'une nouvelle technologie, ces facteurs vont influencer la manière et le moment de la prochaine utilisation, la Figure 3.1 illustre la schématisation principale du modèle.

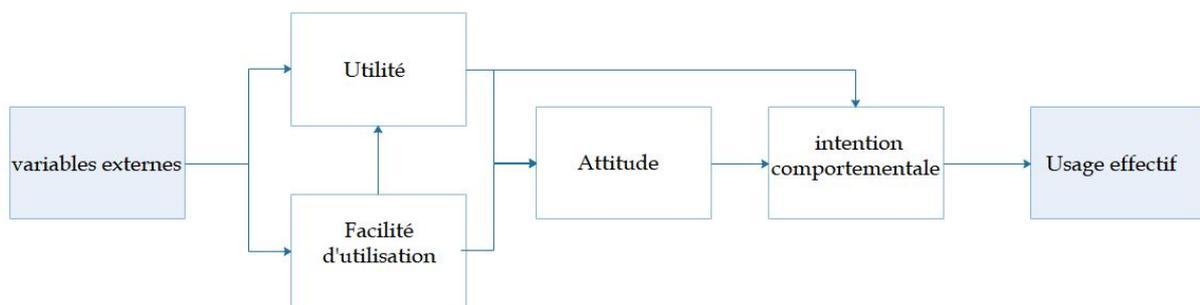


FIGURE 3.1: Le Modèle TAM (Davis et al., 1989)

- Facilité d'utilisation (PEOU) : Le degré auquel une personne pense que l'utilisation d'un système particulier serait sans effort (Davis et al., 1989).
- Utilité (PU) : Le degré auquel une personne pense que l'utilisation d'un système particulier améliorerait sa performance au travail (Davis et al., 1989).
- Attitudes (ATT) : les sentiments positifs ou négatifs d'une personne à propos du comportement ciblé (Fishbein et Icek, 1975).
- Intension (BI) : la force de l'intention d'effectuer un comportement spécifié (Fishbein et Icek, 1975).

3.1.3 Théorie unifiée de l'acceptation et de l'utilisation de la technologie (UTAUT)

Est une théorie du système d'information pour l'étude du modèle d'acceptation de la technologie développé par (Venkatesh et al., 2003). L'UTAUT vise à expliquer les intentions et le comportement des utilisateurs envers un système d'information. Les construc-

tions de la théorie sont : performance espérée (performance expectancy), effort attendu (effort expectancy), influence sociale (social influence), conditions facilitatrices (facilitating conditions), la Figure 3.2, illustre la schématisation principale. Les définitions des

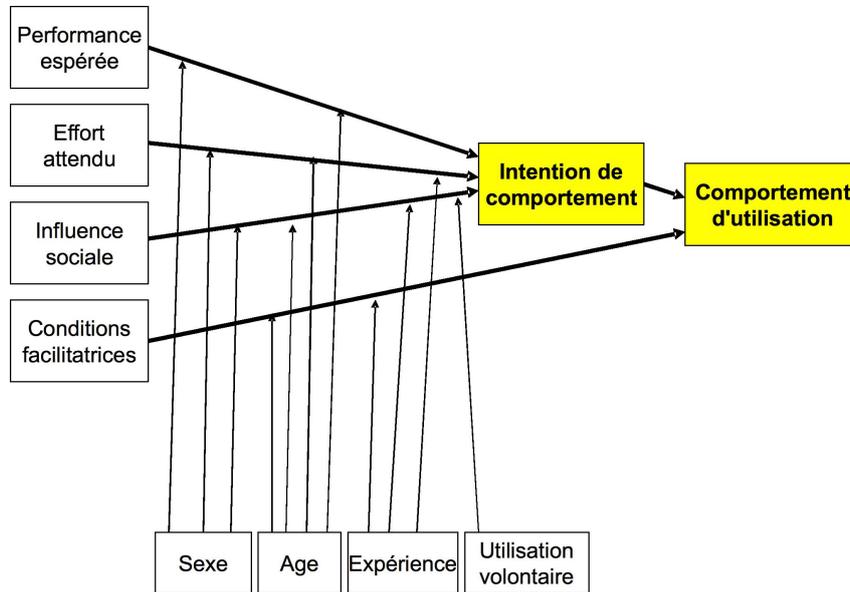


FIGURE 3.2: Le modèle UTAUT (Venkatesh et al., 2003)

constructions sont proposées par (Venkatesh et al., 2003) :

- Performance espérée : le degré auquel une personne croit que l'utilisation du système l'aidera à obtenir des gains et un rendement au travail.
- Effort attendu : le degré de facilité associé à l'utilisation du système.
- Influence sociale : le degré auquel un individu perçoit que d'autres personnes croient qu'il devrait utiliser le nouveau système.
- Conditions facilitatrices : le degré auquel un individu croit qu'une infrastructure organisationnelle et technique existe pour soutenir l'utilisation du système.

Dans la section suivante nous allons décrire les détails de l'utilisation des deux modèles décrits précédemment.

3.2 Matériel et méthode

3.2.1 Participants

L'étude présentée dans cette recherche a interrogé 150 lycéens, issus de deux (02) lycées différents en Algérie, Tebessa, (Hardi Mohamed et Malek Benabi). Les informations démographiques détaillées sont présentées dans la Table 3.1.

Caractéristiques	Descriptions	Fréquence	Pourcentage %
Sexe	F	85	56.7%
	M	65	43.3%
Age	14-15	101	67.3%
	16-17	45	30%
	18-19	4	2.7%
Années	1ère Année	150	100%
Lycées	Hardi Mohamed	108	72%
	Malek Benabi	42	28%
Possession	Smartphone	104	69.33%
	Tablet	11	7.33%
	Smartphone et Tablet	35	23.34%
	Aucun	0	0.00%
Appareils Mobiles	Smartphone	132	88%
	Tablet	18	12%
Système d'exploitation	Android	150	100 %
Heures passées sur mobile par jour	0-1	30	20%
	1-3	45	30%
	3-5	30	20%
	Plus	45	30%
Niveau scolaire (la note sur 20)	7-10	29	19.3%
	10-12	67	44.7%
	12-14	28	18.7%
	14-16	19	12.7%
	Plus de 16	7	4.7%
M-learning	Oui	98	65.33%
	Non	52	34.66%

TABLE 3.1: Description des participants

Nous remarquons que, 85 élèves sont des filles (56,7%) et 65 sont des garçons (43,3%), la plupart des élèves des deux lycées préfèrent utiliser les smartphones (88%) plutôt que les tablettes (12%). Selon (Harfield Nang, 2014), les smartphones sont les appareils les plus utilisés. Pas étonnant, le système d'exploitation le plus utilisé est Android, selon les données mondiales collectées par (Statcounter, 2020) 43.55% des appareils algériens fonctionnant sous Android, ce qui représente (96,02%) des smartphones et (79.98%) des tablettes. Nous avons également remarqué que la plupart des participants (65,33%) avaient de l'expérience dans l'étude des technologies de M-learning, ce qui facilite le processus d'explication.

3.2.2 Constructions

Ce modèle de recherche était basé sur le modèle d'acceptation de la technologie (TAM) et la théorie unifiée de l'acceptation et de l'utilisation de la technologie(UTAUT), composé de six (06) constructions primaires décrites dans la Table 3.2

Les Constructions	ID	Significations
Attitude	ATU	Mesurer l'attitude des élèves envers la technologie
Performance espérée	PE	Mesurer la performance des élèves
Conditions facilitatrices	FC	Mesurer la disponibilité des connaissances, de l'aide et des ressources
Intention	BI	Mesurer l'intention future des élèves d'utiliser la technologie
Facilité d'utilisation	EU	Mesurer la facilité d'utilisation de la technologie
Influence sociale	SI	Mesurer tout type d'influence externe

TABLE 3.2: Description des constructions

Le choix de construction de l'étude :

Le but de cette étude est d'examiner l'acceptation et le comportement des élèves du secondaire envers la technologie mobile lors d'une série de séances en classes et en ligne, utilisant les ressources ouvertes et libres, en se basant sur les constructions des modèles (ATM et UTAUT). Nous avons choisi celle qui correspond le plus à notre but de recherches et nos hypothèses, pour le comportement nous avons utilisé (l'attitude, la performance espérée et l'intention), et pour l'acceptation nous avons opté pour (les conditions facilitatrices, la facilité d'utilisation, l'influence sociale), pourtant ces deux champs sans très liés comme le montre les figures 3.1 et 3.2 mais lors de création de questionnaires nous les avons séparé.

3.2.3 Procédure

Pour faire le questionnaire de notre étude, nous avons eu premièrement la permission des deux directeurs des lycées (Hardi Mohamed et Malek Ben Nabi), ainsi que les parents

des élèves, pour utiliser les données des apprenants dans cette recherche, en leur offrant une copie du questionnaire et en décrivant les buts et les résultats attendus de cette recherche. Ensuite nous avons collaboré avec les enseignants de l'informatique des deux lycées, sur le type de cours, le choix de ressources de partage, et la procédure d'enseignement adéquate. Les enseignants d'informatique dans les deux lycées devaient mettre en place des cours de M-learning pendant trois semaines, les cours étaient administrés par un membre de la recherche qui servait de coordinateur entre les deux lycées. Nous avons utilisé deux types de ressources éducatives libres :

- Diffusion vidéo : les enseignants identifieraient le matériel pédagogique en ligne approprié, représenté par un groupe de vidéos sur l'introduction à la programmation pour les débutants, les cours étaient envoyés en ligne aux étudiants participants via leur courrier électronique, les étudiants devaient travailler sur certaines activités et participer aux discussions en ligne et en classe.
- Logiciel d'apprentissage : Après les trois premières semaines, un examen qui consiste à programmer des problèmes mathématiques simples à l'aide d'un éditeur et interpréteur d'algorithmes a été envoyé aux étudiants, nous avons utilisé le logiciel "Algo Mobile " qui est open sources est facile à utiliser, cette application est développée à la base pour des fins éducatives, il est considéré comme idéale pour les débutants en programmation.

À la fin de l'expérience, les apprenants devaient répondre à l'enquête, qui comprenait 32 questions en deux parties, la première contient les dix (10) questions liées aux informations démographiques et contextuelles. La deuxième partie comprend les constructions définies dans la Table 3.4, une échelle de Likert à 05 points a été utilisée pour concevoir l'enquête afin de donner aux participants un choix plus précis (degré d'accord ou de désaccord) ; *"Tout à fait d'accord, d'accord, ni d'accord ni en désaccord, pas d'accord, Pas du tout d'accord"*, l'enquête a été diffusée via Google Forms.

3.2.4 Hypothèses

Les hypothèses dans une enquête présentent des déclarations claires, pour définir des résultats attendus lors de l'étude d'un phénomène sur une population spécifique. Pour définir les hypothèses de notre étude nous avons présenté des suppositions claires et logiques conformes avec l'objectif de l'enquête, afin d'arriver à des résultats précis. Nous nous sommes aussi basés sur les résultats précédents d'autres recherches (park et al., 2011 ; Briz et al., 2016 ; Nikou et al., 2016 ; Azizi et Khaton, 2019), pour à la fin arriver à dériver huit (08) hypothèses décrites dans la Table 3.3 :

Hypothèses	Variable dépendante	Variable indépendante
H1	Attitude	facilité d'utilisation
H2	Performance espérée	facilité d'utilisation
H3	Intention	facilité d'utilisation
H4	Attitude	Influence sociale
H5	Attitude	Conditions facilitatrices
H6	Performance espérée	Conditions facilitatrices
H7	Intention	Attitude
H8	Performance espérée	Attitude

TABLE 3.3: Description des hypothèses

- H1 : La facilité d'utilisation affecte l'attitude
- H2 : La facilité d'utilisation affecte la performance espérée
- H3 : La facilité d'utilisation affecte l'intention
- H4 : L'Influence sociale affecte l'attitude
- H5 : Conditions facilitatrices affectent l'attitude
- H6 : Conditions facilitatrices affecte la performance espérée
- H7 : L'attitude influence l'Intention
- H8 : L'attitude influence la performance espérée

3.2.5 Création du questionnaire et l'analyse des données

L'enquête a été rédigée dans un premier temps en arabe, deux (02) professeurs d'arabe du secondaire diplômé d'un master 2 ont évalué le questionnaire littérairement, puis deux (02) professeurs d'anglais du secondaire également diplômé d'un master 2 l'ont traduit séparément pour le présenter dans le papier associé à ce travail, ce questionnaire a été traduit en français pour la thèse. Ensuite un groupe de 15 étudiants testent le sondage dans une communication face à face pour éliminer toute ambiguïté, la Table 3.4 présente une partie de cette enquête.

Construction	ID	Question
Attitude	ATU01	Utiliser le mobile m'a motivé
	ATU02	Étudier avec l'apprentissage mobile est une bonne idée
	ATU03	Je suis intéressé par l'utilisation du mobile pour les études
	ATU04	L'utilisation du Algogo rend l'apprentissage plus intéressant
	ATU05	Utiliser un mobile pour l'apprentissage était amusant
Performance espérée	PE01	J'ai trouvé le système utile
	PE02	Utiliser le Mobile me permet de terminer rapidement
	PE03	Utiliser le Mobile améliore ma compréhension
	PE04	L'utilisation du Algogo me permet d'atteindre mon objectif pédagogique
	PE05	J'ai fait mieux que je ne le pense en utilisant le Mobile
	PE06	Je travaille plus dur avec le Mobile qu'avec l'ordinateur
Facilitating Condition	FC01	J'ai les ressources requises pour utiliser Algogo
	FC02	J'ai les connaissances requises pour utiliser Algogo
	FC03	L'assistance est disponible lorsque j'ai des difficultés avec Algogo
Behavior Intention	BI01	J'utiliserai le mobile pour apprendre
	BI02	Je décide d'utiliser le mobile pour apprendre d'autres cours
Ease of Use	EU01	Je suis capable d'utiliser le Mobile pour apprendre
	EU02	J'ai des expériences d'utilisation du mobile pour l'apprentissage
	EU03	Algogo est facile à utiliser
Social Influence	SI01	Je conseillerai mes collègues d'utiliser le mobile pour l'apprentissage
	SI02	Ma société encourage l'utilisation du mobile pour étudier
	SI03	Ma famille encourage l'utilisation du mobile pour étudier

TABLE 3.4: Partie du questionnaire

Afin d'analyser les données reçues des élèves, et étant donné que la recherche dépend fortement du comportement de l'apprenant, des techniques de modélisation d'équations

structurelles (SEM) ont été utilisées. Selon (Hox, Bechger, 2009), « SEM est une technique de modélisation statistique très générale, largement utilisée dans les sciences du comportement ». Il existe deux types de SEM (Tenenhaus, 2008), une modélisation d'équations structurelles basée sur la covariance et une autre basée sur les composants.

- Le SEM basé sur la covariance est utilisé pour la validation du modèle et nécessite un large échantillon.
- Le SEM basé sur les composants est utilisé pour le calcul du score et peut être effectué sur de petits échantillons.

Dans cette recherche, la technique PLS (Partial Least Squares), qui est une technique de modèle SEM basée sur les composants, a été choisie parce qu'elle correspondait aux besoins de la recherche et surtout à la taille de l'échantillon, la méthode PLS se concentre sur la prédiction d'un ensemble spécifique de relations hypothétiques qui maximise la variance expliquée dans les variables dépendantes (Hair et al., 2011). PLS a deux composants (Kock, 2010), le modèle de mesure représente la théorie qui spécifie comment les variables mesurées se combinent pour décrire la méthode, et le modèle structurel qui représente à son tour la théorie qui montre comment les constructions sont liées les unes aux autres. Pour arriver à notre but nous avons utilisé deux logiciels le SPSS (v.25 pour windows) et le SmartPls (v.3.2.8 pour windows) sont tous deux conçus pour la détermination de la valeur prédictive.

3.3 Résultats

Dans cette section, nous allons présenter les résultats statistiques de notre enquête, les présentations des résultats sont divisée en deux parties la première est consacrée aux statistiques descriptives, qui permettent de décrire l'état des réponses des apprenants sans obligatoirement déduire une relation claire entre eux. Par contre dans la deuxième partie, les composant PLS, vont nous présenter le modèle externe et interne qui vont exposer la fiabilité des résultats et des variables, ainsi que la liaison et sa robustesse entre les constructions pour déterminer si les hypothèses sont réellement correctes.

3.3.1 Statistiques descriptives

Le pourcentage des différents résultats de l'enquête est illustré dans la Table 3.5, basé sur une échelle de Likert de 05 points. 62,44% trouvent que la technologie mobile est facile à utiliser ; 52,89% des étudiants ont confirmé que les conditions facilitatrices surtout du soutien externe étaient élevées. 44,93% avaient une attitude positive envers la technologie. 47,56% déclarent que l'utilisation de cette technologie améliore leurs performances, 53,67% ont l'intention de l'utiliser à l'avenir et la recommander. 51,11% des étudiants avouent qu'ils avaient un soutien social pour utiliser la technologie mobile, de plus la la Figure 3.3 représente le pourcentage d'acceptation de OER choisie, on le trouve dans les questions dont les identificateurs sont ATU 04, PE01, PE04, FC02 et FC03, en calculant la moyenne on peut admettre que 50.39% ont été très d'accord avec le choix de la ressource éducative (les vidéos et Algogo Mobile)

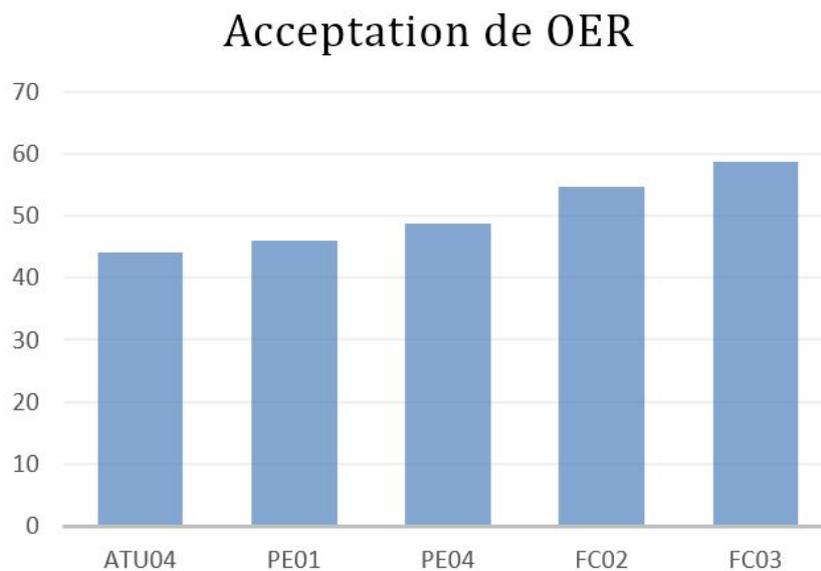


FIGURE 3.3: L'acceptation du type d'OER choisi

Acceptance level	PE	EU	ATU	BI	FC	SI
(Tout à fait d'accord, d'accord) %	47,56	62,44	44,93	53,67	52,89	51,11
Tout à fait d'accord %	17,33	38,44	18,27	28	23,56	22,22
D'accord %	30,22	24	26,67	25,67	29,33	28,89
Ni d'accord ni en désaccord %	21	18	21,60	13,67	18,89	19,78
Pas d'accord %	20,56	8,44	19,73	15,67	17,56	14,22
Pas du tout d'accord %	10,89	11,11	13,73	17,00	10,67	14,89

TABLE 3.5: Pourcentage des réponses de chaque construction

3.3.2 Composants PLS

Dans cette section nous avons suivi la démarche et les critères présentées par (Brise et al. 2016).

— Le modèle de mesure (modèle externe)

Le modèle externe présente une analyse de la fiabilité de l'échelle et la validité des constructions. Tout d'abord, il est nécessaire de calculer le chargement externe, où le résultat accepté devait dépasser 0,5. Les facteurs de chargement doivent dépasser 0.5 pour être significatifs, les T-statistiques doivent être supérieures à $50\% \pm 1.96$ au niveau de 5%, et pour la fiabilité composite, elles doivent être supérieures à 0.7. On remarque que le résultat de tous les éléments de construction a été accepté sauf EU1 où la fiabilité de l'indicateur se réfère à 0.471 (Table 3.6). L'alpha de Cronbach est utilisé pour examiner la cohérence interne et la fiabilité, selon (Nunnally 1978), la valeur alpha est acceptée si elle est égale ou supérieur à 0.70, mais cela dépend du domaine de la recherche, dans cette étude, l'alpha de Cronbach de toutes les constructions est supérieur à 0.777 (Table 3.6), ce qui signifie que la cohérence interne des données est acceptable.

Pour la validité du concept, deux valeurs doivent être vérifiées, la première est la validité convergente; cette valeur est mesurée par la variance moyenne (AVE), la valeur d'acceptation doit être égale ou supérieure à 0,5 est c'est le cas de cette recherche (Table 3.6). La deuxième valeur est la validité discriminante, il est neces-

saire de comparer la racine carrée de l'AVE avec la corrélation factorielle, comme présentée au (Table 3.7), la racine carrée de chaque AVE est supérieure à sa corrélation factorielle, ce qui signifie que l'enquête présentée montre une bonne validité de discernement.

Construct	Indicator	Factor loading	T-statistics	Indicator reliability	Composite reliability	Cronbach's Alpha	Convergent validity AVE
Attitude (ATU)	ATU01	0.924	25.677	0.892	0.974	0.967	0.884
	ATU02	0.946	45.695	0.972			
	ATU03	0.932	26.123	0.887			
	ATU04	0.957	31.190	0.911			
	ATU05	0.944	39.023	0.959			
Performance espérée (PE)	PE01	0.968	31.13	0.943	0.979	0.975	0.890
	PE02	0.933	24.052	0.890			
	PE03	0.944	23.303	0.904			
	PE04	0.957	37.899	0.938			
	PE05	0.958	32.804	0.909			
	PE06	0.901	21.850	0.998			
Conditions facilitatrices (FC)	FC01	0.916	22574	0.882	0.926	0.880	0.806
	FC02	0.863	12573	0.764			
	FC03	0.915	23.724	0.883			
Intention	BI01	0.977	29.566	0.939	0.976	0.952	0.954
	BI02	0.978	35.735	0.968			
Facilité d'utilisation (EU)	EU01	0.731	2.627	0.471	0.867	0.777	0.686
	EU02	0.917	8.366	0.981			
	EU03	0.830	5.5533	0.739			
Influence sociale (SI)	SI01	0.838	10.936	0.943	0.888	0.816	0.725
	SI02	0.856	7.043	0.730			
	SI03	0.866	4.851	0.591			

TABLE 3.6: Le modèle de mesure

	EU	FC	FI	ATU	PE	SI
EU	0.829					
FC	0.340	0.898				
FI	0.334	0.433	0.977			
ATU	0.282	0.606	0.539	0.941		
PE	0.356	0.736	0.474	0.764	0.944	
SI	0.223	0.240	0.416	0.425	0.343	0.852

TABLE 3.7: La valeur de la validité discriminante

— Le modèle structurel (modèle interne)

Dans cette partie, il est nécessaire de calculer le coefficient du chemin, qui permet

de valider les hypothèses présentées au début de l'enquête. Le résultat est présenté dans la Table 3.8, comme nous l'avons remarqué, toutes les hypothèses sont supportées (Y), à l'exception de deux hypothèses : entre EU et PE, et EU et ATU car toutes les valeurs T-statistique sont supérieures à $50\% \pm 1.96$ au niveau de 5%, sauf pour la relation entre EU et PE et la valeur et EU et de l'ATU ou les valeurs sont respectivement 1,409 et 0,556.

Hypothèses	coefficient du chemin	T-statistiquess	supporter
ATU -> BI	0.484	6.569	Y
ATU -> PE	0.493	6.165	Y
EU -> ATU	0.040	0.556	N
EU -> BI	0.196	2.981	Y
EU -> PE	0.077	1.409	N
FC -> ATU	0.526	6.229	Y
FC -> PE	0.411	4.658	Y
SI -> ATU	0.287	3.938	Y

TABLE 3.8: Le modèle structurel

Pour indiquer la force des variables indépendantes on calcule la valeur f^2 la Table 3.9 affiche les résultats. Et finalement le paramètre Q^2 est calculé dans la Table 3.10 pour fournir la pertinence prédictive du modèle, tous les résultats sont positifs, ce qui confirme sa pertinence prédictive.

Construct	ATU	BI	PE
ATU		0.495	0.464
EU	0.009	0.219	0.070
FC	0.566	0.280	0.723
SI	0.310	0.153	0.144

TABLE 3.9: L'effet des variables indépendantes

Construct	Predictive relevance(Q^2)
ATU	0.369
EU	0.290
FC	0.582

TABLE 3.10: Pertinence prédictive (Q^2)

3.4 Discussion

Cette recherche vise à étudier le comportement et l'acceptation des élèves du secondaire de la technologie mobile en associant les ressources éducatives et libres, la plupart des études récentes tentent d'examiner l'acceptation des technologies éducatives en raison de l'influence positive de ce domaine sur l'apprentissage (Azizi et Khaton, 2019 ; Sunar et al. 2020). En outre, l'utilisation des technologies aide les enseignants et les apprenants à avoir une meilleure expérience éducative, où on trouve que la plupart des pays développés utilisent plusieurs types de technologies éducatives, selon (Hamidi et al. 2011), les pays européens utilisent des technologies de l'information avancées pour l'éducation, tandis que les systèmes éducatifs des pays africains ne sont pas basés sur les technologies ; et c'est toujours le cas de nos jours. L'utilisation des technologies éducatives en Afrique est assez rare et lente. Cependant, commencer à choisir les bonnes ressources ouvertes et libres et les combiner avec une technologie familière pour les étudiants (par exemple, le M-learning) aidera à atteindre une meilleure qualité d'apprentissage.

Dans cette étude, la plupart des apprenants sont familiarisés avec le mobile ; 95,33% possèdent un ou plusieurs appareils mobiles, ce qui est similaire à de nombreuses recherches (Boruff et Storie, 2014, Briz et al., 2016). 44,93% des apprenants avaient une attitude positive envers la technologie, ce qui donne à l'élève l'intention de l'utiliser à l'avenir. (Liaw et Huang, 2011), déterminent que l'attitude envers l'apprentissage mobile offre l'opportunité de développer un environnement de M-Learning approprié et joue un rôle important pour l'acceptation de ces systèmes. (Mahat et al., 2012) indiquaient que le M-learning considérablement accepté par les étudiants malaisiens, même chose pour (Abu Al-Aish, 2014), qui a révélé que les étudiants anglais avaient une attitude positive envers le M-learning et étaient prêts à l'utiliser à nouveau.

