République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université A/Mira de Béjaia
Faculté des Sciences Exactes
Département d'Informatique

MÉMOIRE DE MASTER RECHERCHE

En

Informatique

Option

Réseau et Sécurité

Thème

Authentification Biométrique par mouvements oculaires

Présenté par : M. NEMEUR Chems Eddine et M. AMEUR Zineddine

Soutenu le 22 Juin 2023 devant le jury composé de :

Président	**ACHROUFENE Achour**	**U. A/Mira Béjaia.**
Examinatrice	**KESSIRA Dalila**	**U. A/Mira Béjaia.**
Encadrante	**ZEBBOUDJ Sofia**	**U. A/Mira Béjaia.**
Invité	**AKILAL Abdellah**	**U. A/Mira Béjaia.**

Béjaia, Juin 2023.

⁕ *Remerciements* ⁕

Nous remercions Dieu le Tout-Puissant pour Sa guidance et Sa miséricorde qui nous ont permis d'accomplir ce travail.

Nous remercions tous ceux qui ont contribué à notre travail, en commençant par nos encadrants Dr ZEBBOUDJ Sofia et Dr AKILAL Abdellah, pour leurs efforts et leur accompagnement. Nous exprimons également notre gratitude envers les participants qui ont pris part à notre processus de collecte de données, rendant ainsi notre travail possible.

⁂ *Dédicaces* ⁂

Je suis profondément reconnaissant envers ma mère pour son amour inconditionnel, son soutien indéfectible et les nombreux sacrifices qu'elle a consentis pour moi. Ses précieux conseils et sa présence constante dans ma vie ont été essentiels à ma réussite. À travers ce travail humble, je souhaite exprimer mes sentiments et ma gratitude éternelle envers elle. Que Dieu la protège et la bénisse toujours.

Je suis profondément reconnaissant envers mon père pour les nombreuses années de sacrifices qu'il a consenties. Son enseignement des valeurs nobles, son dévouement à mon éducation et son soutien permanent ont été inestimables. Que Dieu le protège et le préserve toujours.

Je suis infiniment reconnaissant envers mes sœurs pour leur amour inconditionnel et leur soutien moral sans faille, particulièrement lors des moments les plus difficiles.

M. ***AMEUR Zineddine***

⁂ *Dédicaces* ⁂

À mes chers parents qui ne sont plus parmi nous dans ce monde, ce travail est dédié à leur souvenir. Je suis éternellement reconnaissant de leur amour, et leur présence, même si elle est spirituelle, continue de me guider vers le succès.
À mes deux frères, ma sœur et mon oncle pour leur encouragement constant et leurs précieux conseils.Je suis chanceux d'avoir pu compter sur leur soutien tout au long de mon parcours académique.

M. ***Chems Eddine Nemeur***

Table des matières

Table des matières **i**

Table des figures **iii**

Liste des tableaux **v**

Introduction générale **1**

1 Apprentissage automatique **2**
1.1 Introduction . . . 2
1.2 Apprentissage Automatique . . . 2
1.3 Types d'apprentissage Automatique . . . 3
1.3.1 Apprentissage Automatique supervisé . . . 3
1.3.2 Apprentissage Automatique non supervisé . . . 6
1.3.3 Apprentissage Automatique par