Les résultats de l'évaluation des hypothèses montrent que l'attitude de l'élève est essentielle pour prédire l'espérance de performance, ce qui s'explique par la façon dont l'attitude positive fait croire à l'individu que le système aide à atteindre l'objectif, ce résultat est également montré dans (Park et al., 2011). L'attitude a également un effet significatif sur

l'intention, ce résultat est en accord avec (Hadadgar et al., 2016 ; Azizi et Khaton, 2019). L'étude a montré que l'intention est également effectuée avec la facilité d'utilisation et cela était aussi présentée dans les résultats précédents (Azizi et Khaton, 2019), (Dissanayake et Velananda, 2020) expliquent cette hypothèse selon eux, les apprenants adopteront plus le M-learning s'ils croient qu'il est facile à utiliser, mais l'attitude exerce plus d'influence que la facilité d'utilisation sur l'intention du comportement (Table 3.8). Idem pour la performance espérée où l'attitude et les conditions facilitatrices ont plus d'influences que la facilité d'utilisation sur elle. Les conditions facilitatrices ont un effet significatif à la fois sur l'espérance de performance et l'attitude, comme le montre le (Table 3.9), selon (Cohen, 1992), la taille de l'effet est petite si la valeur est d'environ 0,1, moyenne si elle est d'environ 0,3 et grande si la valeur dépasse 0,5. La facilité d'utilisation a un impact positif sur l'attitude de l'utilisateur, ce qui est en accord avec (Davis et al. 1989 ; Briz et al. 2016). Cette recherche a prouvé également que l'influence sociale affecte positivement l'attitude chez les lycéens est cela est causé par l'influence élevé des parents et des amis sur cette tranches d'âges. Cette étude a révélé l'intention (BI) des apprenants est largement prévisible par leur attitude et elle est affectée par la facilité d'utilisation, la condition facilitatrice et l'influence sociale.

Limites de M-Learning selon les résultats de l'enquête

Dans le chapitre précédant nous avons entamé les limites de M-learning du point de vue des différents chercheurs dans la littérature, mais dans ce chapitre nous allons voir les limites selon le point de vu des apprenants qui ont essayé cette technologie pendant plus de trois semaines. Nous avons ajouté une 33^{ème} question dans l'enquête de types question ouverte, afin de définir les obstacles vécus par les apprenants. Selon cette étude, les apprenants du secondaire semblent préférer utiliser la technologie dans le processus d'apprentissage, mais certains ont demandé de l'aide pour éviter la distraction, surtout des réseaux sociaux, naturellement, les élèves font face à la distraction en utilisant un appareil qui contient leurs messages, appels et jeux, etc. D'autres ont parlé de la taille de l'écran surtout ceux qui n'ont pas utilisé une tablette, vu que les élèves du secondaire

utilisent leurs téléphones pendant longtemps (30%) plus de 5 heures / jour, mais (18%) des apprenants confirment ne jamais utiliser le mobile pour apprendre à cause de la taille de l'écran. La batterie était un point préoccupant pour la plupart des élèves les étudiants s'interrogent sur la batterie la meilleure façon pour l'optimiser, et est-ce que cela va affecter sa performances par la suite, Shudong et Higgins (2006) ont confirmé que lorsque l'étudiant utilise son téléphone pour étudier, l'écran du téléphone reste déverrouillé pendant toute la période d'étude, ce qui fait que la batterie se décharge rapidement. Une dernière préoccupation était centrée autour du stockage, les étudiants se sont inquiétés de la quantité de données qui peuvent être associées à un cours en ligne, ils se demandent comment utiliser toutes les données sans les télécharger dans leurs appareils, à la fois applications et données.

Pour résumer les résultats de la recherche, les données analysées présentées dans cette contribution ont mis en évidence la cohérence structurelle et la fiabilité de l'étude, cela explique que les élèves du secondaire préfèrent utiliser les M-learning tout en entamant certaines préoccupations, de plus l'utilisation des deux types de ressources éducatives libres (vidéos et logiciels) ajoute une valeur remarquable à l'efficacité de l'apprentissage. Cette recherche permet de créer de nouvelles opportunités pour les apprenants, leurs permet d'être plus impliqués dans le processus d'apprentissage et de profiter de l'expérience du M-learning avec leurs enseignants dans la classe pour les aider à pratiquer de plus en plus l'apprentissage en ligne. La contribution de la deuxième partie présente l'architecture d'un système éducatif personnalisé basé sur le Cloud mobile pour surmonter les limites vécues par les élèves dans cette enquête.

II. Architecture basée sur MCC pour l'apprentissage auto-adaptatif

Le succès de l'apprentissage mobile en ligne augmente rapidement de jour en jour, les chercheurs de ce domaine sont très intéressés par cette technologie, qui offre un apprentissage à la demande, mais comme nous avons présenté dans la partie précédente, les apprenants souffrent de quelques limites lors de l'utilisation de M-learning et ces limites peuvent créer un recul ou un abandon de cours. Un type de cours formel connus sous le nom du MOOC (Massive Open Online Courses), considéré comme une ressource éducative libre basée sur les vidéos et les quiz, englobent des objectifs intellectuels et éducatifs, qui offrent diverses possibilités d'apprentissage tout au long de la vie, peut être très efficace pour résoudre quelques obstacles de M-learning. Cependant, de nombreuses limitations, telles que la diversité des apprenants, le manque de motivation, etc., ont affectés les résultats des apprenants en MOOC, ce qui a malheureusement augmenté le taux d'abandon. La combinaison de la mobilité du M-learning avec la formalité des MOOC dans des systèmes d'apprentissage intelligents peuvent délivrer des résultats prometteurs.

La voie de personnalisation dans l'apprentissage intelligent, est très vaste est comprend beaucoup de types de technologies, ce qui nous a amené vers une spécification de personnalisation afin de résoudre les quatre problèmes discutés avec les élèves du lycée (Taille d'écran, la batterie, le stockage et la distraction), et aussi d'adapter les plateformes selon les besoins et le contexte des apprenants. Dans cette partie nous allons proposer l'architecture d'un service Cloud mobile pour la personnalisation de l'apprentissage intelligent, afin de dépasser les limites de M-learning.

3.5 Relation entre M-learning et MOOC

La technologie M-learning et les cours MOOC, permettent tous deux de créer un apprentissage en ligne qui sert à faciliter et à fournir une éducation de qualité aux apprenants, mais les deux souffrent de quelques limites. D'une part le problème principal des MOOC, et le taux d'abondons, causé principalement par la non-personnalisation de cours ou de la plateforme de diffusion. D'autre part les M-learning présentent des en-

vironnements d'apprentissage où l'appareil collecte des informations importantes sur le contexte de l'apprenant mais ces mêmes apprenants souffrent de la distraction et de la petite taille de l'écran qui ne permet pas de passer des heures continues d'apprentissage. Donc la combinaison de ces deux technologies permet d'établir des cours formels basés sur des vidéos de courtes durées bien organisés dans des environnements personnalisés selon le contexte de l'utilisateur, collecté par les capteurs mobiles, permettant d'élargir la portée de l'éducation grâce à la caractéristique de mobilité, qui consiste à accroître la capacité d'un apprenant à déplacer son environnement physique d'apprentissage librement (Barbosa et Geyer, 2005). (Waard, 2013) ont suggéré l'utilisation du micro-contenu MOOC qui représente des cours de temps très réduit et de même valeur d'apprentissage, lors de l'utilisation de M-learning technologie pour assurer une meilleure interaction de la part des apprenants.

Dans la section suivante nous allons présenter le Context-Driven Mobile Learning, qui présente un apprentissage intelligent et personnalisé, dans le but d'améliorer l'environnement d'apprentissage mobile. L'utilisation des MOOC, permet de formaliser les cours présentés, ce qui aide à diminuer la distraction. De plus le principe de base des MOOC c'est l'utilisation de vidéos comme ressources pédagogiques, regarder les vidéos est un acte habituel chez les apprenants, la plus part des propriétaires du smartphone passent beaucoup de temps à regarder des vidéos en ligne selon (smith, 2020) « Nous regardons plus d'un milliard d'heures de vidéos YouTube par jour, ... 70% de ces vues proviennent d'appareils mobiles. ». Cette habitude va aider à dépasser l'obstacle de la taille de l'écran si les vidéos étaient bien personnalisées selon les besoins de l'apprenant. Le CD-Mobile MOOC, utilise aussi la technologie Cloud mobile qui prouve ces performances dans tous les domaines où ils sont présents, à cause de leur avantages majeurs, où ils réduisent la dépendance au matériel ce qui permet d'optimiser la batterie et la capacité de stockage, tous les problèmes de maintenance et de perte de données. Dans ce qui suit, nous proposons une solution pour combiner ces trois domaines afin d'obtenir un système d'apprentissage sensible au contexte qui optimise les ressources et aide les apprenants lors de leurs processus d'apprentissage pour éviter le scénario d'abandons.

3.6 Architecture CD-Mobile MOOC

Dans cette section nous allons présenter l'architecture du modèle de l'apprentissage intelligent en ligne basée sur la combinaison des cours formel MOOC et de la technologie M-learning dans les Mobile Cloud Computing afin de personnaliser l'apprentissage mobile.

3.6.1 Processus de conception de l'architecture CD-Mobile MOOC

Nous nous sommes concentré d'abord sur les cours formel MOOC, en extraire les fonctionnalités de base de ces cours selon ce qui est présenté dans les différentes plateformes MOOC, Un apprenants qui s'inscrive dans une plateforme MOOC serait capable de :

- S'identifier : cette fonctionnalité permet d'afficher l'interface de l'utilisateur
- Editer le profil : le participant peut changer le nom, le mot de passe et les préférences linguistiques, etc.
- S'inscrire dans un MOOC : choisir depuis un ensemble de cours celui qui intéresse le participant, souvent dans la page d'accueil de l'interface utilisateur, l'apprenant trouve des MOOC recommandés selon ses préférences.
- Suivre le cours : Dans cette phases l'apprenant peut voir des vidéos, répondre au quiz, pour s'auto évaluer.
- Évaluer avec les autres : cette fonctionnalité permet d'avoir un apprentissage collaboratif en échangeant des informations, et en utilisant les réseaux sociaux, et le forum.

Les fonctionnalités des MOOC, avec l'organisation des Cours (vidéo, texte, quiz et examen) lui donnent une valeur de plus par rapport aux autres cours en ligne, car les informations présentées proviennent des enseignants de grandes universités qui suivent une pédagogie d'apprentissage lors de la création des vidéos et des questions et les différentes pratiques. Presque tous les MOOC, utilisent les vidéos comme la méthode principale de diffusion de contenu (Hansch et al., 2015), de plus les vidéos sont de durée courte et focalisent sur une seule unité de cours à la fois, ce qui permet de maximiser l'engagement de l'apprenants (Guo et al., 2014). L'engagement est défini par (Csikszentmihalyi,

1990), comme un état où l'individu est tellement impliqué dans une activité que toutes les distractions sont bloquées. Ce qui a notre avis va permettre de diminuer le taux de distraction lors de l'utilisation de mobile.

Nous avons divisé le fonctionnement du système en trois parties complémentaires principales Figure 3.4 :

- L'acquisition : cette fonctionnalité permet de collecter les différentes informations de contexte de l'environnement depuis les capteurs physiques de l'appareil mobile, et elle permet aussi de rassembler des informations sur le contexte de l'utilisateur et ses traces pour les sauvegarder sur les bases de données du Cloud, afin de créer une vue claire sur les préférences de l'apprenant.
- Le traitement : Dans cette étape, le système reçoit des données sur le contexte de l'apprenant et des traces de son interaction avec l'interface et les capteurs mobiles, puis un processus consécutif via plusieurs modèles sera généré pour adapter la plateforme et les cours aux préférences de l'apprenant, cette étape présente un travail combinatoire de différents modèles, chacun est chargé de déterminer la bonne adaptation d'un contenu spécifique.
- Personnalisation : cette étape présente le modèle final qui permet d'offrir le contenu le plus spécifique à l'apprenant en fonction de ses préférences, de son environnement et de ses besoins. Le moteur d'adaptation reçoit le modèle de profil et le modèle de cours en entrée, ainsi que le style d'apprentissage de chaque utilisateur, son niveau et sa catégorie qui permettent d'adapter la plateforme d'apprentissage et le contenu, et génère une sortie personnalisée du système.

Toutes les étapes précédentes fonctionnent sur le Cloud Mobile, qui est défini dans cette contribution comme une technologie qui offre des services optimaux pour les utilisateurs afin de libérer les appareils mobiles des apprenants et des enseignants du traitement et du stockage.

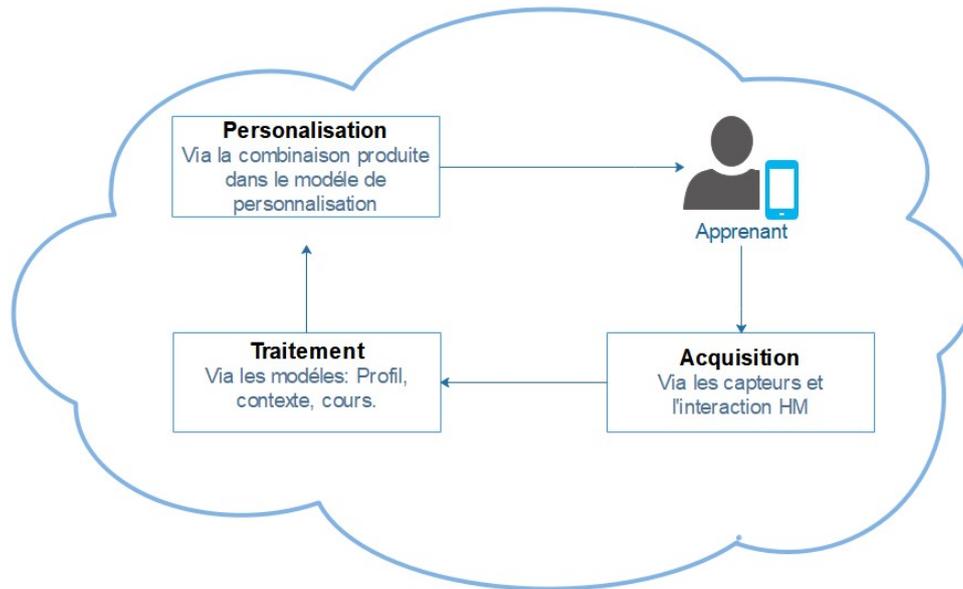


FIGURE 3.4: Les étapes de fonctionnements de l'architecture

3.6.2 Description de l'architecture CD-mobile MOOC

L'architecture du CD-Mobile MOOC (figure 3.7) est basée sur quatre niveaux principaux. Le premier niveau est celui de l'apprenant qui utilise un appareil mobile pour accéder à la plateforme associée avec le modèle CD-Mobile MOOC afin d'assister des cours personnalisés sur une plateforme qui s'adapte à ses préférences et son contexte. L'interaction entre l'apprenant et le système se produit dans le deuxième niveau qui concerne l'interface utilisateur, permet de collecter des données sur l'appareil de l'apprenant, son contexte et ses traces afin de créer le module de traçabilité, et encore de spécifier le comportement général de l'apprenant, qui aide à définir sa catégorie et son niveau. Les données collectées sur cette phase représentent le matériel de fonctionnement du troisième niveau qui représente à son tour un ensemble de modèles de traitement et de personnalisation, ces modèles permettent de s'occuper de la gestion des données extraites du comportement de l'utilisateur et de son contexte. Le dernier niveau est celui du stockage de données.

Le Premier niveau (l'utilisateur) :

Il existe un degré élevé de variabilité dans le comportement des apprenants dans un massive open online course (DeBoer et al., 2013), en raison du nombre massif de ces derniers et de leurs natures différentes (âge, sexe, niveau éducatifs, distribution géo-

graphique, etc.). Sur les plateformes éducatives tel que edX, nous pouvons trouver des informations sur cette variabilité, nous avons extrait de cette plateforme les figures (figure 3.5 et figure 3.6), qui représente consécutivement la distribution géographique et le niveau éducatif des apprenants inscrits, et comme nous pouvons le constater la diversité est énorme. Ce qui rend l'utilisation d'un seul scénario d'apprentissage un facteur majeur d'abandon. Un autre obstacle rencontré lors de l'interaction avec les apprenants en plus de leurs variabilités est la motivation, l'attention des apprenants est difficile à obtenir et encore plus difficile à conserver (Hansch et al., 2015), ce qui présente la personnalisation du contenu comme une solution assez intéressante qui promet beaucoup d'amélioration dans ce domaine.

Dans notre modèle l'apprenant va utiliser son appareil mobile pour accéder à la plateforme éducative qui utilise la technologie M-learning, et contient des MOOC. Dès sa première interaction l'utilisateur remplit un formulaire qui contient des informations sur son âge, son niveau éducatif, etc. Pour donner au système une image sur lui, après cette phase cet utilisateur devient un apprenant lorsqu'il s'inscrit au niveau d'un MOOC et là les différents modules de traitement et de personnalisation gèrent ses informations pour lui fournir le meilleur scénario possible qui s'adapte avec son contexte et ses préférences. L'utilisation de la technologie M-learning avec les MOOC permet aux systèmes d'extraire des informations sur le contexte de l'utilisateur et de son environnement via les capteurs physique du mobile, dans cette proposition nous nous sommes basés sur trois capteurs physiques, le GPS, le microphone et la caméra. Le GPS (Global Positioning System) qui représente un système de positionnement par satellite, permettant de connaître le positionnement d'un appareil partout dans le monde, cette fonctionnalité permet de localiser l'apprenant s'il est à la maison ou en dehors et enregistre cette information dans la base de données du capteur pour l'utiliser dans la phase de traitement. De même, pour le microphone qui représente un appareil capable de convertir un signal acoustique en signal électrique, le principe de l'utilisation de cette fonctionnalité dans notre architecture est de définir le niveau de bruit qui entoure l'utilisateur et ainsi le stocker sur la même base de données de capteur. Pour la caméra l'utilisation de cette dernière permet de capturer

le visage de l'apprenant pour déterminer son niveau d'engagement. Afin de profiter des avantages de l'architecture CD-mobile MOOC, l'apprenant doit être abonné aux services Cloud du système pour sauvegarder ses cours et notes à distance, et aussi pour bénéficier des économies d'énergie.

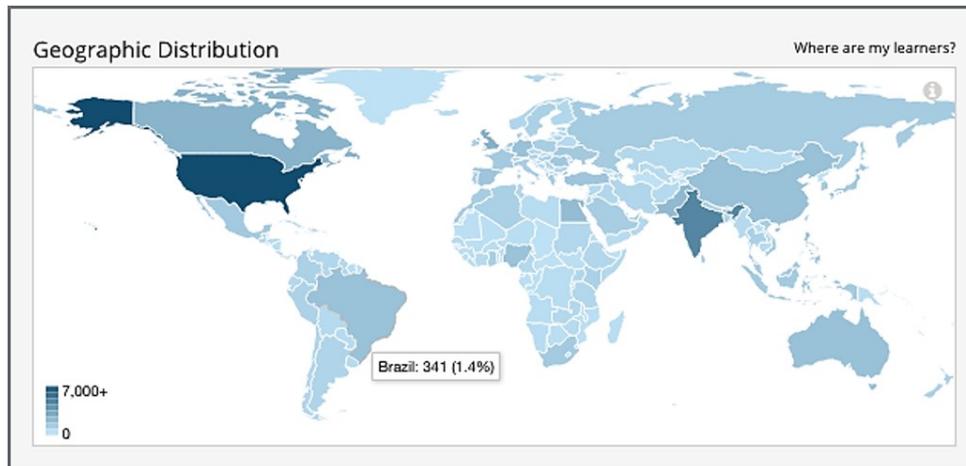


FIGURE 3.5: Distribution géographique

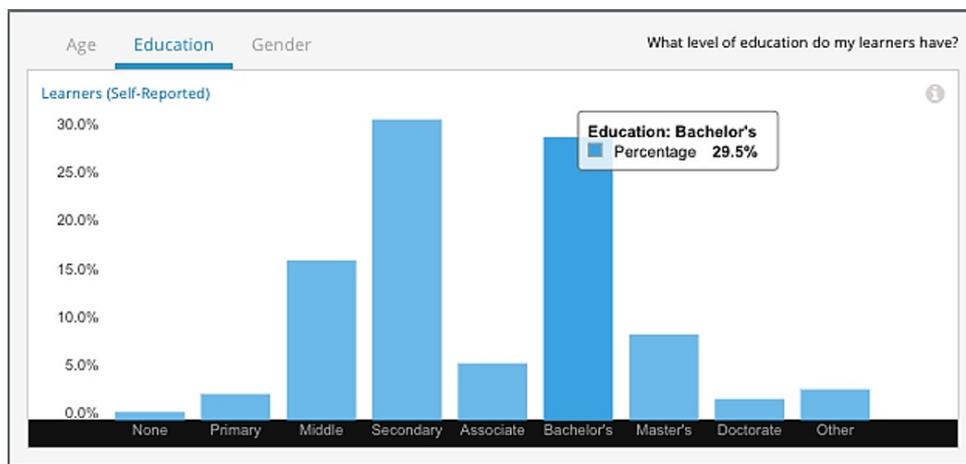


FIGURE 3.6: Le niveau éducatif des apprenants

Le deuxième niveau (l'interface) :

L'interface présente un ensemble des dispositifs matériels et logiciels permettant à un utilisateur humain de contrôler et de communiquer avec une machine. Lorsqu'un apprenant s'inscrit au niveau de la plateforme il devient dans une interaction directe avec l'interface de l'architecture CD-Mobile MOOC, il remplit le formulaire qui contient ses informations de base, ensuite il s'inscrit au niveau du MOOC choisi. Pour toutes les autres

connexions l'apprenant s'identifie au niveau de l'interface et commence son processus d'apprentissage. Le module de l'interface sauvegarde sur la base de données contexte et traces situées dans le Cloud des informations concernant l'état de l'appareil de l'apprenant, ses interactions et traces. Dans ce système les informations retirées de l'appareil sont le débit de connexion pour décider sur le types de cours le plus adéquat pour une faible ou une forte connexion, et la détection de la langue de système d'exploitation afin d'adapter la langue de la plateforme avec celle du système, ces deux informations sur le contexte de l'utilisateur sont manipulées dans le modèle de contexte temps réel, afin de générer des règles conditionnelle appliquées tout au long du processus de traitement. A ce niveau aussi le système collecte les traces des apprenants ces traces sont sauvegardée sur la base de données du contexte et traces et sont manipulées au niveau du module traçabilité dans le modèle d'adaptation du cours, Les traces permettent de définir le niveau d'activation de l'apprenant, sa catégorie, son comportement envers les quiz et les vidéos etc.

Le troisième niveau (modules de traitement et de personnalisation) :

Ce niveau est responsable du traitement des informations acquises des niveaux précédents pour garantir le résultat de personnalisation attendu, nous allons définir donc les quatre modèles représentatifs de ce niveau.

1). Le module du contexte temps réel :

Lorsque l'apprenant se connecte, de nombreuses données changent par rapport à la connexion précédente, tels que la vitesse de connexion de l'internet, la langue du système, son endroit et le niveau de bruit dans son environnement. Pour un meilleur résultat, l'architecture présente le modèle de contexte temps réel qui contient des données relatives à court terme, calculées à la volée pour chaque connexion. Ce modèle est diviser en deux parties, la première parties reçoit des informations depuis la base de données de capteurs et la deuxième partie reçoit des informations depuis la base de données du contexte et traces.

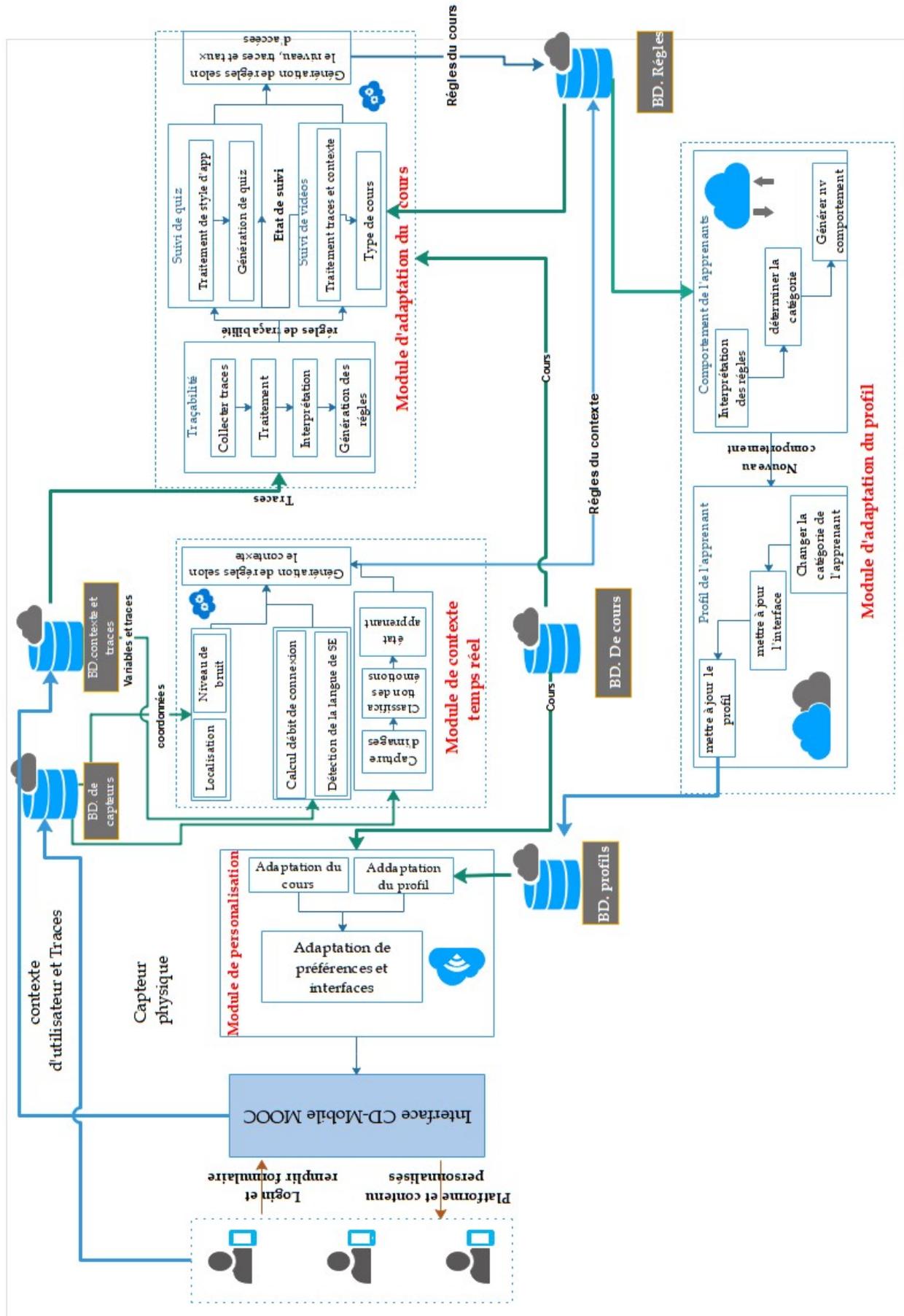


FIGURE 3.7: Architecture de CD-Mobile MOOC

- La première partie extraite à la volée des informations sur le contexte de l'environnement de l'apprenant, dans cette contribution nous nous sommes basés sur trois variables du contexte d'environnement.

Le premier consiste à définir la localisation de l'apprenant, depuis le capteur GPS, si ce dernier se trouve en dehors de la maison le système va calculer le niveau de bruit dans son entourage par la deuxième variable (le microphone), et lui propose un cours en texte à la place de la vidéo, s'il se trouve dans un environnement brouhaha, l'apprenant peut à tout moment revenir vers le cours en vidéo.

La troisième variable est les mouvements faciales extraites depuis la caméra de l'utilisateur, elle capture le visage de l'apprenant à chaque instant, si ses expressions faciales changent, l'image capturée est classifiée pour définir le nouvel état émotionnel, nous détaillerons la partie de classification dans le chapitre suivant.

- La deuxième partie s'intéresse à l'appareil de l'apprenant, nous nous sommes basés sur deux variables :
 - La première variable permet de calculer le débit de connexion :
si les apprenants se trouvent avec une connexion très faible qui mène vers un suivi plein d'interruption, le système lui propose un cours en texte ou une présentation qui contient le même contenu.
 - La deuxième variable concerne la langue de la plateforme éducative :
un acte très important lors de la personnalisation d'un système c'est d'adapter sa langue avec celle de l'apprenant, nous avons choisi la langue du système d'exploitation de l'appareil de l'apprenant en considérant que c'est la langue qui le met le plus à l'aise.

Les informations extraites de ce modèle représente des règles conditionnelles sauvegardées sur la base de données des règles.

2). Le module d'adaptation du cours :

Afin de personnaliser un système d'apprentissage, il faut non seulement personnaliser la plateforme mais aussi il faut faire attention au cours présenté, pour cette raison nous avons créé ce modèle qui contient trois modules :

○ La traçabilité :

Pour bien distinguer les préférences des apprenants et analyser leurs comportements, il est nécessaire de gérer leurs traces. Chaque clic sera enregistré dans la base de données, contexte et trace pour évaluer le progrès et les activités des apprenants. (Champin et al., 2013) présentent un module de traçabilité qui détermine les informations nécessaires sur une trace, dont le module utilisé dans cette contribution est inspiré. La définition du concept de traçabilité présentée comme suit :

- Quoi : identifier les actions.
- Quel : Identifier la ressource.
- Qui : Identifier l'identifiant de l'apprenant.
- Quand : Identifier l'heure.
- Comment : de quelle manière l'apprenant utilise-t-il la ressource.
- Depuis où : Identifier la dernière consultation.

Chaque activité effectuée par un apprenant sera enregistrée dans la base de données et interprétée par des règles conditionnelles, sauvegardé dans la base de données des règles.

Pour bénéficier de ce concept, le modèle de traçabilité d'abord collecte les traces depuis la BD contexte et traces, ensuite il les fait subir à une phase de traitement, qui permet de présenter les informations déjà collectées d'une manière interprétable, et puis la phase d'interprétation permet de donner un sens aux traces afin de les exploiter d'une manière correcte. Enfin la phase de génération des règles d'utilisation, qui permet de générer les règles selon l'interprétation des traces qui vont par la suite être utilisées pour personnaliser les autres modules.

○ Le suivi de vidéo :

Cette phase permet de déterminer la façon dont l'apprenant a suivi le cours, et est ce qu'il a terminé la vidéo entièrement ou pas. Ensuite après avoir traité le suivi de l'apprenant ce module génère les règles conditionnelles qui déterminent le type de cours approprié selon le rythme d'apprentissage, les cours sont obtenus à partir de la base de données de cours.

○ Le suivi de quiz :

Cette phase est responsable de la génération de quiz obtenu à partir de la base de données de cours, après avoir suivi le style d'apprentissage, ce module détermine la progression des apprenants afin d'améliorer leurs niveau en exploitant les traces acquises du module de traçabilité, le traitement dans ce module propose quatre catégories :

- Résultat entre 0%-40% au quiz : les apprenants dans cette situation sont obligés de revoir la vidéo.

- Résultat entre 41%-70% au quiz : les apprenants dans cette situation sont invités à lire un document ci-joint pour renforcer les informations de la vidéo.

- Résultat entre 71%-90% au quiz : une explication des questions sans réponse ou mal répondues est présentée.

- Résultat entre 91%-100% dans le quiz : Redirection vers les forums pour répondre et produire des questions.

NB. Le pourcentage fourni est calculé en fonction des appréciations de diplômés du lycée algérien.

3). Le module d'adaptation du profil :

Ce module (Figure 3.8), permet de personnaliser le profil de l'apprenant selon son contexte et son progrès, en recevant des informations depuis la base de données des règles. Il est divisé en deux parties le comportement de l'apprenant et le profil de l'apprenant. Le module de comportement est responsable de l'interprétation et l'application des règles conditionnelles, qui permet de détecter des nouvelles préférences traduites par un nouveau comportement, ce nouveau comportement contient des informations sur la catégorie de l'apprenant et les données de différentes variables contextuelles. Ensuite cette information circule vers le module profil qui à son tour change la catégorie de l'apprenant si c'est nécessaire et mettre à jour l'interface, le profil et sauvegarde ces informations sur la base de profil.

Les apprenants dans un MOOC peuvent être segmentés selon plusieurs catégories, les apprenants appartiendront donc à l'une de ces quatre catégories :

- Inactif : inscrit sur la plateforme mais ne se connecte jamais, le système ne réalise ce type d'apprenants que lorsqu'ils atteignent un certain seuil, dans ce cas, le système envoie un e-mail dans lequel il propose une liste de choix pour répondre à la question : quel est le problème rencontré sur la plateforme ? Les réponses collectées permettent d'améliorer l'expérience des apprenants.

- Actif : apprenants qui réussissent le MOOC tout en participant activement au forum.

- Moins fréquents : les apprenants qui fréquentent de moins en moins les MOOC, des e-mails de motivation sont souvent envoyés (par exemple, la bonne nouvelle est révélée, rejoignez-nous), lors de leur consultation, le système propose le meilleur scénario pédagogique possible.

- Les apprenants qui utilisent simplement des ressources.

Pour chaque catégorie, un algorithme sera appliqué ; il prend en entrée les traces, les informations de profil, la valeur des variables de contexte et les règles proposées par les autres modèles et délivre en sortie la catégorie qui convient à ce type d'utilisateur.

4). *Le module de personnalisation :*

Le modèle de personnalisation représente la phase finale qui regroupe toutes les informations et les règles conditionnelles provenant des modèles précédents. À partir des éléments d'entrée (base de données de profil et de cours), le système catégorisera tous les utilisateurs, il adapte le profil et les cours, selon le progrès des apprenants, leur contexte et leur niveau d'interaction, ce qui donne comme résultat un changement au niveau de l'interface et des préférences. De cette manière, le système construit un scénario pédagogique unique pour chaque apprenant. Les règles conditionnelles peuvent être modifiées par l'administrateur de la plateforme en fonction d'une demande des formateurs.

Le quatrième niveau (le stockage) :

Ce niveau est celui de stockage, dans notre système nous avons profité des services du Cloud mobile, le modèle présente cinq bases de données ; chacune est responsable d'un type de données :

— Base de données de capteurs : sauvegarde les données provenant des capteurs phy-

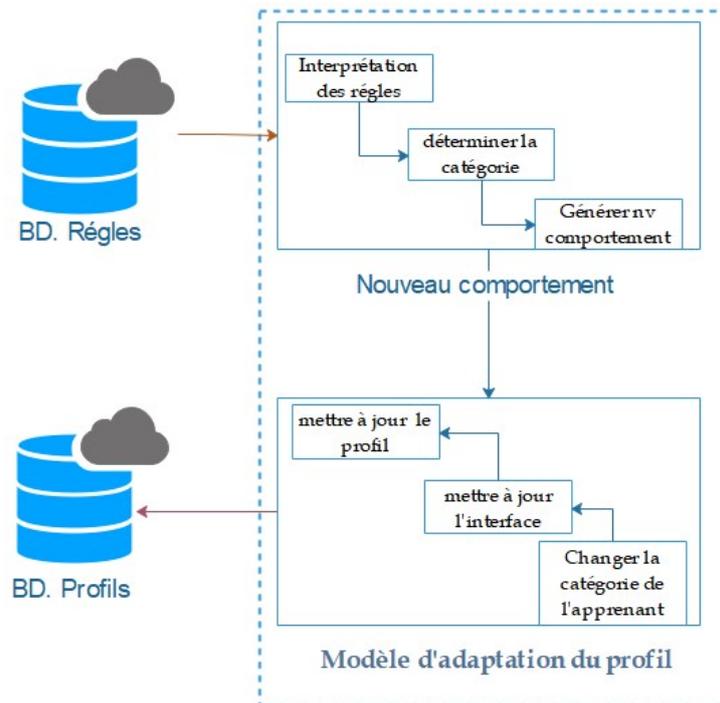


FIGURE 3.8: Le modèle d'adaptation du profil

siques.

- Base de données de contexte et trace : sauvegarde les données provenant de l'interface.
- Base de données de cours : sauvegarde les données provenant des enseignants et des apprenants.
- Base de données de règles : sauvegarde les données provenant du traitement conditionnel des différents modèles.
- Base de données de profil : sauvegarde les données sur le profil de l'apprenant, des informations statiques comme (nom, date de naissance, ..) et dynamiques comme ses préférences et sa catégorie.

3.6.3 Personnalisation selon CD-Mobile MOOC

Cette section décrit les points de personnalisations selon le processus CD-Mobile MOOC en utilisant le contexte des utilisateurs. Le MOOC étant un domaine d'apprentissage qui associe plusieurs types acteurs (enseignants, apprenants, équipe techniques, etc.), nous nous concentrerons seulement sur la personnalisation du point de vue apprenants

comme le but de notre travail est d'améliorer leur expérience d'apprentissage.

— Adaptation de ressources d'apprentissage :

Les ressources pédagogiques du MOOC sont essentiellement représentées dans les vidéos, les examens (Quiz, examen finale) et quelques supports de texte (Hajri, 2018), afin d'adapter les cours sur les préférences et les capacités de l'apprenant, son appareil et son réseau. Le système compte d'abord sur deux points ; le premier est le contexte temps réel de l'apprenant et le deuxième est le niveau de l'apprenant et sa progression. D'une part, les règles générées dans le modèle du contexte changent le type de cours du vidéo au texte en cas de niveau de bruit élevé avec une localisation hors maison, ou en cas de faible vitesse de connexion internet qui peut causer des interruptions successives qui vont à leur tour produire des coupures de compréhension pour l'apprenants où il va sentir un isolement qui va finir par l'abandon (Elliott et al., 2005). Cette technique permet la recommandation des cours personnalisés.

— Adaptation selon le contexte temps réel :

Une fois l'apprenant connecté, le système calcule la vitesse de sa connexion, et via les capteurs mobiles il examine sa position et le niveau de bruit de son entourage, puis décide de proposer le cours en vidéo ou en support de texte. De l'autre part pour éviter l'avancement de l'apprenant dans le cours son avoir compris le maximum des unités abordées, nous comptons sur la détection de niveau d'engagement de chaque apprenant retiré depuis ses émotions faciales, cette procédure est mise en place via un classifieur qui détecte les mouvements de visage de l'apprenant, pour détecter son état émotionnelle via son visage, selon l'émotion détectée le système retire un taux d'engagement.

— Adaptation selon le taux de suivi :

L'exploitation des résultats des quiz est une étape très importante dans la personnalisation des cours en ligne, contrairement aux autres formations notre système propose une redirection selon les résultats des quiz. Les résultats possibles sont catégorisés selon les scores, donc chaque apprenant va être dirigé vers l'étape suivante selon son taux de succès au quiz.

— Adaptation selon la langue :

Le modèles propose une adaptation selon la langue de système d'exploitation de l'appareil mobile de l'apprenant, cette adaptation concerne non seulement l'interface mais même les sous-titres de la vidéo seront délivrés dans cette langue.

— Adaptation selon la catégorie :

Nous avons opté vers une approche de personnalisation par apprenant qui vise à collecter toutes les informations du profil utilisateur, cette approche se divise en deux catégories :

- Le stéréotype (Vincent et Grugeons, 2005) : c'est l'ensemble des caractéristiques rassemblées dans un groupe d'apprenants qui servent à déterminer les capacités du groupe égales ou proches des siennes. Cependant, le problème qui se pose souvent dans ce type est la difficulté qu'il affiche à considérer l'individualité de ces apprenants, ce qui ne répond pas au concept de personnalisation.
- Le profil de l'apprenant : Cette approche est propre à chaque abonné, elle retient plus précisément les informations extraites des apprenants, mais elle risque de tomber sur une analyse erronée pouvant conduire à des constatations invalides.

Nous avons pratiquement utilisé la deuxième catégorie dans toute notre proposition sauf dans le cas de niveau de l'apprenant et le taux de consultation ou nous avons opté vers le stéréotype, pour mieux réagir envers les changements rapides de l'apprenant.

3.6.4 Avantages de l'utilisation de Mobile Cloud Computing dans l'architecture proposée

L'utilisation du Cloud computing mobile permet d'éviter plusieurs problèmes causés par les systèmes M-learning traditionnelles, dans notre architecture, nous avons opté vers un service Cloud orienté infrastructure qui permet de garantir une intégration entre les systèmes M-learning interne et le Cloud. L'organisation du développement du MOOC (université, école, etc. . .) va être responsable de la gestion du contenu d'apprentissage, et

la disponibilité des cours, la gestion, le développement, la maintenance et le stockage sont garantis par les services de Cloud.

Avant d'émigrer vers les services Cloud, il faut assurer une continuité d'échange de flux et la sécurité de transfère et de données des apprenants, c'est pour cela nous avons utilisé deux types d'hébergements. Le premier type est consacré pour les modules du cours en relation directe avec les enseignants et les apprenants, utilisant un Cloud public comme elles nécessitent une disponibilité et un taux de calcul élevé. Les modules de gestion de contexte vont être hébergés dans un Cloud privé manipulé par un administrateur local du service pour des raisons de sécurité. L'utilisation du Cloud hybride est la solution idéale pour cette architecture, d'où la communication entre le Cloud privé et public est garantie via un réseau privé figure 3.9. L'impact de Cloud computing mobile pour les

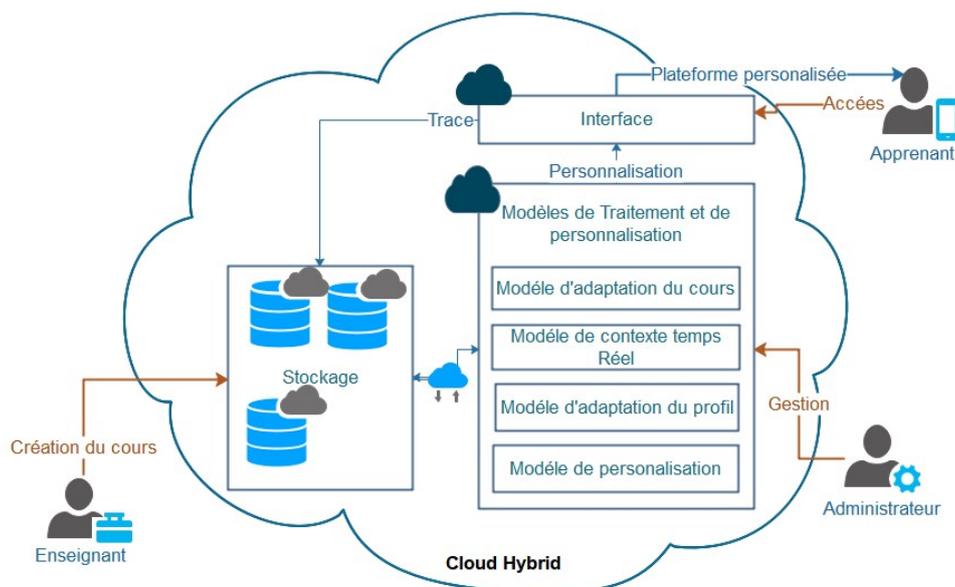


FIGURE 3.9: CD Mobile-Mooc basée sur le cloud computing

systèmes LMS en général est les M-learning en particulier est énorme et il a été discuté dans plusieurs recherches (Hassani et al., 2017; Hossain et al., 2013; Okai-Ugbaje et al., 2020), dans l'architecture proposée l'utilisation du Cloud a permis de dépasser les obstacles mentionnés par les lycéens interrogé dans l'enquête précédente. Nous avons résumé les avantages dans ce qui suit :

- La capacité de stockage et de calcul : les élèves des deux lycées ont mentionné leurs inquiétudes de l'espace est le coût de téléchargement de ressource éducative,

l'utilisation des serveurs Cloud permet de stocker les données et garantit la rapidité de calcul, sans se soucier des ressources de l'appareil.

- Disponibilité des data : Permet aux apprenants et même au enseignants de connecté depuis n'importe quelle autre appareil leurs données et leurs progrès sont tous sauvegardés sur le Cloud.
- L'implémentation est les mise à jours est beaucoup plus facile, vue la diversité des offres disponibles.
- L'utilisation des Cloud est plus ou moins coûteuse que la mise en place des serveurs locaux.
- La sécurité de partage des ressources entre les apprenants et entre les enseignants et les apprenants.

3.7 Discussion

Le système CD-Mobile MOOC offre un environnement en ligne où les apprenants et les enseignants peuvent se réunir pour produire une session pédagogique réussie, dans une méthode personnalisée basée sur l'approche de l'apprenant, en adaptant les services mobiles Cloud computing, les capteurs mobiles et l'environnement de l'apprenant.

En comparaison avec différents travaux, trois critères ont été utilisés : recommandation des différentes ressources efficaces et adéquates avec les préférences de l'apprenant, analyse du contexte de l'utilisateur, la priorisation de la compréhension et la motivation de l'apprenant.

Le modèle proposé dans notre contribution est conforme avec toutes les caractéristiques de comparaison. Une fois le contexte analysé par le modèle de contexte temps réel, le système CD-Mobile MOOC recommande des ressources adéquates pour les apprenants en fonction de leur langue, de leur niveau et de leurs préférences générales dans une série successives de traitements dans les différents autres modèles. La priorisation de la compréhension est observées dans le modèle d'adaptation du cours, l'apprenant ne peut en aucun cas dépasser le quiz s'il n'a pas dépassé le seuil défini dans cette étude par (>40%), de plus le suivi de son engagement permet d'alerter le système en cas de détection d'une

émotion d'apprentissage négative.

Pour la motivation elle est considérée comme l'objectif principal de notre proposition, tous les scénarios précédents sont intéressés par la motivation pour délivrer un bon niveau d'apprentissage qui garantit la meilleure expérience en exploitant toutes les ressources matérielles et logicielles disponibles. L'intégration de ce modèle dans une plateforme pédagogique confère des avantages significatifs.

En général, la plupart des travaux dans ce domaine exploitent les avantages du contexte des utilisateurs dans le système éducatif (Onah et Sinclair, 2015; Florian et al., 2015; Monsalve et al., 2020). Bien que certains aient écarté cette phase très importante en s'appuyant soit sur les traces de l'apprentissage comme le cas de (Montserrat et al., 2015), soit sur un formulaire rempli par l'apprentissage qui permet de détecter le style d'apprentissage de ce dernier (Drissi et Amirat, 2016). (Syed et al., 2017; Syed et Nair, 2018) priorise la motivation des apprenants, d'autres recherches donnent la priorité au contenu d'apprentissage en ligne (Marie et al., 2011; Tarus et al., 2018).

Basculer entre la vidéo et le support du texte est une caractéristique très importante dans notre modèle CD-Mobile MOOC, qui est particulière à la nature de certains endroits dans ce monde, où la connexion internet est très faible. Wu Turner (2020) discutent la relation entre la bande passante, l'interaction et les performances de l'apprenant, et leurs affectations sur les cours en ligne. (Dana et al., 2020) confirment qu'il existe des problèmes de débit pour les communautés rurales où de nombreux étudiants abandonnent les cours en ligne en raison d'un accès Internet médiocre, et la même chose s'est produite dans certains pays du tiers monde. Par conséquent, on a mis l'accent sur cette fonctionnalité qui peut motiver de plus l'apprenant.

De nombreuses architectures et modèles utilisent des techniques d'intelligence artificielle pour proposer un contenu personnalisé aux apprenants (Coppens, 2017), exploitent les avantages des appareils mobiles, (B.Zhang et al., 2016; Shorfuzzaman et al., 2018) et les services du Cloud (Li Gao, 2017; Elkhateeb et al., 2019) afin d'offrir la meilleure expérience pour les apprenants et les enseignants.

Le défi auquel nous sommes confrontés est d'exploiter pleinement ces technologies, pour

soutenir l'apprentissage en ligne, afin d'aider les apprenants à rester motivés et à dépasser les difficultés techniques, et les difficultés des cours.

Conclusion

Dans ce chapitre nous avons discuté la première contribution de cette thèse qui comprend deux parties : une statistique afin d'examiner le comportement et l'acceptation de la technologie par des lycéens, en identifiant les principaux défis et les limites que ces élèves rencontrent lors d'une séance d'éducation ouverte avec l'utilisation du mobile dans un processus d'apprentissage digitalisé. L'autre consiste à présenter une architecture basée sur le Cloud Computing mobile comme solution des défis extraites dans la première partie, nous avons exploité les services MCC, la technologie M-learning et les MOOC pour aboutir à modéliser cette architecture.

Pour résumer, cette modélisation permet aux apprenants de se lancer dans une expérience d'apprentissage à leur manière, en extrayant des informations de leur contexte et suivant leurs styles d'apprentissage individuels.

Dans le chapitre suivant nous allons discuter la deuxième contribution de cette thèse qui a aussi pour but de personnaliser l'apprentissage en ligne, en utilisant le deep learning, pour détecter les émotions faciales de l'apprenant.

Chapitre 4

Nouvelle approche d'apprentissage sensible aux émotions

Contents

4.1	Expressions faciales et apprentissage	124
4.2	Emotions et apprenant	125
4.3	Démarche de creation de classifieur	127
4.3.1	Description des émotions d'apprentissage	127
4.3.2	Description de la base de données	129
4.3.3	Prétraitement des images	133
4.3.4	Création de la nouvelle base de données basée unités d'action .	137
4.4	Modèle de classification des états	143
4.4.1	Deep Residual Learning (Resnet)	143
4.4.2	Architecture de classification des états	146
4.5	expérimentation : Résultat et discussion	148
4.5.1	Métriques d'évaluation	149
4.5.2	Synthèse et discussion	155
	Conclusion	156

L'une des Méthodes d'expressions les plus utilisées est comprise par les humaines peu importe leur origine, leur culture et leur langue et les émotions faciales, cette méthode de communication non verbale permet d'échanger des messages très puissants et naturels entre les personnes, tout le monde fait des expressions faciales qui donnent une vue générale sur leurs états émotionnels, la reconnaissance des expressions faciales (FER) est la méthode la plus puissante et la plus naturelle de communication émotionnelle non verbale. Ekman et Friesen, (1971) présentent un ensemble de six émotions acceptées comme émotions universelles vu en détails dans le chapitre 2, ces émotions dites des émotions de base ; la joie, la tristesse, la colère, la peur, le dégoût et la surprise. Notre visage est capable de créer plus que 7000 différentes expressions. Ces expressions sont traduites par

des émotions qui ajoutent beaucoup d'informations sur l'état de communication entre les personnes.

La reconnaissance des émotions faciales est un domaine qui intéresse beaucoup de chercheurs pour des différents intérêts.

Afin de décrire le mouvement des traits du visage qui crée l'expression faciale, nous utilisons souvent le système de codage de l'action faciale (FACS), qui décrit un ensemble d'unités d'action qui permettent de définir à leur tour les émotions et leur intensité. FACS est une méthode de description des mouvements du visage développée par les psychologues Paul Ekman et Wallace Friesen en 1978, elle est devenue le principal outil descriptif utilisé dans les études sur l'expression faciale, l'unité d'action est l'unité de base de la méthode du système de codage d'action faciale (FACS). Par conséquent, nous pouvons décrire chaque émotion en fonction des unités d'actions qui la représentent.

Dans ce chapitre nous allons présenter la deuxième contribution de cette thèse qui permet aussi d'offrir un environnement d'apprentissage personnalisé en se basant sur la reconnaissance des émotions faciales des apprenants.

Les expressions faciales des personnes changent en parallèle avec leurs comportements, nous pouvons utiliser cette caractéristique lors de la personnalisation de l'environnement d'apprentissage. Lorsqu'un apprenant est en train d'écouter un cours, ses expressions de visage nous donne beaucoup d'informations sur la compréhension générale de cet apprenant et son état émotionnel, nous pouvons aussi remarquer en étudiant son visage s'il a perdu son intérêt ou s'il est ennuyé, ce qui permet à ses enseignants d'ajuster le style d'apprentissage, afin de lui offrir un environnement d'apprentissage personnalisé selon ses besoins.

4.1 Expressions faciales et apprentissage

Comme nous l'avons mentionné dans le chapitre 2, l'utilisation des émotions des apprenants dans l'acte de la personnalisation d'apprentissage en ligne augmenterait considérablement leurs performances, ce qui sollicite plusieurs recherches, suivre cette voie comme celle de (Zatarain Cabada et al., 2020), afin d'adapter et d'ajuster les cours en suivant les

changements émotionnels des apprenants. Dans ce chapitre nous utiliserons les réseaux de neurones convolutifs (CNN) pour l'extraction et la reconnaissance des émotions des apprenants d'une manière instantanée, afin de donner aux enseignants le taux d'engagement total et celui de chaque apprenant. Pour que chaque apprenant reçoive un scénario pédagogique propre à lui, et afin d'appliquer le principe de personnalisation de l'apprentissage en ligne, nous allons nous baser sur le comportement des apprenants, parmi les différentes variables de contexte comportemental, nous comptons sur les expressions faciales.

L'utilisation de l'approche deep Learning cependant nécessite la plupart du temps une base de données dotée d'une taille importante, à ce propos on trouve beaucoup de bases de données pour les expressions faciales à des fins de recherches, mais la majorité de ces dernières ne présentent que les six émotions de base, certaines ajoutent un état neutre. Avantagement, quelques bases de données utilisent des codes d'unités d'action pour décrire les émotions comme Ck+, DISFA... etc, ces unités peuvent représenter toutes les expressions faciales possibles basées sur le système de codage d'action faciale (FACS) via les unités d'action (AU) (Ekman et Friesen, 1971). Actuellement ont trouvé 64 unités d'actions (AU) regroupées sur trois parties ; les principaux codes (44), les mouvements de la tête (13), et les mouvements des yeux (07), chacune d'elles est liée à un ou plusieurs mouvements des muscles faciaux.

4.2 Emotions et apprenant

Les apprenants développent souvent plusieurs émotions dans leurs environnements éducatifs, qui se diffèrent selon la nature du cours suivi, l'émotion sur l'environnement d'apprentissage est divisée en deux parties : des émotions positives et d'autres négatives, ces émotions sont capturées de plusieurs façons telles que les expressions faciales, les mouvements corporels et la parole, mais les plus utilisées sont la reconnaissance de l'émotion selon l'expression du visage (Tarnowski et al., 2017). Les apprenants créent des expressions faciales lorsqu'ils sont intéressés par le sujet, un attachement lorsqu'ils sentent son importance, ce qui permet de créer certains émotions lors du parcours éducatif. Ces émotions peuvent donner une vue générale sur l'engagement des apprenants, pouvoir décrypter le

code des différentes expressions faciales en temps réel permet aux enseignants de les intégrer de plus en plus, ce qui va influencer considérablement le processus d'apprentissage. On trouve que les émotions d'apprentissage utilisées dans la plupart des recherches sont pratiquement les émotions basiques, logiquement pas toutes les émotions basiques sont considérées comme des émotions d'apprentissage, plus précisément ça reste fortement lié à la nature du cours présenté, par conséquent les émotions de bases ne peuvent pas déterminer tout le temps les détails de processus d'apprentissage. Dans notre contribution nous allons nous concentrer sur un type d'émotion dit : émotions de réalisation ou « *Achievement emotions* » présenté par Pekrun (2006), comme un type d'émotions directement liées soit aux activités de la réalisation, soit aux résultats de la réalisation. Les activités de réalisation comme les mathématiques et le physique laissent chez les apprenants des émotions comme la satisfaction, la frustration, l'ennui et le désengagement, les résultats de réalisation laissent chez eux des émotions comme la joie, l'espoir, la fierté, l'anxiété, la nervosité, le désespoir, la honte et la colère, liés au succès ou à l'échec. Nous avons choisi de ce type large d'émotions de réalisation six émotions (Satisfaction, confusion, frustration, nervosité, ennui, surprise), le tableau 4.1 présente la définition des émotions d'apprentissage choisies.

NB. Les définitions du Table 4.1 sont extraites du dictionnaire français Larousse.

Emotions	Définitions
La satisfaction	Action de satisfaire un besoin, un désir ou une demande
La confusion	Action de confondre, de prendre quelque chose ou quelqu'un, pour quelque chose, quelqu'un d'autre, erreur ou méprise
La frustration	Priver quelqu'un d'un bien, d'un avantage dont il croyait pouvoir disposer, qu'il pensait posséder
La nervosité	État permanent ou momentané d'irritabilité ou d'inquiétude
L'ennui	Laisser quelqu'un, ne provoquer chez lui aucun intérêt, lui causer un sentiment de fatigue
La surprise	État de quelqu'un qui est frappé par quelque chose d'inattendu

TABLE 4.1: La définition des émotions d'apprentissage

4.3 Démarche de creation de classifieur

Les émotions des apprenants affectent la réussite de l'apprentissage, elles jouent un rôle important pour comprendre l'orientation éducative des élèves, elles reflètent aussi la plupart du temps le niveau d'attention, motivation et acceptation des apprenants (Lasri, et al., 2019). Les émotions négatives telles que la frustration, l'ennui et la nervosité peuvent affecter négativement ce processus d'apprentissage, donc détecter les transitions émotionnelles et mettre en place des tentatives de régulation des émotions négatives une fois qu'elles se sont produites, peuvent être particulièrement importantes dans un environnement d'apprentissage en ligne.

Le but de ce chapitre, est de présenter un modèle en profondeur pour améliorer la personnalisation dans les plateformes éducatives, via la reconnaissance des émotions des apprenants, cette méthode de reconnaissance des émotions faciales est basée sur la détection des unités d'action. Notre travail va se baser sur les images, le premier défi rencontré est la rareté des données, cependant, il y a très peu de données existantes à apprendre et leur acquisition est difficile et chère. C'est pour cela que la première partie dans notre contribution était de créer une base de données depuis une base déjà existante mais qui sert beaucoup plus nos besoins, ensuite nous avons créé le modèle de reconnaissance.

4.3.1 Description des émotions d'apprentissage

Dans cette recherche, nous sélectionnons l'émotion basée sur l'approche des émotions de réalisation. Ainsi, on reconnaît six émotions d'apprentissage, constituées de Satisfaction, confusion, frustration, nervosité, ennui, surprise. «Satisfaction, confusion, surprise» présente l'émotion positive qui est décrite comme un état où l'apprenant est prêt à accueillir, transférer et digérer des informations. Néanmoins, « la frustration, la nervosité, l'ennui» présentent une émotion négative qui est décrite comme un état de refus ou peur de l'acquisition de connaissances. Nous avons défini chaque émotion dans son contexte éducatif, et nous l'avons décrite avec ses unités d'action appropriées.

- Satisfaction : La satisfaction des apprenants peut être un terme très large qui occupe plusieurs aspects, mais nous allons nous concentrer sur la satisfaction de réalisation

qui doit être non seulement détectée via les expressions faciales mais aussi avec les traces de l'apprenant sur la plateforme éducative, donc la satisfaction de l'apprenant représente son bien être dans l'environnement qui le suit ce qui conduit vers une perception positive. Ce sentiment est souvent traduit par les unités d'action AU6, AU12 (Ekman et Friesen, 1976).

- Confusion : Un apprenant engagé dans un cours se sent confus lorsqu'il confronte une contradiction, une panne du système ou une erreur dont il n'est pas capable d'expliquer (Coppens, 2017). Ce sentiment est souvent traduit par les unités d'action AU1, AU4, AU7 et AU12 (McDaniel et al., 2007)
- Frustration : L'état de frustration survient chez l'apprenant lorsqu'il se trouve dans une situation sans solution, dont il n'a pas de plan, donc il reste figé devant un dilemme quelconque, selon son engagement ce dernier peut fournir plus d'effort et résoudre le problème sinon la frustration pourrait se transformer en ennui (D'Mello et Graesser, 2012). Ce sentiment est souvent traduit par les unités d'action AU4 et AU14 (Grafsgaard et al., 2013).
- Nervosité : Cette émotion est souvent traitée dans le contexte d'apprentissage comme la peur plus précisément c'est le fait d'anticiper et subir une menace physique ou psychique. Selon (Zhang et al., 2014) soit Remontée de la partie interne des sourcils (AU1), Remontée de la partie externe des sourcils (AU2) et Abaissement et rapprochement des sourcils (AU4) ou Remontée de la partie interne des sourcils (AU1), Remontée de la partie externe des sourcils (AU2) et Ouverture entre la paupière supérieure et les sourcils (AU5) doit apparaître lorsque on traite la nervosité. Nous pouvons aussi ajouter Étirement externe des lèvres (AU20) et Ouverture de la bouche et séparation légère des lèvres (AU25) à l'ensemble précédent (Gagnon et al., 2015).
- L'ennui : Aucune caractéristique faciale corrélative n'a été trouvée. (McDaniel et al., 2007). Depuis l'ensemble de différentes unités d'action, nous extrayons celles qui montrent un acte hors écran, comme AU43, et les mouvements de yeux AU61, AU62, AU63 et AU64, qui présentent respectivement ; les yeux fermés, les yeux vers

la gauche, les yeux vers la droite, les yeux levés, les yeux baissés.

- Surprise : La surprise est l'une des émotions qui prouve le plus l'engagement des apprenants, une pédagogie d'apprentissage discutée dans un colloque organisé par l'association québécoise de pédagogie collégiale (AQPC), par Carol Fortin et Esteban Gonzalez, qui s'appelle la pédagogie de l'inattendu, affirme qu'afin d'attirer l'attention des élèves il faut utiliser des activités non habituelles, cet acte prouvé son efficacité. Les tâches monotones entraînent l'ennui ce qui amène vers le désengagement, par contre la surprise donne aux apprenants l'envie de découvrir. Dans les systèmes d'apprentissage actuels à l'école ou hors l'école l'émotion surpris est devenue primordiale. Ce sentiment est souvent traduit par les unités d'action AU1, AU2, AU5 et AU26. Remontée de la partie interne des sourcils (AU1), Remontée de la partie externe des sourcils (AU2), Ouverture entre la paupière supérieure et les sourcils (AU5), Ouverture de la mâchoire (AU26) (Ekman et Friesen, 1976).

Il faut toujours interpréter l'expression faciale, comme des humaines on peut le faire naturellement, nous observons d'abord les parties du visage qui changent d'apparence, puis nous déduisons que cela signifie une certaine émotion, c'est plus simple de dire qu'une personne à l'air surpris que qu'il a activé AU1, AU2, AU5 et AU26. Afin d'équiper la machine de cette particularité il faut envisager toutes possibilités de présence des unités d'action pour chaque émotions en étude, Valstar (2008) a étudié les possibilités de mapper les AU reconnues aux six émotions de base, de plus (Craig et al.,2008), ont de même présenté les possibles AU pour la confusion et la frustration pour le cas de cette dernière il ont déduit sa forte relation avec les AU1 et AU2 quand ils se produisent ensemble. Le tableau 4.2, montre les possibilités de présence de chaque unité d'action pour les émotions d'apprentissage choisis.

4.3.2 Description de la base de données

Le principe de notre travail est d'entraîner un modèle en profondeur de reconnaissance des unités d'action, la plus part des bases de données se concentre sur les émotions universelles, on trouve ces dernières sur plusieurs formes (vidéo, image, 2D, 3D, ou même

des dessins animés), dans cette contribution nous nous sommes basés sur un autre type d'émotion dites émotions d'apprentissage.

Emotion	Unité d'Action	La description de l'unité d'action
Satisfaction	6+12	AU6 : Remontée des joues
	12	AU12 : Étirement du coin des lèvres
Confusion	1+4+7+12	AU1 : Remontée de la partie interne des sourcils
	4	AU4 : Abaissement et rapprochement des sourcils
	7	AU7 : Tension de la paupière
	4+7	AU12 : Étirement du coin des lèvres
Frustration	4+14	AU4 : Abaissement et rapprochement des sourcils
	1+2	AU14 : Plissement externe des lèvres
		AU1 : Remontée de la partie interne des sourcils AU2 : Remontée de la partie externe des sourcils
Nervosité	1+2+4+5+20+25	AU1 : Remontée de la partie interne des sourcils
	1+2+4	AU2 : Remontée de la partie externe des sourcils
	1+2+5	AU4 : Abaissement et rapprochement des sourcils AU5 : Ouverture entre la paupière supérieure et les sourcils AU20 : Étirement externe des lèvres AU25 : Ouverture de la bouche et séparation légère des lèvres
Ennuie	43	AU43 : Les yeux fermés
	61	AU61 : Les yeux vers la gauche
	63	AU63 : Les yeux levés
	64	AU64 : Les yeux baissés
Surprise	1+2+5+26	AU1 : Remontée de la partie interne des sourcils
	1+2+26	AU2 : Remontée de la partie externe des sourcils
	5+26	AU5 : Ouverture entre la paupière supérieure et les sourcils AU26 : Ouverture de la mâchoire

TABLE 4.2: Les possibilités de présence de chaque unité d'action pour les émotions d'apprentissage choisis

L'utilisation du deep Learning nécessite un taux raisonnable de données qui permet d'entraîner le modèle sans se soucier de problème comme overfitting (Bonnes performance sur les données d'entraînement, mauvaise prédiction) ou underfitting (mauvaise performance sur les données d'entraînement, mauvaise prédiction), c'est pour cela la plupart des recherches utilise une ou plusieurs bases de données destinées au recherches scientifique. Dans le cas des émotions d'apprentissage ces bases de données ne servent pas complètement les besoin de la recherche, ce qui permet de définir plusieurs solution. Certains chercheur Mohamad Nezami et al., (2020) entraînent leur modèle avec la base de données FER-2013 (qui contient des images des émotions universelles) ensuite ils ont créé leur propre base qui contient des émotions d'apprentissage (satisfait, confus, ennuyé) et re-entraîner le modèle avec la nouvelle base afin de détecter l'engagement des apprenants. D'autres comme Candra et al., (2018), créent leur propre base de données qui contient deux émotions : intérêt et ennui, d'abord ils ont enregistré les émotions des élèves au format vidéo, ces vidéos sont ensuite converties en un seul frame de visage toutes les 25 secondes de manière itérative et automatique. Il existe aussi plusieurs recherches qui ont utilisé les émotions basiques directement comme dans le cas de (Bahreini et al., 2016 ; El Hammoumi et al., 2018 ; Li et al., 2021).

La méthode proposée :

Dans cette étude, nous développons un système de reconnaissance des unités d'actions faciales responsables des émotions d'apprentissage basé sur les réseaux de neurones convolutifs, dans le but de l'intégrer dans un système d'apprentissage en ligne. Notre contribution se base sur la détection et la reconnaissance des unités d'action, le principe de ce travail se base sur le faite que chaque émotion faciale cause un mouvement musculaire, ce mouvement est traduit par une ou plusieurs unités d'action (Figure 4.1), ces unités d'action sont codées selon un système de codage des actions faciales (Figure 4.2), donc la détection et le regroupement des unités d'action d'une émotion quelconque va permettre de la reconnaître. Donc nous avons utilisé la base de données Extended Cohn-Kanade (Lucey et al., 2010), car elle contient le codage des unités d'action qui va nous

servir pour la création de la nouvelle base de données, malgré qu'elle est considérée comme une petite base de données. Bien connue que, souvent les architectures deep Learning ne sont pertinents que lorsque nous disposons d'une énorme quantité de données, mais cela n'est pas toujours le cas, certes, l'apprentissage en profondeur nécessite la capacité d'apprendre automatiquement des caractéristiques à partir des données d'entraînement, ce qui n'est généralement possible que lorsque de nombreuses données d'apprentissage sont disponibles. Cependant, les réseaux de neurones convolutifs (CNN) sont par conception l'un des meilleurs modèles disponibles pour la plupart des problèmes de classification d'images, même avec très peu de données à apprendre lors de bons choix des filtres (Zhang et al., 2016). C'est pour cela que nous avons opté pour l'utilisation d'un modèle basé sur les réseaux de neurones convolutifs.

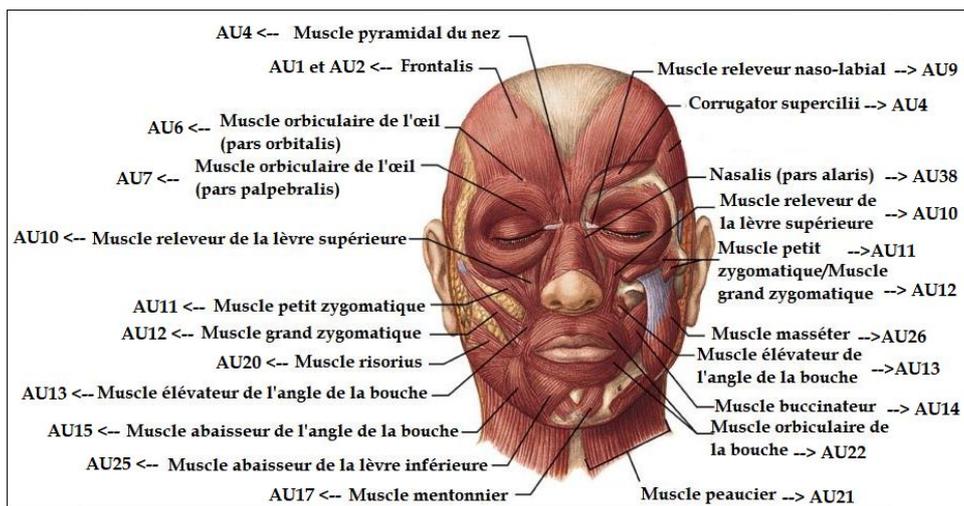


FIGURE 4.1: Les muscles avec la liste des unités d'actions associées

Extended CohnKanade CK+

(Lucey et al., 2010), est une base de données qui comprend un total de 593 séquences d'images qui varient en durée (de 10 à 60 frames) de 123 personnes, âgés de 8 à 50 ans, 69% sont des femmes, 81% d'Euro-Américains, 13% d'Afro-américains et 6% d'autres groupes. Cette base de données contient une variété d'émotions et d'unités d'action (07 émotions et 30 AU), elle se compose de sept émotions (les six émotions principales plus le mépris) avec le visage neutre qui représente la première image de la séquence dont la fin

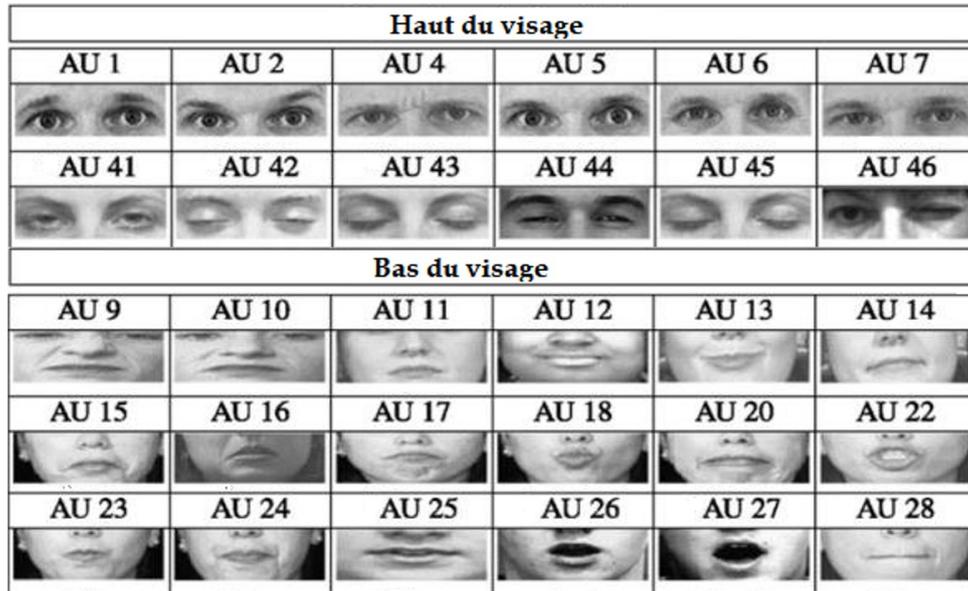


FIGURE 4.2: Facial Action Coding System (FACS) (Yin et al., 2017)

est désignée par l'image qui représente le pic d'émotion. Ces images ont été enregistrées, certaines en niveau de gris et d'autres en couleur, en 640×490 ou 640×480 pixels. Parmi les 593 images 327 représentent les pics d'émotions, elles se composent de 45 images pour la colère, 18 images pour le mépris, le dégoût se figure sur 59 images, 25 images présentent la peur, 69 images pour la joie, 28 images présentent des expressions tristes et enfin la surprise est vue sur 83 images, La figure 4.3 présente un échantillon de six expressions de la base de données CK+ et le visage neutre.



FIGURE 4.3: Les émotions universelles depuis la base de données CK+

4.3.3 Prétraitement des images

La classification des unités d'action est un problème de classification multi-label, bien que le CNN est l'approche de classification la plus populaire pour ce type de problèmes, son utilisation pour des images hautement structurées comme le visage humain peuvent ne pas capturer tous les détails du changement. Pour cette raison nous avons proposé

une règle pour la décomposition de l'image d'un visage en groupes chacun d'entre eux contient un ou plusieurs états selon les unités d'action qui peuvent apparaître a la région qu'il couvre.

AU	Description de l'unité d'action	Parties du visage
1	Remontée de la partie interne des sourcils	Sourcil et paupière
2	Remontée de la partie externe des sourcils	Sourcil et paupière
4	Abaissement et rapprochement des sourcils	Sourcil et entre sourcil
5	Ouverture entre la paupière supérieure et les sourcils	Paupière et yeux
6	Remontée des joues	Les yeux et les joues
7	Tension de la paupière	Paupière et yeux
12	Étirement du coin des lèvres	Du bout du nez à la bouche
14	Plissement externe des lèvres	Du bout du nez à la bouche
20	Étirement externe des lèvres	Le menton et la bouche
25	Ouverture de la bouche et séparation légère des lèvres	Le menton et la bouche
26	Ouverture de la mâchoire	Le menton et la bouche
43	Yeux fermés	Paupière et yeux
61	Yeux tournés vers la gauche	Paupière et yeux
63	Yeux levés	Paupière et yeux
64	Yeux baissés	Paupière et yeux

TABLE 4.3: Les unités d'action et leur région

D'abord nous avons examiné les unités d'action en question et leurs positions et région d'apparition dans le visage (Table 4.3), en suite nous avons regroupé les régions similaire ensemble, les unités d'action qui se trouve dans des régions proches et qui peuvent apparaître en même temps sont regroupées sous le nom d'état le tableau (Table 4.4) résume ce travail.

Description des groupes

Comme nous l'avons mentionné dans la table 4.4, nous avons retiré sept (07) groupes du visage selon les unités d'action traitées. Ces sept groupes forment ensemble 13 états :

- Groupe 01 : Responsable de la région qui couvre les sourcils et les paupières comme dans les unités d'action 1 et 2, qui présentent respectivement la remontée de la

Les groupes	Les états	Les unités d'action	Région sur le visage
G01	E1	AU1	Sourcil et paupière
G01	E2	AU1 + AU2	Sourcil et paupière
G02	E3	AU5	Paupière et yeux
G03	E4	AU1+AU2+ AU5	Sourcil, paupière et yeux
G02	E5	AU7	Paupière et yeux
G03	E6	AU1 + AU7	Sourcil, paupière et yeux
G04	E7	AU4	Partie de sourcil et entre sourcil
G05	E8	AU6	Les yeux et les joues
G06	E9	AU12	Du bout du nez à la bouche
G06	E10	AU14	Du bout du nez à la bouche
G07	E11	AU26	Le menton et la bouche
G07	E12	AU20 + AU25	Le menton et la bouche
G02	E13	AU43 ou AU61 ou AU63 ou AU 64	Paupière et yeux

TABLE 4.4: Les états et les groupes selon les régions des unités d'action

partie interne des sourcils, et la remontée de la partie externe des sourcils.

- Groupe 02 : Responsable de la région des paupières et des yeux, dans notre contribution nous avons groupé les unités action AU5, AU7, AU43, AU61, AU63 et AU64 sous ce groupe, la description des AU mentionnée est fortement liée à la partie soit des paupières (AU5, AU7) soit des yeux (AU43, AU61, AU63, AU64) le mouvement des paupières affectées le mouvement des yeux et vice versa.
- Groupe 03 : Ce groupe est responsable des deux états E4 et E6 qui eux même décrivent respectivement la situation où les unités d'action 1, 2 et 5 sont présentées ensemble ou la situation dont AU1 et AU7 sont déclenchés en même temps, vue la description des deux groupes précédents (1 et 2) le G03 couvre les zones (yeux, paupières et sourcils).
- Groupe 04 : Couvre une seule unité d'action selon notre contribution. Les deux muscles responsables du mouvement de AU4 sont (Muscle pyramidal du nez et Corrugator supercilii) figure 4.4, dont vient les régions sourcils et entre sourcils.
- Groupe 05 : Responsable de la région qui couvre les yeux et les joues comme dans l'unité d'action 6, qui est définie par la remontée des joues, ce mouvement mène vers un mouvement du muscle Orbicularis oculi (Muscle orbiculaire de l'œil) présent au

tour des yeux (Figure 4.1).



FIGURE 4.4: Les muscles responsables d'AU4

- Groupe 06 : Regroupe AU12 et AU14 respectivement défini par l'étirement du coin des lèvres et le plissement externe des lèvres, la région dont ce groupe est responsable est la zone du bout de nez à la bouche, le mouvement des lèvres lors de l'AU12 cause le mouvement du muscle grand zygomatique , et pour le cas de l'AU14 le muscle en question est le muscle buccinateur, la figure 4.5 montre ces deux muscles et leurs positions.



FIGURE 4.5: Les muscles responsables d'AU12 et AU14

- Groupe 07 : Le dernier groupe est responsable des unités d'action 20, 25 et 26, leurs définitions respectives sont ; Étirement externe des lèvres, ouverture de la bouche et séparation légère des lèvres et l'ouverture de la mâchoire qui montre que le mouvement de ces trois, passe dans la zone de la bouche mais nous avons ajouté le menton aussi car il est directement affectée par l'ouverture et la fermeture de la bouche.

4.3.4 Création de la nouvelle base de données basée unités d'action

Chaque image de la base de données CK+ a été traitée comme un groupe d'images d'état, selon la définition des unités d'action spécifiques fournies avec la base de données, pour chaque une des (593) séquences d'images, les créateurs de la base de données ont défini un seul fichier FACS, qui regroupe les unités d'actions présentes dans l'image pic. Notre modèle est responsable de reconnaître chaque état présenté dans les émotions d'apprentissage choisi (Table 4.5), avec aussi l'état neutre (un état ou il n'y a pas une présence d'unités d'action). La création de la nouvelle base de données d'état basée sur CK+, passe par plusieurs étapes :

Emotion	Satisfaction	Confusion	Frustration	Nervosité	Ennui	Surprise
Les états	E8+E9 E9	E6+E7+E9 E7 E5 E7+E5	E7+E10 E2	E4+E7+E12 E2+E4 E4	E13	E4+E11 E2+E11 E3+E11

TABLE 4.5: Les possibilités de présence de chaque état pour les émotions d'apprentissage choisi

— Échelle de gris

L'image en niveau de gris n'a qu'un seul canal et l'intensité des pixels varie entre 0 (noir) et 255 (blanc), donc moins d'informations sont fournies pour chaque pixel, les résultats expérimentaux ont montré que la classification avec des images en niveau de gris entraînait une classification plus précise qu'avec des images RGB à travers les différents types de classificateurs (Bui, 2016). Les images de CK+ sont un mélange entre des images en couleurs et autres en niveau de gris, par conséquent, pour unifier le type d'entrée, les images ont toutes été converties en gris.

— Détection de visage et recadrage

Nous avons utilisé dlib toolkit (King, 2009) pour obtenir 68 points de repère (landmark) Figure 4.6, ces points caractérisent les zones où les AU peuvent apparaître, toutes en fournissant des informations importantes sur le visage. Le fonctionnement ici était de recadrer l'image, à l'aide de `dlib.get_frontal_face_detector`, qui ren-

verra les informations sur les visages, en particulier elle permet de trouver le visage humain.

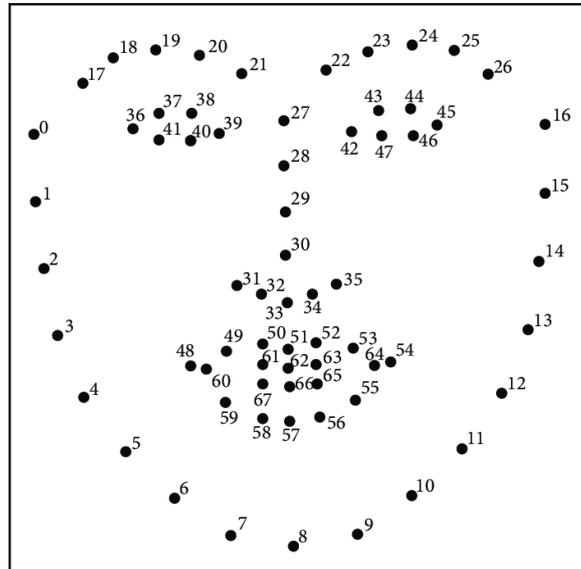


FIGURE 4.6: Face Landmark 68 points

— Création des images d'état (Extraction de régions)

Après avoir défini les régions nous avons partitionné le visage à l'aide des landmark, ces derniers sont retirés grâce à la méthode *dlib.shape_predictor*. Nous extrayons la position de 68 repères (les coins de la bouche, le nez, les sourcils, les yeux, etc.), ensuite l'image subit une division en sept groupes (Figure 4.7), selon les règles de décomposition et les coordonnées des landmarks. Il est important de mentionnée dans cette phase que l'utilisation des point de repère seulement n'est pas suffisant pour créer les rectangles de division de la façon la plus correcte, les lieux réels où se déroulent les activités peuvent manquer, d'où les erreurs du système peuvent augmenter. Au lieu d'utiliser les cordonnées exactes des landmark, nous adoptons selon la définition des groupes déjà mentionnés à créer des régions d'intérêt liées au unités d'action, puis la division est faite selon ces ROI, les nouvelles cordonnées sont présentées dans le tableau 4.6. Enfin selon la définition des unités d'action fournis pour cette image en particulière nous créons l'ensemble d'image étiqueté par les états qu'elles représentent selon son fichier FACS, les nouvelles images enregistrées sont seulement celles qui ont des états actifs.

Comme résultat, pour chaque image de la base de données CK+, nous allons avoir dix (10) images selon les sept groupes en bias de fichier FACS de cette image, nous pouvons définir les étiquettes de l'ensemble de ces dix images, seuls ceux qui ont un état actif sont sauvegardés les autres seront ignorés.

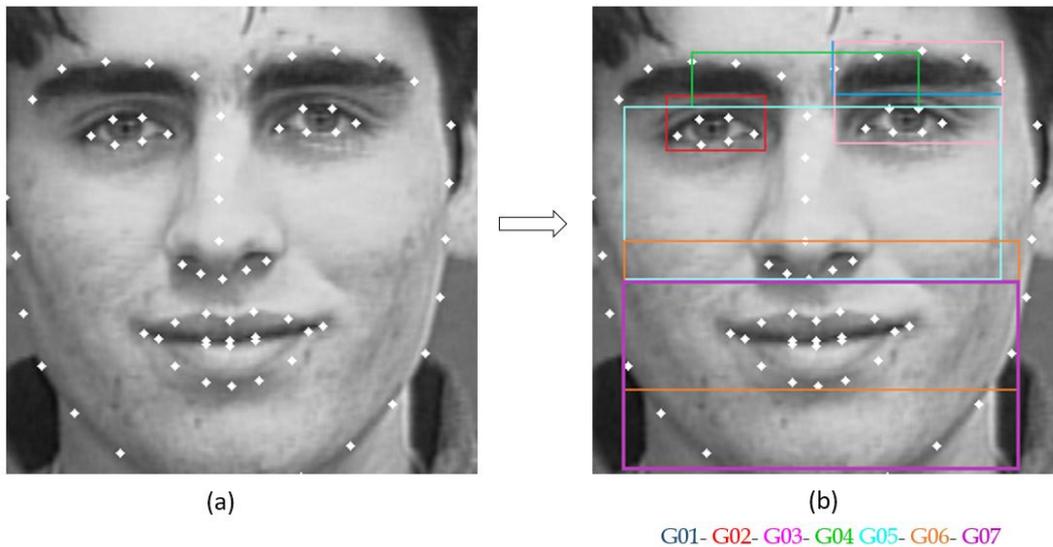


FIGURE 4.7: Landmark et rectangles de division (a) : Les points de repère. (b) : Les rectangles de chaque groupe

- Désignation des coordonnées des rectangles de division

Afin de bien entourer la région des unités d'action, nous avons utilisées deux critères ; les points de repère et la moitié de la distance entre deux points, dont nous avons utilisé la distance euclidienne (4.1), le tableau 4.6 présente les coordonnées utilisées dans la division des images d'état.

$$\sqrt{\sum_{i=1}^n (x_i + y_i)} \quad (4.1)$$

La fonction $\text{dis}()$, est la fonction qui permet de calculer la moitié de distance entre les deux point en argument par exemple $\text{dis}(x_{17}, x_0)$ est définis par :

$\text{math.sqrt}((x_{17} - x_0)^2)/2$. L'utilisation de : α est due au ajustement des bordures.

- Ajustement des bordures

Dans quelques cas la région des yeux n'est pas complètement claire (Figure 4.8), et cela est remarqué surtout lors de l'émotion surprise où les yeux sont grand ouvert,

G01 : (G)	A : $(x_{17} - \text{dis}(x_{17}, x_0) ; y_{19} - \alpha)$ B : $(x_{21} + \text{dis}(x_{27}, x_{21}) ; y_{19} - \alpha)$ C : $(x_{21} + \text{dis}(x_{27}, x_{21}) ; y_{37})$ D : $(x_{17} - \text{dis}(x_{17}, x_0) ; y_{37})$	G03 : (G)	A : $(x_{17} - \text{dis}(x_{17}, x_0) ; y_{19} - \alpha)$ B : $(x_{21} + \text{dis}(x_{27}, x_{21}) ; y_{19} - \alpha)$ C : $(x_{21} + \text{dis}(x_{27}, x_{21}) ; y_{41} + \alpha)$ D : $(x_{17} - \text{dis}(x_{17}, x_0) + \alpha ; y_{41} + \alpha)$
G01 : (D)	A : $(x_{22} - \text{dis}(x_{22}, x_{27}) ; y_{24} - \alpha)$ B : $(x_{26} + \text{dis}(x_{26}, x_{16}) ; y_{24} - \alpha)$ C : $(x_{26} + \text{dis}(x_{26}, x_{16}) ; y_{43})$ D : $(x_{22} - \text{dis}(x_{22}, x_{27}) ; y_{43})$	G03 : (D)	A : $(x_{22} - \text{dis}(x_{22}, x_{27}) ; y_{24} - \alpha)$ B : $(x_{26} + \text{dis}(x_{26}, x_{16}) ; y_{24} - \alpha)$ C : $(x_{26} + \text{dis}(x_{26}, x_{16}) ; y_{46} + \alpha)$ D : $(x_{22} - \text{dis}(x_{22}, x_{27}) ; y_{46} + \alpha)$
G02 : (G)	A : $(x_{17} + \text{dis}(x_{17}, x_{36}) ; y_{37} - \text{dis}(y_{37}, y_{19}))$ B : $(x_{21} - \alpha ; y_{37} - \text{dis}(y_{37}, y_{19}))$ C : $(x_{21} - \alpha ; y_{41} + \alpha)$ D : $(x_{17} + \text{dis}(x_{17}, x_{36}) ; y_{41} + \alpha)$	G04 :	A : $(x_{19} ; y_{19})$ B : $(x_{24} ; y_{19})$ C : $(x_{24} ; y_{37})$ D : $(x_{19} ; y_{37})$
G02 : (D)	A : $(x_{22} + \text{dis}(x_{22}, x_{42}) ; y_{43} - \text{dis}(y_{43}, y_{23}))$ B : $(x_{26} - \alpha ; y_{43} - \text{dis}(y_{43}, y_{23}))$ C : $(x_{26} - \alpha ; y_{46} + \alpha)$ D : $(x_{22} + \text{dis}(x_{22}, x_{42}) ; y_{46} + \alpha)$	G05 :	A : $(x_{17} ; y_{37})$ B : $(x_{26} ; y_{37})$ C : $(x_{26} ; y_{30})$ D : $(x_{17} ; y_{30})$
G06 :	A : $(x_4 ; y_{30})$ B : $(x_{12} ; y_{30})$ C : $(x_{12} ; y_{57})$ D : $(x_4 ; y_{57})$	G07 :	A : $(x_4 ; y_{33})$ B : $(x_{12} ; y_{33})$ C : $(x_{12} ; y_8)$ D : $(x_4 ; y_8)$

TABLE 4.6: Les coordonnées de rectangles de division des images d'état

l'ajustement des bordures est garantie par une petite valeur $\pm\alpha$ varie entre $[3,5]$ cet intervalle est choisi après avoir testé sur plusieurs entier.

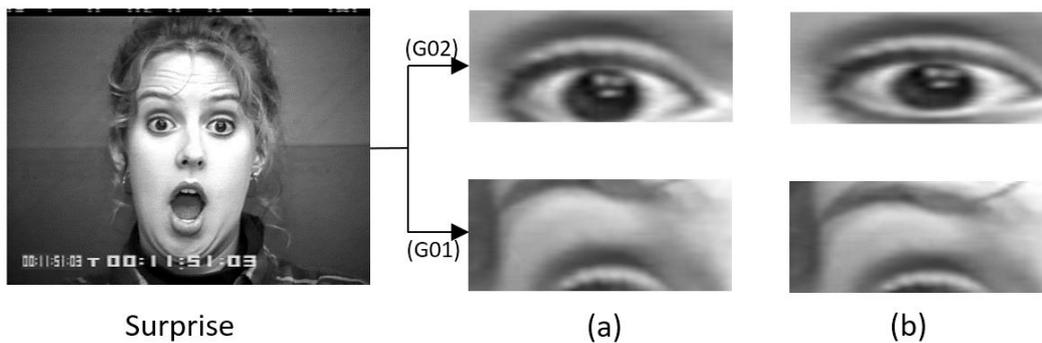


FIGURE 4.8: Recadrage de l'image sans la valeur α (a). Recadrage de l'image avec la valeur α (b)

— Redimensionnement

Cette étape consiste à redimensionner les images d'état avec les bordures ajustées, cela a été fait pour s'assurer que les images étaient toutes de même taille car elle représente dix régions différentes ce qui mène vers des tailles différentes en raison

de la nature et de la taille du visage et la région d'intérêt. Par conséquent, le redimensionnement a dû être fait pour atteindre une entrée d'image standard. Afin de décider la taille la plus adéquate nous avons utilisé 50 images neutres qui nous a donné 500 images d'état, (état 0), les tailles de sortie ont été enregistrées ensuite nous avons retiré la moyenne de la longueur et la largeur depuis ces enregistrements. Nous avons utilisé les images neutres où tous les états sont E0, afin d'augmenter le taux de cohérence entre les images.

La moyenne était de 117×50 ; Ainsi, toutes les images de tous les ensembles de données devraient être de cette taille pour assurer la cohérence et garantir le même nombre de caractéristiques pour toutes les images, l'utilisation de cette technique n'a pas affectée la résolution.

— Augmentation des données

L'augmentation des données est une stratégie utilisée pour d'une part augmenter la quantité de données, et de l'autre part, rendre le modèle plus robuste et empêcher l'overfitting en créant des versions modifiées d'images, en utilisant plusieurs techniques. Dans notre cas nous avons utilisé sept (07) techniques différentes d'augmentation, il s'agit de (décalage en largeur, décalage en hauteur, transformation de cisaillement, zoom, luminosité et retournement horizontal et vertical), le choix de ces techniques s'explique par les types d'images d'état.

- Décalage en largeur : l'image est décalée horizontalement avec un taux varie de $]0,1]$, nous avons utilisé ce type de décalage pour toutes les régions du groupes 4, 5, 6 et 7 avec un décalage qui ne dépasse pas le 0.2.
- Décalage en hauteur : l'image est décalée verticalement avec un taux varie de $]0,1]$, nous avons utilisé ce type de décalage pour les régions du groupes 4 et 05 avec un décalage qui ne dépasse pas le 0.2.
- Transformation de cisaillement : incline la forme de l'image pour fixer un axe et étirer l'image à un certain angle dans ce cas l'angle est de 20° , nous avons utilisé cette technique pour tous les groupes.
- Zoom : un zoom pour agrandir l'image est appliquer pour une plage de $[0.8,1.0]$,

cette technique est appliquée pour tous les groupes.

- Luminosité : la plage de luminosité appliquée est de $[0.1, 1.0]$, ce qui permet de générer des images aléatoires avec un décalage de lumière différent, nous l'avons appliqué sur tous les groupes.
- Retournement horizontal et vertical : cette stratégie permet de générer des images inversées horizontalement d'une manière aléatoire, nous l'avons utilisée pour tous les groupes. Nous pouvons également appliquer un retournement vertical de la même façon.

Certaines techniques comme le décalage en largeur, décalage en hauteur, transformation de cisaillement, créent des points sans valeur, quoi que nous ayons utilisé des petites plages pour ces derniers afin de ne pas perdre des informations importantes, mais les points vides restent une partie de la procédure. Il existe plusieurs types de remplissage nous avons choisi d'utiliser, 'nearest' ou le plus proche qui est considéré comme la valeur par défaut où les points vides sont remplis par la valeur de pixel le plus proche, la figure 4.9 montre un échantillon d'images générées. Il faut aussi noter que nous avons combiné certaines stratégies, pour avoir encore plus de possibilités de génération comme (luminosité avec zoom, le retournement horizontal avec le cisaillement). Les résultats ont été beaucoup améliorés après l'utilisation de l'augmentation de données qui a permis au modèle de découvrir plus de caractéristiques et de renforcer la prédiction.

A la fin de cette phase, nous avons eu une nouvelle base de données fondée sur les images et les informations de la base CK+, cette BD est étiquetée avec treize (13) états, qui définissent la présence de quinze (15) unités d'action et un autre état neutre, (Figure 4.10) résume le processus détaillé afin d'obtenir la base de données d'état. La création de ces données permet de reconnaître à peu près toutes les émotions qui contiennent une combinaison des unités d'action présentées dans l'ensemble de données, ce qui contribuera à étendre la recherche dans le domaine de la reconnaissance faciale.

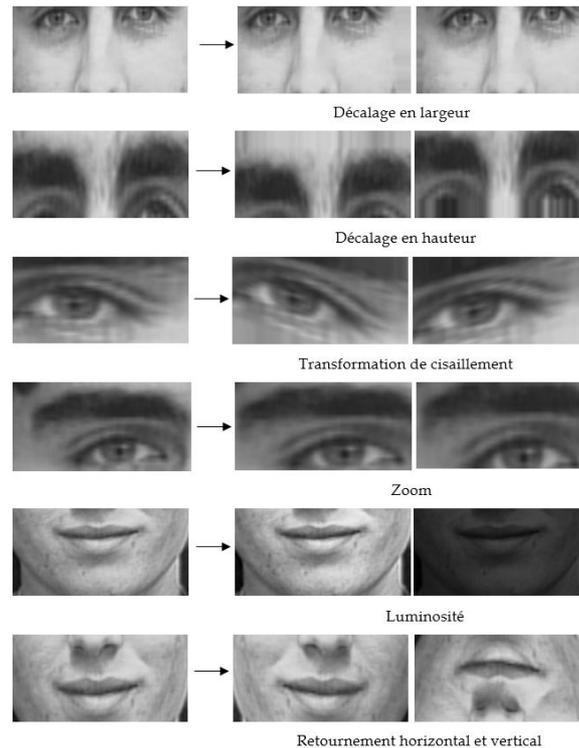


FIGURE 4.9: Ensemble d'images générées par les techniques d'augmentation

4.4 Modèle de classification des états

Nous avons fait le choix de développer un modèle de réseau de neurones basés sur l'architecture **ResNet**, nous avons été influencé par ses excellents résultats dans les domaines lié au visage (Li Lima, 2021 ; Mandal et al., 2021 ; Gruber et al., 2017), aussi vu le type de classification souhaitée où les images des classes sont de plusieurs structures détaillées, ce types de données nécessitent un modèle robuste et performant.

4.4.1 Deep Residual Learning (Resnet)

ResNet a été conçu par Kaiming He et son équipe en 2015 dans un article intitulé Deep Residual Learning for Image Recognition (He et al., 2015). Ils ont proposé une reformulation des couches comme des fonctions résiduelles d'apprentissage en référence aux entrées précédentes, ce qui permet de continuer à apprendre les résidus pour faire correspondre les prévisions avec la réalité, cette méthode a fait ses preuves et montre que ces réseaux résiduels sont plus faciles à optimiser et peuvent gagner plus d'exactitude tout en augmentant la profondeur.

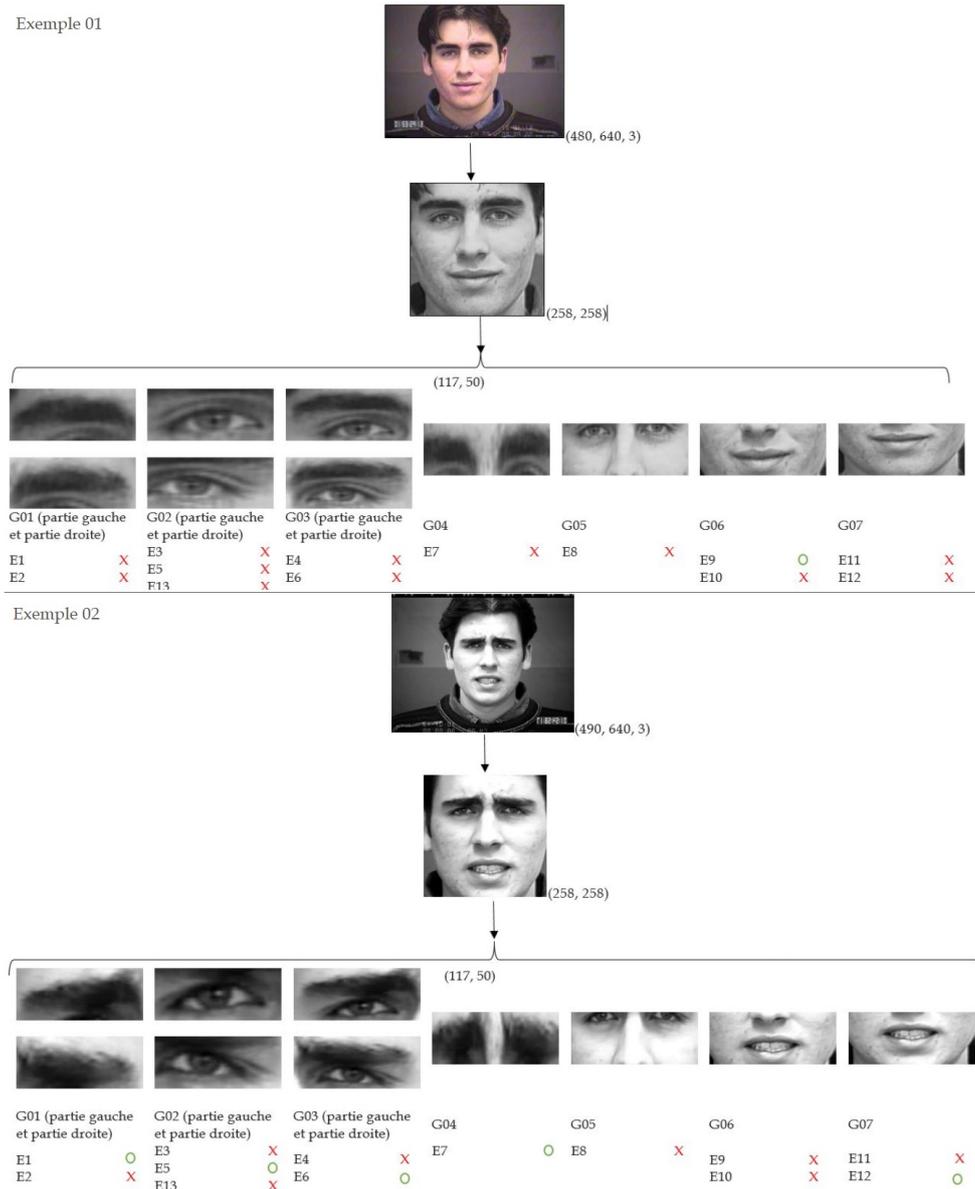


FIGURE 4.10: Traitement d'images : image en gris, extraction du visage puis la division d'images en groupe et redimensionnement (Exemple 01 permet d'extraire l'état 09, l'exemple 02 permet d'extraire les états E1, E5, E6, E7, E12)

Ce résultat a remporté la 1ère place sur la tâche de classement de la compétition ILSVRC 2015 (ImageNet Large Scale Visual Recognition Challenge 2015), ainsi ils ont également remporté les 1ères places sur les tâches de détection ImageNet, de localisation ImageNet, de détection COCO et de segmentation COCO.

L'architecture ResNet conduit à des résultats positifs, le processus devient plus rapide, permet d'éviter l'overfitting et encore augmente l'exactitude contrairement au principe de CNN traditionnelle ou théoriquement l'augmentation du nombre de couches dirige vers

la diminution d'erreur, mais en pratique on observe le contraire la figure 4.11 montre un exemple type. Donc Resnet rend pratique l'affirmation théorique, par l'ajout de saut

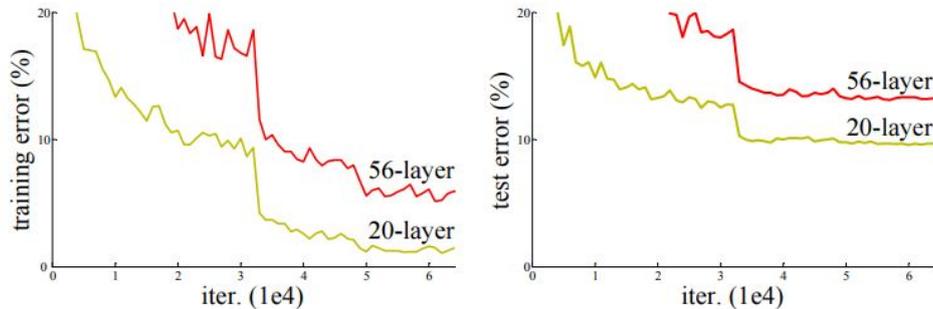


FIGURE 4.11: Erreur d'entraînement (gauche) et erreur de test (droite) sur CIFAR-10, (He et al., 2015)

de couches, la figure 4.12 présente la fonction d'un réseau résiduel, comme il est défini dans le papier original. La formulation de $F(x)+x$ peut être réalisée par des réseaux de neurones à propagation avant avec des connexions raccourcies qui représentent les connexions qui sautent une ou plusieurs couches, ce qui effectuent un mappage d'identité et leurs sorties sont ajoutées aux sorties des couches empilées, cette méthode n'ajoute ni paramètre supplémentaire ni complexité de calcul, d'où vient la force de cette technique. Notre contribution se base sur des données de différentes formes et structures l'utilisation du principe de Resnet, lors de la création de notre architecture permet de bien envisager toutes les particularités des images d'état ce qui nous a aidé à améliorer les résultats.

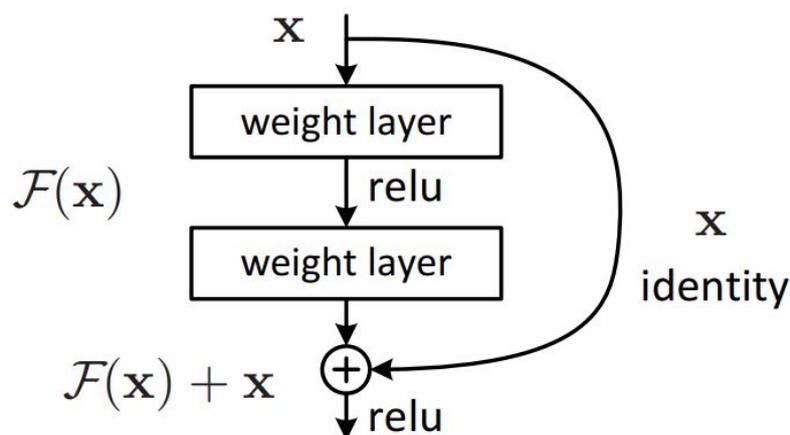


FIGURE 4.12: Residual learning : exemple d'un bloc, (He et al., 2015)

4.4.2 Architecture de classification des états

Nous entraînons le modèle proposé, en utilisant l'ensemble de données créées précédemment, les nouvelles images sont étiquetées avec les états appropriés, puis elles sont alimentées au modèle, comme le montre la figure 4.13 . Cela donne seize blocs de taille de filtre différents, dont on a implémenté un triple layer skips (figure 4.14) qui permet par définition d'éviter plusieurs problèmes comme vanishing gradients (gradient de disparition) ou la saturation de précision et l'overfitting. Nous avons également une couche Max pooling au début et une couche AveragePooling à la fin, malgré la différence de travail entre les deux, ils ont le même objectifs de réduire les dimensions de feature maps créer par la couche convolutionnelle. Au début nous avons utilisé Max pooling afin de détecter si les fonctionnalités apparaissent ou non, par contre l'average pooling se concentre davantage sur les fonctionnalités globales ce qui peut mieux représenter la résistance d'une caractéristique, surtout lorsque la taille de fenêtre est petite ce qui est conforme avec la phase finale de l'architecture. Chaque couche conv contient une phase de normalisation de données garanties par BatchNormalisation (BN), qui présente une technique qui permet de rendre le modèle plus rapide et plus stable grâce à la normalisation des entrées des couches (Ioffe Szegedy, 2015), il permet aussi de diminuer l'importance d'initialisation du poids, et de Régularisé un peu le modèle (Chen et al., 2019). Cette même couche contient aussi la fonction d'activation Relu qui assure la non-linéarité des données diffusées, BN et ReLU sont les deux utilisés pour améliorer la capacité de convergence du modèle. L'utilisation de la couche flatten, permet de convertir la matrice 2D en un vecteur appelé Flatten qui est ensuite traité par la couche entièrement connectée (fc). Enfin, la couche de sortie comporte 14 neurones pour les 14 cas possibles et une fonction d'activation softmax pour produire des prédictions de type probabilité pour chaque classe. Une fois l'apprentissage terminé, la prédiction est effectuée au niveau des images d'état.

La figure 4.14 détaille plus la construction d'un bloc de couches conv2D, la capture commence par une entrées $a^{[l]}$, qui va par la suite représenter l'identité (x) à ajouter à la couche finale avant l'application de la fonction d'activation Relu équation (4.4), pour les entrées et les sorties qui circulent aux couches du centre leur calcul est le même, éq(4.2) et (4.3).

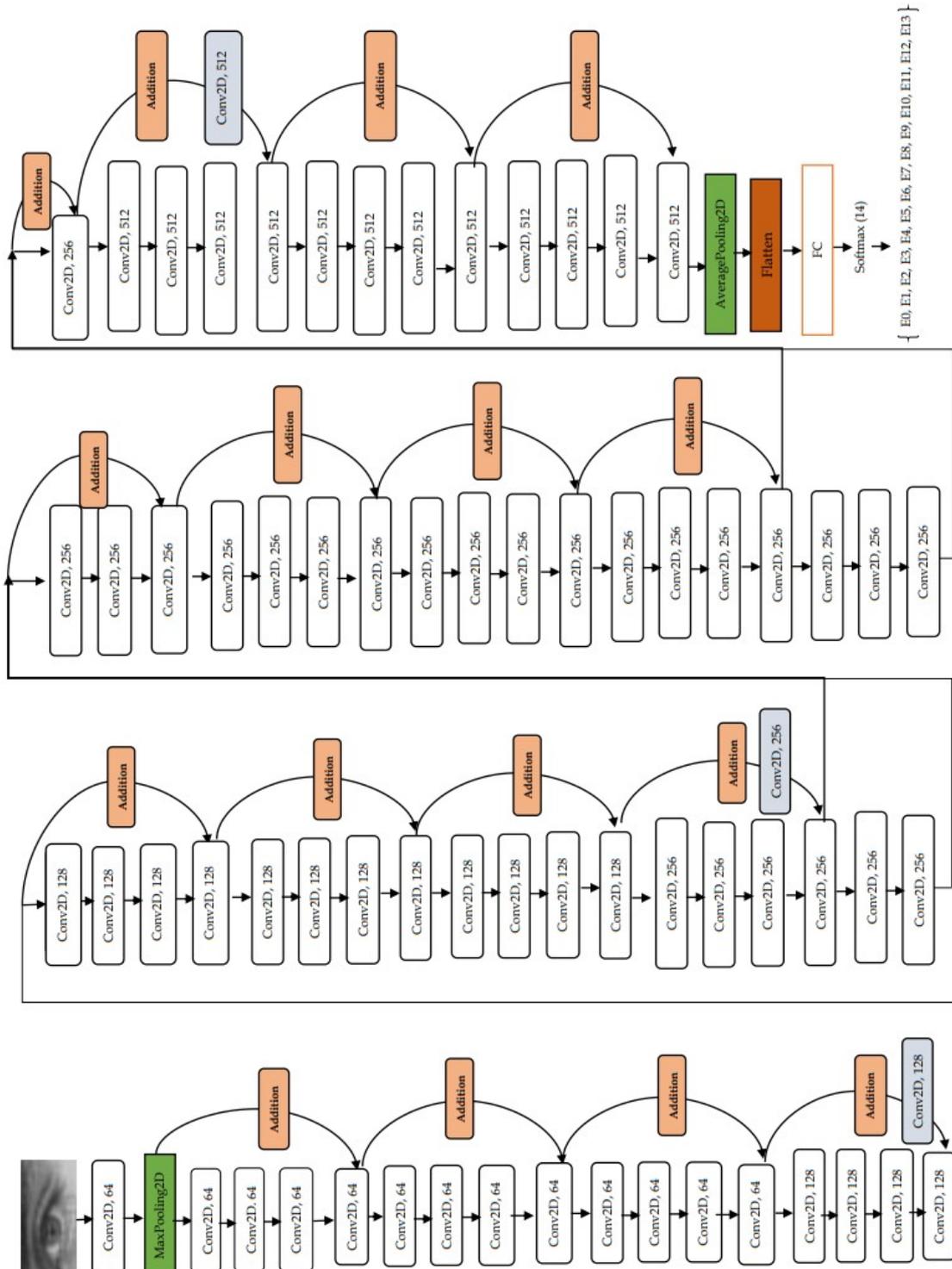


FIGURE 4.13: L'architecture de notre modèle de reconnaissance d'états faciaux adaptés au ResNet. Chaque rectangle est un bloc Conv comprenant une couche convolutionnelle, BatchNormalization et une fonction d'activation (Relu) ces derniers ne sont pas affichés par souci de simplicité, pour le bloc conv en bleu il contient seulement une couche convolutionnelle

$$a^{[l+1]} = g(w^{[l+1]} * a^{[l]} + b^{[l+1]}) = g(z^{[l+1]}) \quad (4.2)$$

$$a^{[l+2]} = g(w^{[l+2]} * a^{[l+1]} + b^{[l+2]}) = g(z^{[l+2]}) \quad (4.3)$$

.

.

$$a^{[l+5]} = g(w^{[l+5]} * a^{[l+4]} + b^{[l+5]}) = g(z^{[l+5]}) \quad (4.4)$$

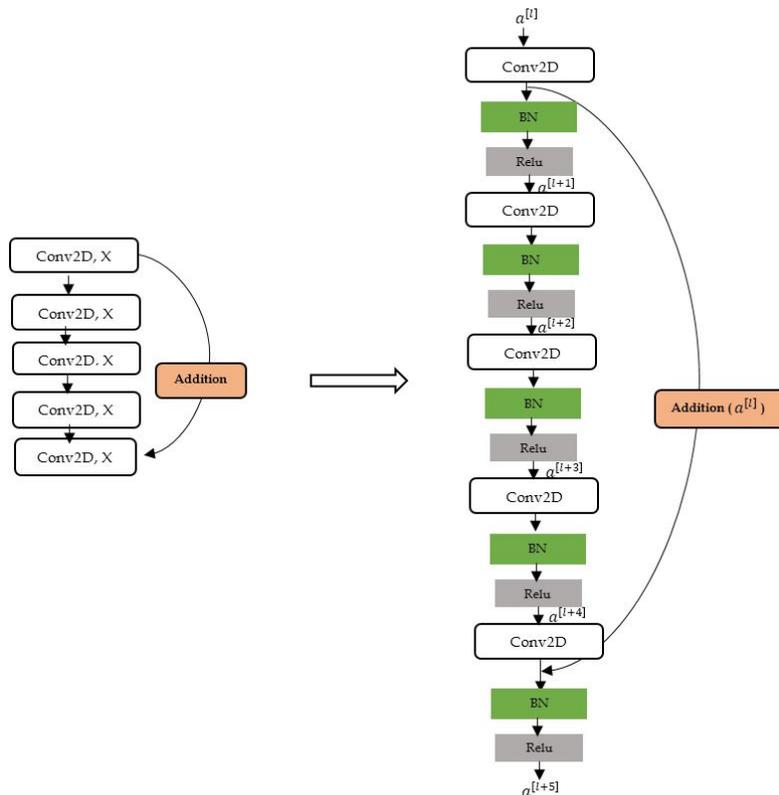


FIGURE 4.14: La description détaillée de la construction d'un bloc

4.5 expérimentation : Résultat et discussion

Nos expérimentations et évaluations de la performance de cette approche de classification d'état faciale ont été réalisées sur Google Colaboratory utilisant Python 3.7.10, comprenant toutes les étapes proposées pour le prétraitement des images (conversion en niveau de gris, détection de visage et recadrage selon les points de repère, extraction de régions, ajustement des bordures, augmentation des données et redimensionnement), et la création, l'apprentissage et le test du modèle, avec une étape de prédiction. Nous avons obtenu 6923 images partagées sur les quatorze états d'une manière équilibrée, ces données

ont été divisé d'un rapport de 70 :30 l'une pour la phase d'apprentissage et l'autre pour la phase d'évaluation. Le modèle présente un taux d'exactitude de 98.36% et un taux d'erreur de 0.11. Les courbes présentées dans la figure 4.15, montre respectivement la courbe d'exactitude de validation et d'erreur de validation sur 200 époques. Dans notre

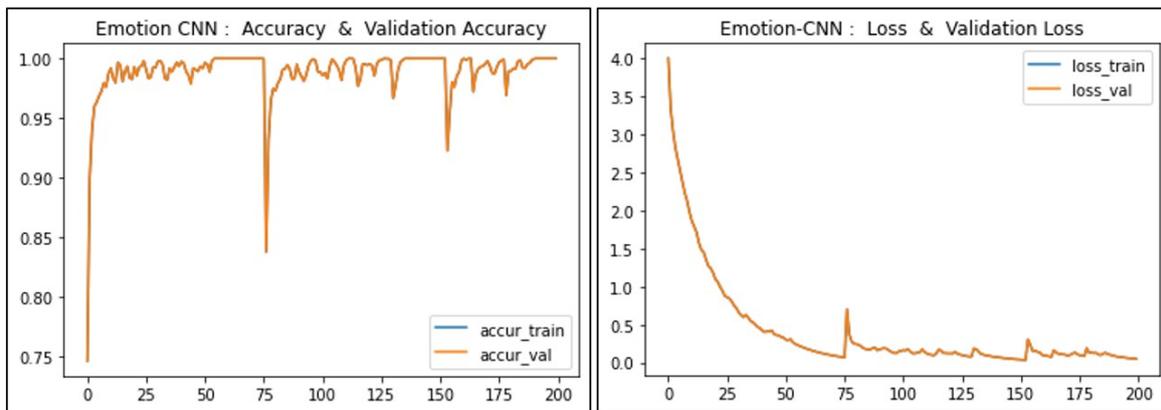


FIGURE 4.15: Les courbes : d'exactitude de la validation (a) et d'erreur de validation (b)

cas, l'exactitude de la validation (a) augmentant progressivement à chaque époque ajoutée, ce qui montre que le modèle a été bien entraîné, de même la courbe (b) montre que l'erreur de validation diminue jusqu'à stable, ce qui montre à nouveau que le processus d'apprentissage se déroule dans les meilleures conditions.

4.5.1 Métriques d'évaluation

Afin d'évaluer l'efficacité et l'effectivité des descripteurs proposés, nous avons utilisé cinq mesures d'évaluation pour bien cerner les performances de notre approche :

Exactitude (Accuracy) :

cette métrique d'évaluation (4.5) est considérée comme la plus fréquente elle dépend du rapport entre le nombre des états bien classés et le nombre total d'états (Zhu et al., 2010).

$$\text{Acc} = \frac{TN+T}{TN+TP+FN+FP} \quad (4.5)$$

La Table 4.7 présente le taux d'exactitude de quelques travaux pour la classification des AU dans le jeu de données CK+, le résultat que nous avons obtenu représente le taux

le plus élevé dans cette liste est considéré comme une bonne réalisation. Chaque méthode citée dans ce tableau présente une approche différente des autres par exemple Wei (2021) présente une nouvelle méthode de reconnaissance des expressions faciales basées sur la connaissance préalable des cartes de saillance et des caractéristiques profondes multicouches dans le réseau CNN. Cui et al (2020) comptent sur la méthode d'optimisation des contraintes pour coder les connaissances génériques sur les dépendances probabilistes des unités d'action dans un réseau bayésien qui va être ensuite intégré dans un cadre d'apprentissage profond. Zhou et al (2017), établissent l'utilité de cartes de caractéristiques sélectives pour les unités d'action dans un CNN profond formé par l'apprentissage par transfert. Chaque approche a son côté positif et négatif, la façon dont les chercheurs traitent ces deux côtés dans leur études contrôle directement le taux de précision, nous présentons dans cette section d'autres mesures pour avoir une meilleure vision du résultat de notre approche.

Méthodes	Exactitude
(Liu et al., 2015)	93.70%
(Khorrami et al., 2015) 06-Classes	95.7% \pm 2.5%
(Zhou et al., 2017) 08-Classes	96.95%
(Dagher et al., 2019)	93.29%
(Xie et Wu, 2019)	95.88%
(Wang et Peng, 2019)	77.07%
(Cui et al., 2020)	97.59%
(Wei, 2021)	93.61%
Notre modèle	98.36%

TABLE 4.7: Comparaison des résultats d'exactitude avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score

Précision et Rappel (precision Recall) :

- Précision : cette mesure d'évaluation (4.6) présente le rapport entre le nombre d'états correctement classés (les vrais positifs) et l'ensemble de tous les états classés (tous les positifs prédits par le modèle), plus le modèle prédit de faux positifs, plus la précision est faible (Hossin et al., 2015).

$$\text{Acc} = \frac{TP}{TP+FP} \quad (4.6)$$

- Rappel : La métrique (4.7) est utilisée pour mesurer la fraction des états positifs qui sont correctement classés (Hossin et al., 2015).

$$\text{Rec} = \frac{TP}{TP+FN} \quad (4.7)$$

La plupart du temps, la précision et le rappel sont des métriques inséparables, la première, mesure la pertinence des résultats, tandis que la deuxième, mesure le nombre de résultats réellement pertinents. Un modèle avec une faible précision et un rappel élevé renvoie à de nombreux résultats faussement prédit, bien que le modèle qui a un taux de précision élevé et un rappel faible, renvoie très peu de résultats la plupart d'entre eux sont correctement prédits. Alors le modèle idéal est celui avec un taux élevé de précisions et de rappels ce qui permet d'envoyer nombreux résultats étiquetés correctement la plus part du temps, comme le cas de nos résultats ce qui peuvent favoriser la reconnaissance des unités d'action liées aux expressions faciales d'apprentissage. Le tableau 4.8, montre les résultats de précision et de rappel pour les 14 états en étude, le modèle proposé présente des résultats très intéressants.

Etat	Précision	Rappel
E0	0.87	0.95
E1	1.00	0.97
E2	0.99	0.93
E3	1.00	1.00
E4	0.97	0.98
E5	0.99	0.98
E6	0.99	0.98
E7	1.00	1.00
E8	0.99	0.99
E9	1.00	1.00
E10	0.99	1.00
E11	0.98	0.99
E12	1.00	1.00
E13	1.00	1.00

TABLE 4.8: Comparaison des résultats d'exactitude avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score

F1-score :

cette mesure (4.8) représente la moyenne harmonique entre les valeurs de rappel et de précision, le nombre de vrais négatifs n'est pas pris en compte c'est pour cela qu'elle est des fois considérée comme une mesure plus intéressante que l'exactitude. Elle est calculée comme suit :

$$F1=2 * \frac{Pre*Rec}{Pre+Rec} \quad (4.8)$$

Même si nous avons calculé la mesure de la précision et du rappel, où ils étaient tous les deux élevés dans les quatorze classes, nous calculons également le score f1 qui est une moyenne harmonique de la précision et du rappel du modèle.

Nous collectons les scores F1 de deux autres modèles d'état d'art (tableau 4.9), dont nous avons présenté leur exactitude dans la section précédente et qui ont utilisé le même ensemble de données (ck+), afin de comparer notre approche avec d'autres méthodes. Ce qui montre que notre méthode proposée surpasse toutes les autres méthodes présentées, en raison du long prétraitement des data, de l'augmentation des données qui assure l'équilibrage des classes et du modèle profond basé sur ResNet qui s'est avéré un excellent résultat lorsqu'il est utilisé pour le domaine de la reconnaissance des expressions faciales.

Matrice de confusion :

Est une matrices carrée avec des lignes et des colonnes qui indiquent les classes du modèle, elle permet de générer le nombre de faux positifs, de faux négatifs, de vrais positifs et de vrais négatifs, ce qui permet d'avoir plus de détails et retirer les conflits afin d'améliorer le modèle au maximum.

Afin de terminer les mesures résultantes de notre modèle nous allons calculer la matrice de confusion. Le tableau 4.10et la figure 4.16 représentent la matrices de confusion correspondant au taux de reconnaissances atteint pour le jeu de données CK+, comme nous pouvons voir, nous avons obtenu de très bons résultats pour la majorité des états avec un taux moyen proche de 100%, à l'exception de quelques états surtout l'état neutre E0, et E2 (AU1+ AU2), reconnus respectivement avec des taux moyens de 93.79% et 93.1%. E0

Etat	Unité d'action	(Wang et Peng, 2019)	(Cui et al., 2020)	Notre modèle
E0	Neutre	-	-	0.91
E1	AU1	0.93	0.93	0.99
E2	AU1 +AU2	0.89(AU2)	-	0.96
E3	AU5	0.79	-	1.00
E4	AU1+AU2+ AU5	-	-	0.97
E5	AU7	0.44	0.63	0.99
E6	AU1 + AU7	-	-	0.99
E7	AU4	0.74	0.81	1.00
E8	AU6	0.53	0.78	0.99
E9	AU12	0.82	0.92	1.00
E10	AU14	-	-	1.00
E11	AU26	-	-	0.99
E12	AU20 + AU25	0.93(AU25)	-	1.00
E13	AU43 ou AU61 ou AU63 ou AU64	-	-	1.00

TABLE 4.9: Evaluation et comparaison de f1-score avec des méthodes d'état d'art sur le jeu de données CK+. Les chiffres en gras indiquent le meilleur score

est confondues avec les états (E1, E4, E8, E10, et E11) avec des taux d'erreurs très petits qui ne dépassent pas le 2%. Pour le cas d'E2, cet état est confondu qu'avec l'état E0 avec un taux d'erreur de 6.8%.

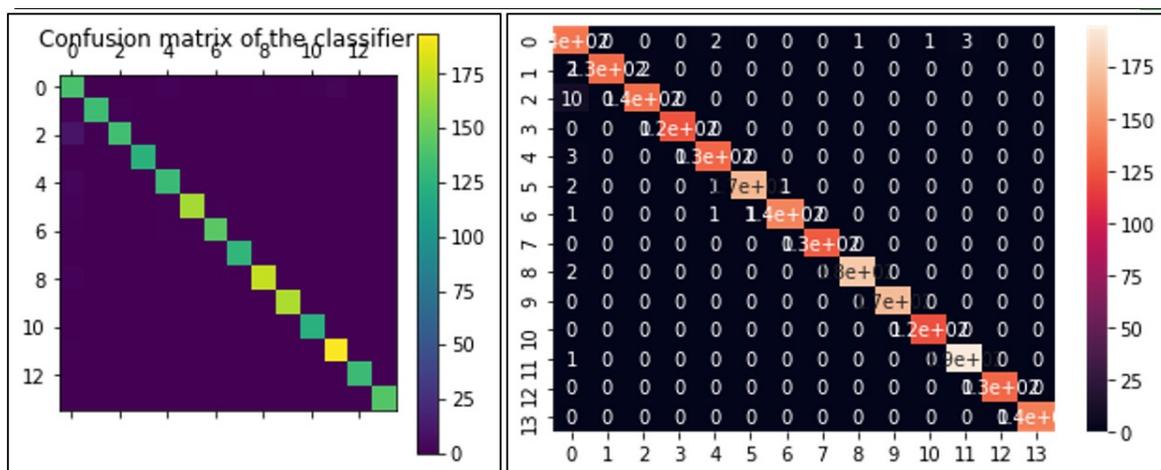


FIGURE 4.16: Matrice de confusion

E	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	136	2	0	0	2	0	0	0	1	0	1	3	0	0
1	2	134	2	0	0	0	0	0	0	0	0	0	0	0
2	10	0	135	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	125	0	0	0	0	0	0	0	0	0	0
4	3	0	0	0	132	0	0	0	0	0	0	0	0	0
5	2	0	0	0	0	169	1	0	0	0	0	0	0	0
6	1	0	0	0	1	1	144	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	128	0	0	0	0	0	0
8	2	0	0	0	0	0	0	0	177	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	170	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	123	0	0	0
11	1	0	0	0	0	0	0	0	0	0	0	194	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	132	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	142

TABLE 4.10: Matrice de confusion

Remarque :

- TP : le nombre d'états correctement classés (True positives).
- FP : le nombre d'états mal classés (False positives).
- TN : le nombre d'états correctement classés mais qui ne sont pas classés par l'approche proposée (True negatives).
- FN : le nombre d'états mal classés et qui ne sont pas classés par l'approche proposée (False negatives).

La prédiction

L'objectif de cette étude est d'utiliser uniquement une caméra pour détecter les expressions faciales des apprenants relatives à l'apprentissage. Au cours de la phase de prédiction, les images d'états de prédiction sont collectées de deux façons :

- La première partie a été collectée à partir de deux jeux de données Ck+ et une partie de la base de donnée DISFA dont elle subit les mêmes changements décrits précédemment, ce qui aide à évaluer l'efficacité et l'effectivité du modèle, la prédiction est réalisée sur 225 images cette phase donne un taux positif de 81.33%.
- La deuxième partie a été collectée à l'aide de quatre lycéens volontaires pour imiter

les émotions d'apprentissage. Nous avons expliqué les mouvements à faire et utilisé un smartphone Samsung M-31s pour capturer les images. La figure 4.17 montre deux d'entre eux qui expriment les émotions d'apprentissage, nous avons collecté 50 images au total, dans des conditions différentes de la base de données Ck+ (différent éclairage, meilleure résolution), ensuite ces images ont subi le même prétraitement discuté précédemment pour extraire la base de données d'états, le taux de prédictions pour cette expérience était 65.21%.



FIGURE 4.17: Quelques émotions d'apprentissage, imitées par des lycéens

4.5.2 Synthèse et discussion

L'objectif principal de ce travail est de réaliser un modèle permettant de classifier les unités d'actions faciales afin de l'utiliser prochainement pour la reconnaissance des émotions faciales d'apprentissage dont nous avons limité spécifiquement dans cette étude par (satisfaction, frustration, confusion, ennui, nervosité et surpris).

Pour aboutir à notre objectif nous avons défini des états qui représentent des groupes d'unités d'actions qui se déroulent en même temps dans la même région ou dans deux régions voisines. En effet, l'approche présentée vise à étendre l'utilisation des données faciales, en particulier les émotions non habituelles, nous pouvons donc utiliser les bases de données ouvertes, disponibles et massives, pour entraîner nos systèmes, même si nous ne cibons pas nécessairement le même type d'émotions, pour la classification des états choisis, l'évaluation du système a montré des résultats très intéressants.

Le résumé de ce travail est décrit ci-dessous :

- La classification des expressions faciales d'apprentissage est une tâche complexe vu la rareté des données d'un côté et la complication de la tâche elle-même de l'autre

côté, plusieurs solutions ont été proposées mais chacune a ses propres inconvénients. Donc, nous avons essayé de trouver une solutions permettant d'étendre l'utilisation des données des base de données qui existent déjà et les rendre plus flexibles pour notre objectif, nous avons d'abord sélectionné les émotions faciales d'apprentissage avec lesquelles nous allons travailler, ensuite nous avons extrait leurs unités d'action, en définissant leurs états et régions adéquates, et enfin nous avons choisi la base de données CK+, largement utilisée pour la classification des émotions faciales et des unités d'action.

- Ce travail nécessite un prétraitement d'images afin de créer la nouvelle base de données qui sera par la suite l'ensemble de données d'entrée du modèle.
- Nous avons choisi de créer le modèle d'apprentissage profond en se basant sur l'architecture ResNet, notre choix est influencé d'une part, par les structures détaillées des images des classes et de l'autre part, par les excellents résultats de ce réseau profond dans le domaine lié au visage.
- Nous avons abouti à des résultats satisfaisants, après être passé par une comparaison des performances selon différents métriques ; l'exactitude, la précision et le rappel et encore le F1-Score et la matrice de confusion.

Conclusion

Ce chapitre propose un système de classification d'états faciaux définis selon les unités d'actions qui présentent des émotions complexes (émotions d'apprentissage), basé sur l'architecture d'apprentissage en profondeur ResNet. Ce systèmes est composé de deux phases complémentaires, la première est celle de prétraitement d'images de la base de données CK+, qui passe par plusieurs étapes (conversion en niveau de gris, détection de visage et recadrage selon les points de repère, extraction de régions, ajustement des bordures , augmentation des données et redimensionnement), pour créer la nouvelle base de données qui va à son tour alimenter notre architecture d'apprentissage en profondeur, la deuxième consiste à la création de modèle d'apprentissage.

Les résultats extraits depuis plusieurs métriques montrent une bonne qualité en mesure

de performance et d'efficacité de classification, ainsi que les résultats de prédiction.

Comme perspectives à courts termes, nous allons tester notre méthode sur un environnement d'apprentissage réel afin de mesurer la robustesse de notre système sur des conditions différentes l'éclairage, la faible résolution, et les actions spontanées des apprenants.

Conclusion générale

Les travaux qui ont été présentés dans cette thèse se sont inscrits dans le contexte des systèmes d'apprentissage en ligne, qui de nos jours devient un domaine de recherches très actif contrairement aux années précédentes, nous avons présenté deux contributions complémentaires.

Architecture CD-Mobile MOOC

Alors que certaines personnes peuvent choisir de ne pas utiliser l'apprentissage en ligne, il reste toujours un domaine évolutif qui s'introduit dans nos vies d'une manière continue. À la suite de nos recherches préliminaires, nous avons pris soin de déterminer si les lycéens algériens sont intéressés par ce type d'apprentissage et est-ce que l'utilisation de ressources d'apprentissage ouvert (OER) en mobile (M-learning) améliore la qualité éducative de ces derniers. En effet, nous avons commencé en première contribution à poser cette question en créant un questionnaire pour les élèves de deux lycées différents, et en essayant de déduire leur aptitude d'utiliser les plateformes éducatives, comme il reste encore un domaine pas trop utilisé en Algérie. Les résultats étaient affirmatifs. L'utilisation des ressources éducatives ouvertes sur mobile n'améliore pas seulement les conditions d'apprentissage, mais elle introduit aussi un impact positif sur la qualité éducative de tous les élèves, pour le reste de la contribution nous avons proposé une architecture basée sur le Cloud computing mobile afin de surpasser quelques lacunes déduites lors de la première expérience avec les élèves et aussi dériver d'autres recherches.

L'architecture proposée sous le nom de CD-Mobile MOOC (Context Driven-Mobile Massive Open Online Cours), est particulièrement présentée pour éviter les limites étudiées comme, la taille d'écran, la batterie, le stockage et la distraction, en étudiants attentive-

ment chaque composant de l'architecture et leur coordination, ensemble nous avons abouti à une modélisation sur quatre niveaux. D'abord le niveau apprenant chargé de créer une liaison de connaissance entre cet individu et le système, dans ce dernier ont rempli un formulaire classique (nom, niveau, genre, éducation ...), à ce niveau le système permet même d'extraire des informations sur l'environnement d'apprenants via les capteurs mobile GPS et le microphone afin de définir le niveau du bruit dans son entourage, et sa caméra pour capturer le visage. Ensuite le niveau interface qui assure l'interaction entre l'apprenant et le système, permet de collecter des données sur l'appareil, l'apprenant, son contexte, ses traces afin de créer le module de traçabilité et la catégorisation. Le troisième niveau introduit les différents modèles de traitement et de personnalisation, ces modèles permettent de s'occuper de la gestion des données extraites du comportement de l'utilisateur et de son contexte, pour enfin arriver au niveau de stockage de données circulant dans le système assurée par le Cloud. Toutes les étapes précédentes fonctionnent sur le Cloud Mobile qui libère les appareils mobiles des apprenants et des enseignants du traitement et du stockage permettant de fournir des services d'apprentissages personnalisés selon les préférences de l'utilisateur. De plus l'utilisation des MOOC comme des cours formalisés basés sur les vidéos aide à diminuer la distraction, regarder les vidéos est un acte habituel chez les apprenants, 70% des vidéos regardées sur YouTube proviennent de smartphone, cette habitude va aider à dépasser l'obstacle de la taille de l'écran.

Pour la première contribution, nous avons proposé une étude statistique afin de déterminer le comportement et l'acceptation d'OER sur mobile pour les lycéens. Grâce à cette étude c'était aussi possible de dériver les limites rencontrées dans cette expérience, nous avons remarqué un comportement positif et prometteur envers ce type d'apprentissage, Pour la deuxième partie de contribution nous avons proposé une architecture dont nous avons étudié toutes les possibilités de la mettre en places afin de surpasser les lacunes rencontrées par les élèves.

Limitations

Bien que nous ayons comparé notre travail avec plusieurs autres architectures, certains ont été implémentés et d'autres n'ont pas été mis en place, nous n'avons aucune certitude d'avoir trouvé la solution optimale pour la personnalisation de l'apprentissage en ligne. Nous croyons que ce problème est fortement liée à la nature des apprenants ce qui nous a dirigé à la base vers le questionnaire, nous étions très objectifs lors de la discussion de notre proposition, de même lors de la comparaison avec les autres recherches. Le fait que c'était impossible de tester l'ensemble de propositions du modèle sur une plateforme réelle dans les conditions actuelles de la recherche, cela nous a ajouté plus de responsabilité pour discuter chaque fonctionnalité avec une précision et une exactitude majeure.

Les émotions chez les apprenants

L'un des modules de l'architecture précédente est le module de reconnaissance des émotions d'apprenants, qui représente la méthode d'expression la plus utilisée et celle comprise par les humains. Cette méthode de communication est très utilisée lors du processus d'apprentissage, tous les enseignants adaptent leurs cours selon les expressions faciales de leurs élèves, dans le domaine de l'apprentissage en ligne aussi, l'utilisation de émotions faciales est très répandue. Tandis que certains chercheurs du domaine peuvent choisir d'utiliser les émotions universelles pour la reconnaissance des émotions des apprenants, il reste un choix non exact qui peut être plus logiquement catégorisé sous le domaine de FER. À la suite de nos recherches préliminaires, nous nous sommes intéressés à déterminer les émotions qui ont une relations avec le domaine éducatif et les expressions faciales de l'apprenant, si la machine sera capable de bien déterminer les émotions nous allons remarquer des améliorations primordiales dans la qualité d'interactions au niveau des plateformes éducatives.

L'objectif de cette contribution est de réaliser un modèle permettant la reconnaissance des unités d'actions faciales afin de l'utiliser prochainement pour classifier des émotions d'apprentissage que nous avons limité par (satisfaction, frustration, confusion, ennui, nervosité et surpris), ce travail est composé de deux parties ; la première est celle de pré-

traitement elle permet de générer la base de données. La deuxième consiste à la création de modèle d'apprentissage. Comme les bases de données qui contiennent des émotions propres à l'apprentissage sont rares, pour aboutir à notre but nous sommes partis de l'idée ; que chaque émotion faciale cause un mouvement musculaire, ces mouvements sont codés par FACS (Facial action coding system) qui génère des unités d'action. Nous avons développé un système qui permet de créer nos données à partir d'un jeu de données qui existe déjà (CK+), et qui contient des émotions de base et les descriptions des unités d'action dans chaque image. Ce processus passe par plusieurs étapes, (la conversion d'images en niveau de gris, la détection de visage et le recadrage selon les points de repères, l'extraction de régions, l'ajustement des bordures, l'augmentation des données et le redimensionnement). Cette nouvelle base de données est ensuite alimentée à notre modèle d'apprentissage profond qui se base sur l'architecture ResNet, le choix de cette architecture est influencé par le type de données et aussi par ses excellents résultats déjà aboutis au domaine lié au visage. La reconnaissance et le regroupement des unités d'action dans un visage quelconque permettent de reconnaître des nouvelles émotions. Nous avons abouti à des résultats satisfaisants, après être passés par une comparaison des performances selon différents métriques ; l'exactitude, la précision et le rappel et encore le F1-Score et la matrice de confusion.

Limitations

Il est important de garder à l'esprit que la classification des unités d'actions que nous avons effectuée, a été dans le contexte d'apprentissage précis et limitée par les émotions choisies au début de la contribution, dont nous avons opté vers des émotions de réalisation, mais l'ensemble de données résultants reste tout de même valide pour une grande échelle d'émotions dont leurs unités d'actions existent dans cette base de données. De plus notre travail n'a pas été testé sur un environnement d'apprentissage réel afin de mesurer sa robustesse sur des conditions différentes comme l'éclairage, la faible résolution, et les actions spontanées des apprenants.

Perspectives

Après avoir détaillé nos contributions, leurs objectifs et certaines de leurs limites, nous présentons les perspectives à court terme qui en découlent. Nous proposons d'abord notre système de reconnaissance des unités d'action pour la classification des émotions choisies préalablement, ensuite nous comptons élargir ce système afin de pouvoir reconnaître les émotions en considérant de nouveaux domaines éducatifs.

De plus, la prise en compte des images réelles où les images sont plus spontanées, dans diverses conditions, se diriger naturellement vers l'utilisation des vidéos à la place de photos, ce qui va donner des résultats plus exacts. Nous proposons également d'intégrer les deux contributions et les tester dans un cours en ligne réel.

Enfin, nous revenons vers l'architecture Cloud mobile proposées, nous avons confiance que sa mise en œuvre va ajouter des qualités à l'expérience d'apprentissage en ligne, comme elle se base sur une proposition réaliste liée à l'apprenant, son environnement et les capteurs de son mobile, nous allons examiner les potentiels de l'implémenter et l'intégrer au niveau de la plateforme éducative de l'université.

Publications et Communications

Communication nationale :

Nadira, B., Makhlouf, D., Amroune, M. (2016). CD-MOOC : Context driven Massive Open Online Courses Amélioration du taux d'abandon des apprenants. Séminaire national sur la simulation numérique dans les sciences appliquées, (SNSA I-2016), Guelma, Algeria, Décembre 2016.

Communication internationale :

Nadira, B., Makhlouf, D., Amroune, M. (2017). Recent Learning Environments : A Survey. 3rd International Conference on Networking and Advanced Systems, (ICNAS-2017), Annaba, Algeria 13-14 December 2017

https://www.researchgate.net/publication/321531740_Recent_Learning_Environments_A_Survey

Publication Internationales :

Nadira, B., Makhlouf, D., Amroune, M. (2021). Personalized Online Learning : Context Driven Massive Open Online Courses. International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 16(6), 1-15.

<http://doi.org/10.4018/IJWLTT.20211101.oa8>

Bibliographie

- Abbas K., Verdier C. and Flory A. 2007. "Exploiting profile modeling for web-based information system". In Proceedings of the International Workshop on Personalized Access to Web Information (PAWI'07), p. 591- 596, Nancy, France
- Abu-Al-Aish, A., (2014). Toward mobile learning deployment in higher education. Brunel University, School of Information Systems. Computing and Mathematics.
- Agrawal R, Srikant R (1995) Mining sequential patterns. In : Proceedings of the eleventh international conference on data engineering. pp 3-14
- Ahmed, I., Yakubu, A., Yakubu, M. (2020). Cloud-based mlearning : A pedagogical tool to manage infrastructural limitations and enhance learning Safiya. 16(2), 48-67.
- Al-Darraji, S., Berns, K., & Rodi, A. (2017). Action Unit Based Facial Expression Recognition Using Deep Learning. International Conference on Robotics in Alpe-Adria Danube Region, 540(november), 413-420. https://doi.org/10.1007/978-3-319-49058-8_45
- Ali, R. A., Rafie Muhammad, M. A. (2017). Investigating the Perception of Students Regarding MLearning Concept in Egyptian Schools. 112-122.
- Alkore Alshalabi, Ibrahim Hamada, Samir Elleithy, Khaled Badara, Ioana Moslehpour, Saeid. (2018). Automated Adaptive Mobile Learning System using Shortest Path Algorithm and Learning Style. International Journal of Interactive Mobile Technologies (iJIM). 12. 4. 10.3991/ijim.v12i5.8186.
- Ameyed, D., M. Miraoui and C. Tadj (2015). "A Spatiotemporal Context Definition for Service Adaptation Prediction in a Pervasive Computing Environment." arXiv preprint arXiv :1505.01071.
- Amin, Mohammed Abu Bakar, Kamalrulnizam AL-Hashimi, Haider. (2013). A review of mobile cloud computing architecture and challenges to enterprise users. 240-244. 10.1109/IEEEGCC.2013.6705783.
- Aneja, D., Colburn, A., Faigin, G., Shapiro, L., Mones, B. (2017). Modeling stylized character expressions via deep learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10112 LNCS, 136-153. https://doi.org/10.1007/978-3-319-54184-6_9
- Antipov, G. (2017). Deep learning for semantic description of visual human traits Gregory Antipov To cite this version : HAL Id : tel-01725853 Apprentissage Profond pour la Description Sémantique des Traits Visuels Humains.

- AWT (agence wallonne des telecommunications) Qu'est ce que l'e-learning 28/01/2008 consultée 12/02/2017
- Azizi, S. and Khatony, A., (2019). Investigating factors affecting on medical sciences students' intention to adopt mobile learning. *BMC Medical Education*, 19(1).
- Bahreini, K., Nadolski, R., Westera, W. (2016). Towards multimodal emotion recognition in E-learning environments. *Interactive Learning Environments*, 24(3), 590-605. <https://doi.org/10.1080/10494820.2014.908927>
- Baldauf, M., S. Dustdar and F. Rosenberg (2007). "A survey on context-aware systems." *International Journal of Ad Hoc and Ubiquitous Computing* 2(4) : 263-277.
- Barbosa, D. N. F., et Geyer, C. F. R. (2005). Pervasive personal pedagogical agent : mobile agent shall always be with a learner. In *Proceedings IADIS International Conference Mobile Learning* (pp. 281-285).
- Baumard, M. (2020). Privé d'école, l'Afrique francophone parie sur la Khan Academy et son enseignement des sciences. *Le Monde.fr*. Retrieved 16 April 2020, from <https://www.lemonde.fr/afrique/article/2020/03/30>
- Beniwal, R., Sharma, S., Raj, D. (2018). M-Learning in Education : Benefits and Challenges. *2nd International Conference on Advancement in Engineering, Applied Science and Management (ICAEASM 2018)*, April, 36-40.
- Bezboruah, Tulshi. (2011). *Mobile Computing : The Emerging Technology, Sensing, Challenges and Applications*. 4. 165-174.
- Boruff, J. and Storie, D. (2014). Mobile devices in medicine : a survey of how medical students, residents, and faculty use smartphones and other mobile devices to find information. *Journal of the Medical Library Association : JMLA*, 102(1), pp.22-30.
- Briz -Ponce, L., Pereira, A., Carvalho, L., Juanes-Méndez, J. and Garca-Pealvo, F. (2017). Learning with mobile technologies - Students' behavior. *Computers in Human Behavior*, 72, pp.612-620.
- Buche, C. (2005). *Un système tutoriel intelligent et adaptatif pour l'apprentissage de compétences en environnement virtuel de formation : Vol. Thèse en I.* <http://www.enib.fr/buche/>
- Bui, H. M., Lech, M., Cheng, E., Neville, K., Burnett, I. S. (2016). Using grayscale images for object recognition with convolutional-recursive neural network. *2016 IEEE 6th International Conference on Communications and Electronics, IEEE ICCE 2016*, 321-325. <https://doi.org/10.1109/CCE.2016.7562656>
- Buyya, R., 2014. Heterogeneity in Mobile Cloud Computing : Taxonomy and Open Challenges. *IEEE COMMUNICATIONS SURVEYS TUTORIALS*, 16(1), pp. 1-2
- Campbell, M., Hoane, A. J., Hsu, F. H. (2002). *Deep Blue. Artificial Intelligence*. [https://doi.org/10.1016/S0004-3702\(01\)00129-1](https://doi.org/10.1016/S0004-3702(01)00129-1)
- Canalblog. (2017). Octobre écoles d'autrefois terres d'aujourd'hui. *Lartésien.canalblog.com*. Retrieved 16 April 2020, from <http://lartésien.canalblog.com/archives/>

2017/10/22/35794861

- Candra Kirana, K., Wibawanto, S., Wahyu Herwanto, H. (2018). Facial Emotion Recognition Based on Viola-Jones Algorithm in the Learning Environment. Proceedings - 2018 International Seminar on Application for Technology of Information and Communication : Creative Technology for Human Life, ISEmantic 2018, 406-410.
<https://doi.org/10.1109/ISEMANTIC.2018.8549735>
- Chaari, T., Laforest, F., Flory, A. (2005). Adaptation des applications au contexte en utilisant les services web. ACM International Conference Proceeding Series, 120, 111-118.
<https://doi.org/10.1145/1102613.1102638>
- Champin, P., Mille, A. and Prié, Y. (2013). Vers des traces numériques comme objets informatiques de premier niveau : une approche par les traces modélisées. *Intellectica. Revue de l'Association pour la Recherche Cognitive*, 59(1), pp.171-204.
- Chuang, Y. T. (2015). SSCLS : A smartphone-supported collaborative learning system. *Telematics and Informatics*, 32(3), 463-474. <http://doi.org/10.1016/j.tele.2014.10.004>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Coppens, L. (2017). Detecting affective states from facial expressions in second language education (Issue January). <http://arno.uvt.nl/show.cgi?fid=143349>
- Coutaz, J. R. G. (2005). *L'informatique sensible au contexte.*, p.7
- Cui, Z., Song, T., Wang, Y., Ji, Q. (2020). Knowledge Augmented Deep Neural Networks for Joint Facial Expression and Action Unit Recognition. *NeurIPS*.
- Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (2008). Emote aloud during learning with AutoTutor : Applying the Facial Action Coding System to cognitive Affective states during learning. *Cognition and Emotion*, 22(5), 777-788.
<https://doi.org/10.1080/02699930701516759>
- Csikszentmihalyi, M. (1990). *Flow : The psychology of optimal experience*. New York : Harper Row.
- D. Jajal, T. (2018). Distinguishing between Narrow AI, General AI and Super AI. Mapping Out 2050. Retrieved 14 April 2020, from <https://medium.com/mapping-out-2050/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22>
- Dana, G., Adam, P., Nikole, H. (2020). As School Moves Online, Many Students Stay Logged Out. *Nytimes.com*. Retrieved 19 August 2020, from <https://www.nytimes.com/2020/04/06/us/coronavirus-schools-attendance-absent.html>
- Davies, B. S., Rafique, J., Vincent, T. R., Fairclough, J., Packer, M. H., Vincent, R., Haq, I. (2012). Mobile Medical Education (MoMed) - how mobile information resources contribute to learning for undergraduate clinical students - a mixed methods study. *BMC Medical Education*, 12(1), 1. <http://doi.org/10.1186/1472-6920-12-1>
- Davis, F., Bagozzi, R. and Warshaw, P., 1989. User Acceptance of Computer Technology : A Comparison of Two Theoretical Models. *Management Science*, 35(8), pp.982-1003.

- DeBoer, Jennifer Stump, Glenda Breslow, Lori Seaton, Daniel. (2013). Diversity in MOOC Students' Backgrounds and Behaviors in Relationship to Performance in 6.002x.
- Dey, K. A., Abowd, D. G. (2000). Towards a better understanding of context and context awareness. In *Computer Human Interactions (CHI2000) Workshop on the What, Who, Where and How of ContextAwareness*.
- Di Leo, I., Muis, K.R., Singh, C.A., Psaradellis, C., Curiosity... Confusion? Frustration! The role and sequencing of emotions during mathematics problem solving, (2019), *Contemporary Educational Psychology*,
doi : <https://doi.org/10.1016/j.cedpsych.2019.03.001>
- Dissanayake, D. and Velananda, Y., 2020. Critical Success Factors for Performance Oriented M-Learning in Sri Lanka. *Journal of Educational and Social Research*, 10(2), p.112.
- D'Mello, S., Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), 145-157.
- Domingo, M. G., Garganté, A. B. (2016). Exploring the use of educational technology in primary education : Teachers' perception of mobile technology learning impacts and applications' use in the classroom. *Computers in Human Behavior*, 56, 21-28.
<http://doi.org/10.1016/j.chb.2015.11.023>
- Drissi, S. and Amirat, A. (2016) 'An experimental study to evaluate learning style personalisation in web-based adaptive e-learning systems', *Int. J. Innovation and Learning*, Vol. 20, No. 1, pp.1-25.
- Ekman, P. Friesen, W. V. (1976). Measuring Facial Movement. *Environmental Psychology and Nonverbal Behavior*, 1(1), pp. 56-75.
- Ekman, P., Friesen, W.V. (1971). Constants across cultures in the face and emotion. *J. Personal. Soc. Psychol.* 17, 124-129. DOI <https://doi.org/10.1037/h0030377>
- Ekman, P., Friesen, W.V. (1978). *Facial Action Coding System : A Technique for the Measurement of Facial Movement ; Consulting Psychologists Press : Palo Alto, CA, USA*
- El Hammoumi, O., Benmarrakchi, F., Ouherrou, N., El Kafi, J., El Hore, A. (2018). Emotion Recognition in E-learning Systems. 2018 6th International Conference on Multimedia Computing and Systems. <https://doi.org/10.1109/icmcs.2018.8525872>
- El Naqa I., Murphy M.J. (2015) What Is Machine Learning?. In : El Naqa I., Li R., Murphy M. (eds) *Machine Learning in Radiation Oncology*. Springer, Cham.
https://doi.org/10.1007/978-3-319-18305-3_1
- Elkhateeb, M., Shehab, A., El-bakry, H. (2019). Mobile Learning System for Egyptian Higher Education Using Agile-Based Approach. *Education Research International*, 2019, 1-13. doi :10.1155/2019/7531980
- Elliot N., Friedman R., Brillet V. (2005), " Irony and Asynchronicity : Interpreting Withdrawal Rates in e-Learning Courses ", *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, p. 459-465.

- Fan, Y., Lam, J. C. K., Li, V. O. K. (2018). Multi-region ensemble convolutional neural network for facial expression recognition. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11139 LNCS, 84-94. https://doi.org/10.1007/978-3-030-01418-6_9
- Fawaz Paraiso. *soCloud : une plateforme multi-nuages distribuée pour la conception, le déploiement et l'exécution d'applications distribuées à large échelle*. Génie logiciel [cs.SE]. Université des Sciences et Technologie de Lille - Lille I, 2014. Français.tel-01009918v1
- Ferrer, D. (2019). History of Online Education. *TheBestSchools.org*. Retrieved 16 April 2020, from <https://thebestschools.org/magazine/online-education-history/>.
- Fishbein, M., and Icek, A. (1975). *Belief, Attitude, Intention and Behavior : An Introduction to Theory and Research*. Reading, MA : Addison-Wesley.
- Florian, C., Marie, L., Nathalie, G. and Jean-Charles, M. (2015). Mise en place de la personnalisation dans le cadre des MOOCs. 7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain. Agadir, Maroc. pp.291-300
- Fouh, E., Breakiron, D. A., Hamouda, S., Farghally, M. F., Shaffer, C. A. (2014). Exploring students learning behavior with an interactive etextbook in computer science courses. *Computers in Human Behavior*, 41, 478- 485.
<http://doi.org/10.1016/j.chb.2014.09.061>
- France Intelligence artificielle Rapport de synthese (2017). Depuis <https://www.vie-publique.fr/rapport/36456-france-intelligence-artificielle-rapport-de-synthese>.
- G.J Hwang, C.C Tsai, S.J.H Yang, "Criteria, strategies and research Issues of context aware ubiquitous learning". *Educ. Technol. Soc*,2008.
- Gagnon, M., Gosselin, P., Maassarani, R. (2015). *The Journal of Genetic Psychology : Research and Theory on Human Development Children's Ability to Recognize Emotions From Partial and Complete Facial Expressions*. September.
<https://doi.org/10.1080/00221325.2014.941322>
- Gehrig, T., Al-Halah, Z., Ekenel, H. K., Stiefelhagen, R. (2015). Action unit intensity estimation using hierarchical partial least squares. 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015.
<https://doi.org/10.1109/FG.2015.7163152>
- Giesbers, B., Rienties, B., Tempelaar, D. and Gijsselaers, W. (2013). A dynamic analysis of the interplay between asynchronous and synchronous communication in online learning : The impact of motivation. *Journal of Computer Assisted Learning*, 30(1), pp.30-50.
- Grafsgaard, J. F., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., and Lester, J. C. (2013). Automatically Recognizing Facial Indicators of Frustration : A Learning-centric Analysis. *Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, 2013, pp. 159-165, doi : 10.1109/ACII.2013.33.
- Grefenstette, E., Blunsom, P., de Freitas, N., Hermann, K. M. (2015). A Deep Architecture for Semantic Parsing. 22-27. <https://doi.org/10.3115/v1/w14-2405>

- Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S. (2010). Multi-PIE. Proceedings of the ... International Conference on Automatic Face and Gesture Recognition. IEEE International Conference on Automatic Face Gesture Recognition, 28(5), 807-813. <https://doi.org/10.1016/j.imavis.2009.08.002>
- Guo, P. J., Kim, J., Rubin, R. (2014, March). How video production affects student engagement : An empirical study of MOOC videos. In Proceedings of the first ACM Conference on Learning@Scale Conference (pp. 41-50). ACM.
- H. Gwo-Jen, "Criteria and Strategies of Ubiquitous Learning," International Conference on Sensor Networks,2006
- Hadadgar A, Changiz T, Masiello I, Dehghani Z, Mirshahzadeh N, Zary N., (2016). Applicability of the theory of planned behavior in explaining the general practitioners eLearning use in continuing medical education. BMC Med Educ. ;16(1) :215.
- Hair, J., Ringle, C. and Sarstedt, M. (2011). PLS-SEM : Indeed a Silver Bullet. Journal of Marketing Theory and Practice, 19(2), pp.139-152.
- Hajri, H. (2018). Personnalisation des MOOC par la réutilisation de Ressources Educatives Libres. Thèse de doctorat de l'université paris-saclay préparée à centralesupélec
- Hamidi, F., Ghorbandordinejad, F., Rezaee, M. and Jafari, M., 2011. A comparison of the use of educational technology in the developed/developing countries. Procedia Computer Science, 3, pp.374-377.
- Hansch, A., Hillers, L., McConachie, K., Newman, C., Schildhauer, T., Schmidt, P. (2015). The Role of Video and Online Learning : Findings From the Field and Critical Reflections. SSRN Electronic Journal. <http://www.ssrn.com/abstract=2577882>
- Harfield, A., Nang, H. (2014). A Survey of Technology Usage by Primary and Secondary School Children in Thailand. December, 12-13.
- Hassani, A. L., Mansouri, K. (2017). Towards a Model of a Scalable Middleware Architecture Based on Cloud Computing, Application Reliable Integration and E-Learning Platforms in Academic Information Systems. Transactions on Machine Learning and Artificial Intelligence, 5(4).
<https://doi.org/10.14738/tmlai.54.3329>
- Hossain, M. S., Muhammad, G. (2019). Emotion recognition using deep learning approach from audio-visual emotional big data. Information Fusion, 49(September), 69-78. <https://doi.org/10.1016/j.inffus.2018.09.008>
- Hossain, M. A., Huang, X. (2013). A cloud based m-learning architecture for higher education. Archives Des Sciences, 66(1), 751-760.
- HOX,J.J., and Bechger,T.M.(2009). An introduction to structural equation modeling family science review, 11,354-373.
- Hussain, N., Ujir, H., Hipiny, I., Minoi, J. L. (2017). 3D facial action units recognition for emotional expression. Advanced Science Letters (ICCSE2017), 24

- Hutter, F., Hoos, H. H., Leyton-Brown, K. (2011). Sequential model-based optimization for general algorithm configuration. In *Lecture Notes in Computer Science* (pp. 507-523).
- Huynh, X.-P., Tran, T.-D., Kim, Y.-G. (2016) Convolutional Neural Network Models for Facial Expression Recognition Using BU-3DFE Database. In : Kim K., Joukov N. (eds) *Information Science and Applications (ICISA) 2016. Lecture Notes in Electrical Engineering*, vol 376. Springer, Singapore. https://doi.org/10.1007/978-981-10-0557-2_44
- Hwang, G.J., Lai, CL. Wang, SY. J. (2015) Seamless flipped learning : a mobile technology-enhanced flipped classroom with effective learning strategies. *Journal of Computers in Education*. <https://doi.org/10.1007/s40692-015-0043-0>
- IBM, <https://www.ibm.com/fr-fr/cloud/learn/what-is-mobile-cloud-computing>, dernière consultation 25/05/2020
- ImageNet Large Scale Visual Recognition Competition 2010 (ILSVRC2010). Image-net.org. (2010). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2010/results>
- ImageNet Large Scale Visual Recognition Competition 2011 (ILSVRC2011). Image-net.org. (2011). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2011/results>
- ImageNet Large Scale Visual Recognition Competition 2012 (ILSVRC2012). Image-net.org. (2012). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2012/results>
- ImageNet Large Scale Visual Recognition Competition 2013 (ILSVRC2013). Image-net.org. (2013). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2013/results>
- ImageNet Large Scale Visual Recognition Competition 2014 (ILSVRC2014). Image-net.org. (2014). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2014/results>
- ImageNet Large Scale Visual Recognition Competition 2015 (ILSVRC2015). Image-net.org. (2015). Retrieved 18 April 2020, from <http://image-net.org/challenges/LSVRC/2015/results>
- ImageNet Large Scale Visual Recognition Competition 2016 (ILSVRC2016). Image-net.org. (2016). Retrieved 18 April 2020, from <http://www.image-net.org/challenges/LSVRC/2016/results>
- ImageNet Large Scale Visual Recognition Competition 2017 (ILSVRC2017). Image-net.org. (2017). Retrieved 18 April 2020, from <http://image-net.org/challenges/LSVRC/2017/results>
- J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, "ImageNet : A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, doi : 10.1109/CVPR.2009.5206848.
- J. Hundebol and N.H. Helms, " Pervasive e-learning- In situ Learning in Changing Contexts,

- "2006.
- Jarek, K., Mazurek, G. (2019). Marketing and artificial intelligence. *Central European Business Review*, 8(2), 46-55. <https://doi.org/10.18267/j.cebr.213>
 - Jason, B. (2019) pooling layers for convolutional neural networks- machine learning mastery. Retrieved 20 May 2020, from <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>
 - Jones, A. C., Scanlon, E., Clough, G. (2013). Mobile learning : Two case studies of supporting inquiry learning in informal and semiformal settings. *Computers and Education*, 61(1), 21-32. <http://doi.org/10.1016/j.compedu.2012.08.008>
 - K. Peters, "M-learning : Positioning educators for a mobile, connected future", *International Journal Of Research in Open and Distance Learning*, 2007
 - K.B. Santosh, "E-and M-Learning : A comparative study", *IJONTE*, Vol.4, pp 86. July 2013
 - K.Matthew, S.Sandra, B.Kevin and A.Peter, "Viewing mobile learning from a pedagogical perspective, *Research in Learning Technology*," 2012
 - Kaplan, A., Haenlein, M. (2019a). Siri, Siri, in my hand : Who's the fairest in the land ? On the interpretations, illustrations, and implications of artificial intelligence. In *Business Horizons*. <https://doi.org/10.1016/j.bushor.2018.08.004>
 - Kaplan, A., Haenlein, M. (2019b). Rulers of the world, unite ! The challenges and opportunities of artificial intelligence. *Business Horizons*. <https://doi.org/10.1016/j.bushor.2019.09.003>
 - Khanna, S. K. (2019). Machine Learning v / s Deep Learning. *International Research Journal of Engineering and Technology (IRJET)*, 06(02), 455-458.
 - Khokhar, R. H., 2012. Tripod of requirements in horizontal heterogeneous mobile cloud computing. *Proc. 1st International Conference on Computing, Information Systems and communications*, pp. 1-9.
 - King, D. E. (2009). Dlib-ml : A machine learning toolkit. *Journal of Machine Learning Research*, 10, 1755-1758.
 - Kock, N., 2010. What Are The Inner And Outer Models In SEM ?. [online] [Warppls.blogspot.com](http://warppls.blogspot.com). Available at : <http://warppls.blogspot.com/2010/02/what-are-inner-and-outer-models.html> [Accessed 19 january 2020].
 - Koelstra, S. (2005). *MMI Facial Expression Database - Home*. [Mmifacedb.eu](http://mmifacedb.eu). Retrieved 5 September 2019, from <https://mmifacedb.eu/>.
 - Kollias, D., Zafeiriou, S. (2019). Expression, Affect, Action Unit Recognition : Aff-Wild2, Multi-Task Learning and ArcFace., *British Machine Vision Conference (BMVC)*
 - Kumar, S., Gupta, A. (2015). Facial expression recognition : A review. In *Proceedings of the National Conference on Cloud Computing and Big Data*, Shanghai, China, 4-6 November 2015 ; pp. 159-162

- Larousse. (2020). Encyclopédie Larousse en ligne - intelligence artificielle. Larousse.fr. Retrieved 7 April 2020, from <https://www.larousse.fr/encyclopedie>
- Lasri, I. Solh, A.R., Belkacemi, M. E., (2019) Facial Emotion Recognition of Students using Convolutional Neural Network. In : 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS), Xian
- Lecun, Y Bengio, Y.. (1995). Convolutional Networks for Images, Speech, and Time-Series. *The Handbook of Brain Theory and Neural Networks*.
- Lecun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436-444). <https://doi.org/10.1038/nature14539>
- Lefevre, M., Broisin, J., Valentin, B., Daubias, P., Daubigney, L., Loria, M., Elliadd, F. G., Guin, N., Jean-daubias, S., Monod-ansaldi, Terrat, H. (2012). Personnalisation de l'apprentissage : comparaison des besoins et approches à travers l'étude de quelques dispositifs 3 . étude d 'EIAH présentant des besoins de. *Sciences et Technologies de L'Information et de La Communication Pour L'éducation et La Formation*, 19, 1-23
- Lemlouma T. 2004. "Architecture de négociation et d'adaptation de Services Multimédia dans des Environnements Hétérogènes". Thèse de doctorat, Institut National Polytechnique de Grenoble, France.
- Li, Q., Liu, Y. Q., Peng, Y. Q., Liu, C., Shi, J., Yan, F., Zhang, Q. (2021). Real-time facial emotion recognition using lightweight convolution neural network. *Journal of Physics : Conference Series*, 1827(1). <https://doi.org/10.1088/1742-6596/1827/1/012130>
- Li, W., Abtahi, F., Zhu, Z. (2017). Action unit detection with region adaptation, multi-labeling learning and optimal temporal fusing. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January*, 6766-6775. <https://doi.org/10.1109/CVPR.2017.716>
- Li, W., Abtahi, F., Zhu, Z., Yin, L. (2018). EAC-Net : Deep Nets with Enhancing and Cropping for Facial Action Unit Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(11), 2583-2596. <https://doi.org/10.1109/TPAMI.2018.2791608>
- Li, Y., Gao, W. (2017). Interconnecting heterogeneous devices in the personal mobile cloud. *Proceedings - IEEE INFOCOM*. <https://doi.org/10.1109/INFOCOM.2017.8057083>
- Liaw, Shu-Sheng Huang, Hsiu-Mei. (2011). Exploring Learners' Acceptance Toward Mobile Learning. [10.1007/978-94-6091-487-4_8](https://doi.org/10.1007/978-94-6091-487-4_8).
- Lieberman, H., Selcker, T. (2000). Out of context : Computer systems that adapt to, and learn from, context. In *IBM System Journal*. Juillet 2000.
- Linear Regression - ML Glossary documentation. ml-cheatsheet.readthedocs.io. (2017). Retrieved 18 April 2020, from https://ml-cheatsheet.readthedocs.io/en/latest/linear_regression.html.
- Lopes, A. T., de Aguiar, E., De Souza, A. F., Oliveira-Santos, T. (2016). Facial expression

- recognition with Convolutional Neural Networks : Coping with few data and the training sample order. *Pattern Recognition*, 61, 610-628. <https://doi.org/10.1016/j.patcog.2016.07.026>
- Lucey, P. ; Cohn, J.F. ; Kanade, T. ; Saragih, J. ; Ambadar, Z. ; Matthews, I. The extended Cohn-Kanade dataset (CK+) : A complete dataset for action unit and emotion-specified expression. In *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, San Francisco, CA, USA, 13-18 June 2010 ; pp. 94-101.
 - Lyons, M., Kamachi, M., Gyoba, J. (1998). The Japanese Female Facial Expression (JAFFE) Database. doi :10.5281/zenodo.3451524
 - M, H., M.N, S. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining Knowledge Management Process*, 5(2), 01-11. <https://doi.org/10.5121/ijdkp.2015.5201>
 - Mahat, J., Ayub, AFM., Luan, S., (2012). An assessment of students' mobile self-efficacy, readiness and personal innovativeness towards mobile learning in higher education in Malaysia. *Procedia Soc Behav Sci*.
 - Marie, L., Jean-Daubias, S. Nathalie, G. (2011). PERSUA2, un modèle pour unifier le processus de personnalisation des activités d'apprentissage. *Environnements Informatiques pour l'Apprentissage Humain*, Conférence EIAH'2011, Mons, Belgique.
 - Markoff, J. (2012). How Many Computers to Identify a Cat ? 16,000. *Nytimes.com*. Retrieved 18 April 2020, from <https://www.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html?pagewanted=all>.
 - Martinez, B., Valstar, M. F. (2016). Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition. *Advances in Face Detection and Facial Image Analysis*, 63-100. doi :10.1007/978-3-319-25958-1_4
 - Matsumoto, D. (1992). More evidence for the universality of a contempt expression. *Motiv Emot* 16, 363-368. <https://doi.org/10.1007/BF00992972>
 - Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P., Cohn, J. F. (2013). DISFA : A spontaneous facial action intensity database. *IEEE Transactions on Affective Computing*, 4(2), 151-160. <https://doi.org/10.1109/T-AFFC.2013.4>
 - McAuley, A., Stewart, B., Siemens, G., Cormier, D. (2010). The MOOC model for digital practice (CCattribution)
 - McDaniel, B., D'Mello, S.K., King, B., Chipman, P., Tapp, K.M., Graesser, A. (2007). Facial Features for Affective State Detection in Learning Environments.
 - McDaniel, B., D'Mello, S.K., King, B., Chipman, P., Tapp, K.M., Graesser, A. (2007). Facial Features for Affective State Detection in Learning Environments.
 - Mehmood, R., Alam, F., Albogami, N. N., Katib, I., Albeshri, A., Altowajri, S. M. (2017). UTiLearn : A Personalised Ubiquitous Teaching and Learning System for Smart Societies. *IEEE Access*, 5, 2615-2635. <https://doi.org/10.1109/ACCESS.2017.2668840>

- Mell, P. and T. Grance, The NIST definition of cloud computing. 2011.
- Meng, Z., Liu, P., Cai, J., Han, S., Tong, Y. (2017). Identity-Aware Convolutional Neural Network for Facial Expression Recognition. IEEE International Conference on Automatic Face Gesture Recognition, 558-565. <https://doi.org/10.1109/FG.2017.140>.
- Miettinen A., 2010. Energy efficiency of mobile clients in Cloud Computing. 2nd USENIX Conference on Hot Topics in Cloud Computing, 1(1), pp. 4-11.
- Miraoui, M. (2009). Architecture logicielle pour l'informatique diffuse : Modélisation du contexte et adaptation dynamique des services,cole de technologie supérieure.
- Miraoui, M., C. Tadj, and C.B. Amar. 2008. Context Modeling and ContextAware Service Adaptation for Pervasive Computing Systems.International Journal of Computer and Information Science and Engineering, 2008. 2(3) : p. 148-157.
- Mohamad Nezami O., Dras M., Hamey L., Richards D., Wan S., Paris C. (2020) Automatic Recognition of Student Engagement Using Deep Learning and Facial Expression. In : Brefeld U., Fromont E., Hotho A., Knobbe A., Maathuis M., Robardet C. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2019. Lecture Notes in Computer Science, vol 11908. Springer, Cham. https://doi.org/10.1007/978-3-030-46133-1_17
- Mollahosseini, A., Chan, D., Mahoor, M. H. (2016). Going deeper in facial expression recognition using deep neural networks. 2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016. <https://doi.org/10.1109/WACV.2016.7477450>
- Monsalve-Pulido, J., Aguilar, J., Montoya, E., Salazar, C. (2020) Autonomous Recommender System Architecture for Virtual Learning Environments, Applied Computing and Informatics
- Monerrat, B. Lavoué, E. George, S. (2015). Toward an Adaptive Gamification System for Learning Environments. [10.1007/978-3-319-25768-6_8](https://doi.org/10.1007/978-3-319-25768-6_8).
- Moser, G. (1990). In Memoriam, Arthur Samuel : Pioneer in Machine Learning. AI Magazine (AAAI), 11(3), 10-11. <https://doi.org/0738-4602>
- Moshiri, N. (2017). The Era of Online Learning. TEDxUCSD [Video]. Retrieved 20 June 2018, from <https://youtu.be/5JKgUoY9pTg?list=TLPQMDIxMDIwMjAEuzdW38P7FA>.
- Ms. J. Goyal, Ms. A.Dadhich, "Pervasive Computing," IJARCCCE, Vol. 4, Issue 12, December 2015
- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. Proceedings of the National Academy of Sciences of the United States of America, 116(44), 22071-22080. <https://doi.org/10.1073/pnas.1900654116>
- Nadira, B. (2016). CD-MOOC Context-Driven Massive Open Online Courses Amélioration du taux d'abandon des apprenants.

- Nations Unies, (1948). Déclaration universelle des droits de l'homme, Retrieved 01 septembre 2019, depuis <https://www.un.org/fr/universal-declaration-human-rights/>
- Nguyen, " Conception d'un système d'apprentissage et de travail pervasif et adaptatif fondé sur le modèle de scénario, " Thèse présentée à l'école Nationale Supérieure des Télécommunications de Bretagne, 26 mars 2010.
- Nikou, S. and Economides, A. (2016). An Outdoor Mobile-Based Assessment Activity : Measuring Students' Motivation and Acceptance. *International Journal of Interactive Mobile Technologies (iJIM)*, 10(4), p.11.
- Nonis, F., Dagnes, N., Marcolin, F., Vezzetti, E. (2019). 3D approaches and challenges in facial expression recognition algorithms-A literature review. *Applied Sciences (Switzerland)*, 9(18). <https://doi.org/10.3390/app9183904>
- Nyamen Tato, Ange Adrienne (2016). " Développement d'un système tutoriel intelligent pour l'apprentissage du raisonnement logique " Mémoire. Montréal (Québec, Canada), Université du Québec à Montréal, Maîtrise en informatique.
- O. Fezile, C. Nadire, " Basic elements and characteristics of mobile learning", WCETR,2011.
- Okai-Ugbaje, S., Ardzejewska, K., Imran, A., Yakubu, A., Yakubu, M. (2020). Cloud-Based M-Learning : A Pedagogical Tool to Manage Infrastructural Limitations and Enhance Learning. *International Journal of Education and Development Using Information and Communication Technology*, 16(2), 48.
- Omar Balghosoon, A., Jameel Qureshi, M. R. (2014). A Novel Model for Cloud Based Mobile Learning System. *International Journal of Information Engineering and Electronic Business*, 6(6), 40-46. <https://doi.org/10.5815/ijieeb.2014.06.05>
- Onah, D., Sinclair, J. (2015). Massive open online courses ? : an adaptive learning framework. *INTED2015 Proceedings*, March, 1258-1266. <https://doi.org/10.13140/RG.2.1.4237.0083>
- Oz, H. (2015). An Investigation of Preservice English Teachers' Perceptions of Mobile Assisted Language Learning. *English Language Teaching*, 8(2).
- P. Seppele, Alameki, H, "Mobile learning in teacher training", *Journal of Computer Assisted Learning*, vol. 19,2003.
- P.Dan, "Flat Army, Creating a connected and engaged organization," wiley Imprint, 2013
- Panesar, A. (2019). Machine Learning and AI for Healthcare : Big Data for Improved Health Outcomes. In *Machine Learning and AI for Healthcare : Big Data for Improved Health Outcomes*. <https://doi.org/10.1007/978-1-4842-3799-1>
- Pantic, M., Valstar, M., Rademaker, R., and Maat, L. (2005). Web-based database for facial expression analysis. *IEEE International Conference on Multimedia and Expo, Amsterdam, 2005*, pp. 5 pp.-, doi : 10.1109/ICME.2005.1521424.
- Paris-Normandie, P. (2019). Rentrée des classes 2019 en douceur pour les écoliers de Verneuil-d'Avre-et-d'Iton. www.paris-normandie.fr. Retrieved 16 April 2020,

- from <https://www.paris-normandie.fr/actualites/societe/rentree-des-classes-2019-en-douceur-pour-les-ecoliers-de-verneuil-d-avre-et-d-iton-ND15528885>
- Park, S., Nam, M. and Cha, S. (2011). University students' behavioral intention to use mobile learning : Evaluating the technology acceptance model. *British Journal of Educational Technology*, 43(4), pp.592-605.
 - Pekrun, R. (2006). The control-value theory of achievement emotions : Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341
 - Pennachin, C., Goertzel, B. (2007). Contemporary Approaches to Artificial General Intelligence 1 A Brief History of AGI. *Artificial General Intelligence*.
 - Plutchik, R. The Nature of emotions : Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am. Sci.* 2001, 89, 344-350.
 - Prabhu. (2018). Understanding of Convolutional Neural Network (CNN) - Deep Learning. Medium. Retrieved 20 May 2020, from <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>.
 - Punjabi, R. J. (2019). A Study Of Attitude Of Secondary School Teachers Towards M-Learning. *AED Journal of Educational Studies*, 8(1), 35-39.
 - Razzaque, M. A., S. Dobson and P. Nixon (2006). "Categorization and modelling of quality in context information."
 - Reich, J., Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science*, 363(6423), 130–131. <https://doi.org/10.1126/science.aav7958>
 - Rumelhart, D., Hinton, G. Williams, R. Learning representations by back-propagating errors. *Nature* 323, 533-536 (1986). <https://doi.org/10.1038/323533a0>
 - Ryan, N., Pascoe, J., Morse, D. (1997). Enhanced reality fieldwork : the contextaware archaeological assistant. In *Computer applications in archaeology*.
 - S.Thomas, "From Theory to Practice : On Designing a Pervasive Learning Game" Institute of Education, University of London,2006.
 - Salber, D., Dey, K. A., Abowd, D. G., (1998). Ubiquitous Computing : Defining an HCI Page 33 research agenda for an emerging interaction paradigm. In *Georgia Tech GVU technical report*. Janvier 1998
 - Sarah Perez, August 4, 2009, Why cloud computing is the future of mobile, <http://www.readwriteweb.com/>, Retrieved on February 2015.
 - Sarode, N., Bakal, J. W. (2020). A Review on use of Mobile Cloud System in Educational Sector. 2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020, 715-720. <https://doi.org/10.1109/ICACCS48705.2020.9074167>

- Savran, A., Alyüz, N., Dibekliolu, H., eliktutan, O., Gkberk, B., Sankur, B., Akarun, L.. (2008) Bosphorus Database for 3D Face Analysis. In : Schouten B., Juul N.C., Drygajlo A., Tistarelli M. (eds) Biometrics and Identity Management. BioID 2008. Lecture Notes in Computer Science, vol 5372. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/978-3-540-89991-4_6
- Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., Parrott, D. J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior*, 25(3), 167-185.
<https://doi.org/10.1023/A:1010671109788>
- Schilit, B., Theimer, M. (1994). Disseminating active map informations to mobile hosts. In *Institute of Electrical and Electronics Engineers (IEEE) Networks*. v. 5, p. 22-32
- Shamim, S. M. Sarker, Angona Bahar, Ali Newaz Rahman, Md. (2015). A Review on Mobile Cloud Computing. *International Journal of Computer Applications*. 113. 4-9. 10.5120/19908-1883.
- Shen, L., Wang, M., Shen, R. (2009). Affective e-Learning : Using "Emotional" data to improve learning in pervasive learning environment. *Educational Technology and Society*, 12(2), 176-189.
- Shih-hsien Yang. (2012). EXPLORING COLLEGE STUDENTS' ATTITUDES AND SELF-EFFICACY OF MOBILE LEARNING *Assoc. The Turkish Online Journal of Educational Technology* -, 11(4), 148-154.
- Shorfuzzaman, M., Hossain, M. S., Nazir, A., Muhammad, G., Alamri, A. (2018). Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment. *Computers in Human Behavior*.
doi :10.1016/j.chb.2018.07.002
- Shudong, W., Higgins, M. (2006). Limitations of mobile phone learning. *The JALT CALL Journal*, 2(1), 3-14.
- Siau, K., Yang, Y. (2017). Impact of Artificial Intelligence, Robotics, and Machine Learning on Sales and Marketing. *Twelve Annual Midwest Association for Information Systems Conference (MWAIS 2017)*, 48(June), 18-19. [https://doi.org/10.1016/S0008-6215\(00\)80957-X](https://doi.org/10.1016/S0008-6215(00)80957-X)
- Smith, k. (2020). 57 Fascinating and Incredible YouTube Statistics. *Brandwatch*. Retrieved 13 May 2020, from <https://www.brandwatch.com/blog/youtube-stats/>
- Sobaih, A. E. E., Moustafa, M. A., Ghandforoush, P., Khan, M. (2016). To use or not to use? Social media in higher education in developing countries. *Computers in Human Behavior*, 58, 296-305. <http://doi.org/10.1016/j.chb.2016.01.002>
- Soualah-alila, F. (2016). *CAMLearn : Une Architecture de Système de Thèse de Doctorat*.
- Statcounter, S. (2020). Mobile Operating System Market Share Algeria | StatCounter Global Stats. *StatCounter Global Stats*. Retrieved 11 September 2020, from <https://gs.statcounter.com/os-market-share/mobile/algeria>.

- Statista. (2020a). Smartphone users worldwide 2020 | Statista. Statista. Retrieved 8 August 2020, from <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide>
- Statista. (2020b). Algeria mobile cellular subscriptions 2000-2018 | Statista. Statista. Retrieved 8 August 2020, from <https://www.statista.com/statistics/647398/number-of-mobile-cellular-subscriptions-in-algeria/>.
- Statista. (2020b). Millennials : daily mobile access 2017 | Statista. Statista. Retrieved 8 August 2020, from <https://www.statista.com/statistics/283138/millennials-daily-mobile-usage/>.
- Steffemel, L. A. (2018). Contributions à la Gestion de l'Hétérogénéité dans les Environnements Distribués et Pervasifs.
- Stephen, D. (2013). What the 'x' in 'xMOOC' Stands for. Retrieved 13 May 2020, from <https://plus.google.com/+StephenDownes/posts/LEwaKxL2MaM>
- Sunar, A., Abbasi, R., Davis, H., White, S. and Aljohani, N., (2020). Modelling MOOC learners' social behaviours. *Computers in Human Behavior*, 107, p.105835.
- Suresh Bhosale, Y., Livingston, J. (2014). Mobile Cloud Computing for M-learning Applications. *International Journal of Engineering Research Technology (IJERT)*, 3(2), 2800-2804. www.ijert.org
- Syed, T. A., Nair, S. S. K. (2018). Personalized recommendation system for advanced learning management systems. *ACM International Conference Proceeding Series*, pp. 90-95,.
- Syed, T. A., Palade, V., Iqbal, R., Sunil, S., Nair, K. (2017). A Personalized Learning Recommendation System Architecture for Learning Management System. *Kdir*, 275-282. <https://doi.org/10.5220/0006513202750282>
- Taigman, Y., Yang, M., Ranzato, M., Wolf, L. (2014). DeepFace : Closing the gap to human-level performance in face verification. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1701-1708. <https://doi.org/10.1109/CVPR.2014.220>
- Tarnowski, P., Kolodziej, M., Majkowski, A., Rak, R. J. (2017). Emotion recognition using facial expressions. *Procedia Computer Science*, 108(June), 1175-1184. <https://doi.org/10.1016/j.procs.2017.05.025>
- Tarus, J. K., Niu, Z., Kalui, D. (2018). A hybrid recommender system for e-learning based on context awareness and sequential pattern mining. *Soft Computing*, 22(8), 2449-2461. <https://doi.org/10.1007/s00500-017-2720-6>
- Tawalbeh, Loai Alassaf, Norah Bakhader, Waseem Tawalbeh, Alaa.(2015). Resilience Mobile Cloud Computing : Features, Applications and Challenges. *10.1109/ECONF.2015.59*.
- Tenenhaus, M., (2008). Component-based Structural Equation Modelling. *Total Quality Management Business Excellence*, 19(7-8), pp.871-886.

- Turing, A.M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460
- U.Lucke, Rensing, "A survey on pervasive education. *Pervasive Mobile Computer*". 2014
- UNESCO (2002). Forum on the impact of Open Courseware for higher education in developing countries final report. Retrieved from <http://unesdoc.unesco.org/images/0012/001285/128515e.pdf>
- UNESCO, U. (2017). Open Educational Resources. UNESCO. Retrieved 6 April 2020, from <https://en.unesco.org/themes/ict-education/oer>
- University of London. History of the University of London. University of London. Retrieved 16 April 2020, from <https://london.ac.uk/about-us/history-university-london>
- V. Jones Jo, J.H. "Ubiquitous learning environment : An adaptive teaching", *ASCILITE*, VOL.21, december 2004
- Valstar, M. (2008). Timing is everything A spatio-temporal approach to the analysis of facial actions Imperial College of Science , Technology and Medicine Department of Computing Timing is everything A spatio-temporal approach to the analysis of facial actions Michel Fran , (Issue February). Imperial College of Science, Technology and Medicine.
- Venkatesh, Viswanath ; Morris, Michael G. ; Davis, Gordon B. ; Davis, Fred D. (2003). "User Acceptance of Information Technology : Toward a Unified View". *MIS Quarterly*. 27 (3) : 425-478. JSTOR
- Vinay, S., Vaseekharan, M., Mohamedally, D. (2013). RoboRun : a gamification approach to control flow learning for young students with TouchDevelop, <https://arxiv.org/abs/1310.0810>
- Vincent , C., Delozanne E., Grugeon, B., Gélis , J.-M., Rogalski, J., and Cou- lange, L. (2005) Des erreurs aux stéréotypes : Des modèles cognitifs de différents niveaux dans le projet pépite, in *Environnements Informatiques pour l'Apprentissage Humain*. NAT LIP6 MOCAH
- W. Mark, "the computer for the 21 century," *Scientific American*, Vol. 265, No. 3, pp. 94-104, September 1991
- Waard, I. d. (2013). Analyzing the Impact of Mobile Access on Learner Interactions in a MOOC. Athabasca : Unpublished Masters Thesis. Athabasca University.
- Wallach, I., Dzamba, M., Heifets, A. (2015). AtomNet : A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery. 1-11. <http://arxiv.org/abs/1510.02855>
- Weiser.M, Some computer science issues in ubiquitous computing. *Communications of the ACM*, 1993, pp75-84.
- West, T. and Paine, C. (2010). The opportunities of mobile learning for executive education. *The Ashridge Journal*, 8(1), pp.1-7
- White, M. (2017). A History of Intranets. <https://doi.org/10.4324/9780203102565>

- Wu, Y., Turner, P. (2020). The Relationship of Bandwidth, Interaction and Performance in Online Courses : A Study. Westga.edu. Retrieved 19 August 2020, from <https://www.westga.edu/distance/ojdl/spring91/wu91.htm>.
- Julie. Y, I.C, "Effects of Interactive Multimedia in E-Learning On Learners and Developers", 2012
- Saadiah, Y, Arniza. E, A, d and Kamarularifin. A, "The definition and characteristics of ubiquitous learning : A discussion" MARA, Malaysia, (IJEDICT), Vol. 6, Issue 1, pp. 117-127, 2010.
- Y.Tzu-Chi Y, F.Ray Kuo, G. Jen Hwang and H.Chun Chu, "A Computer-Assisted Approach for Designing Context-Aware Ubiquitous Learning Activities", 2008
- Yan, F.; Iliyasu, A.; Liu, Z.; Abuhasel, K.; Jiang, Z.; Dong, F.; Hirota, K. (2014). Representation of quantum emotion space using Bloch sphere. Proceedings of the Joint International Conference of ITCA and ISCIIA 2014, Hunan, China, 16-21 September 2014; pp. 1-6.
- Yin, D. B. M., Omar, S., Talip, B. A., Muklas, A., Norain, N. A. M., Othman, A. T. (2017). Fusion of face recognition and facial expression detection for authentication. Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication - IMCOM '17. <https://doi.org/10.1145/3022227.3022247>
- Yin, L.; Wei, X.; Sun, Y.; Wang, J.; Rosato, M.J. A 3D facial expression database for facial behavior research. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06), Southampton, UK, 10-12 April 2006; pp. 211-216
- Yousafzai, A., Chang, V., Gani, A., Noor, R. M. (2016). Multimedia augmented m-learning : Issues, trends and open challenges. International Journal of Information Management, 36(5), 784-792. <https://doi.org/10.1016/j.ijinfomgt.2016.05.010>
- Zatarain Cabada, R., Rodriguez Rangel, H., Barron Estrada, M. L., Cardenas Lopez, H. M. (2020). Hyperparameter optimization in CNN for learning-centered emotion recognition for intelligent tutoring systems. Soft Computing, 24(10), 7593-7602. <https://doi.org/10.1007/s00500-019-04387-4>
- Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., Dobaie, A. M. (2017). Facial expression recognition via learning deep sparse autoencoders. Neurocomputing, 273, 643-649. <https://doi.org/10.1016/j.neucom.2017.08.043>
- Zhang, B., Yin, C., David, B., Xiong, Z., Niu, W. (2016). Facilitating professionals' work-based learning with context-aware mobile system. Science of Computer Programming, 129, 3-19. <https://doi.org/10.1016/j.scico.2016.01.008>
- Zhang, G., Kato, J., Wang, Y., Mase, K. (2016). How to initialize the CNN for small datasets : Extracting discriminative filters from pre-trained model. Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015, 2, 479-483. <https://doi.org/10.1109/ACPR.2015.7486549>
- Zhang, S., Pan, X., Cui, Y., Zhao, X., Liu, L. (2019). Learning Affective Video Features for Facial Expression Recognition via Hybrid Deep Learning. IEEE Access, 7, 32297-32304.

<https://doi.org/10.1109/ACCESS.2019.2901521>

- Zhang, X., Yin, L., Cohn, J. F., Canavan, S., Reale, M., Horowitz, A., Liu, P., Girard, J. M. (2014). BP4D-Spontaneous : A high-resolution spontaneous 3D dynamic facial expression database. *Image and Vision Computing*, 32(10), 692-706.
<https://doi.org/10.1016/j.imavis.2014.06.002>
 - Zhao, X., Chen, G., Chuang, Y., Tao, X., Zhang, S. (2020). Learning Expression Features via Deep Residual Attention Networks for Facial Expression Recognition From Video Sequences. *IETE Technical Review*, 1-9. doi :10.1080/02564602.2020.1814168
 - Zhu, W., Zeng, N., Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations. *Northeast SAS Users Group 2010 : Health Care and Life Sciences*, 1-9
 - Zhou, Y., Shi, B. E. (2017). Action Unit Selective Feature Maps in Deep Networks for Facial Expression Recognition. 2031–2038.
-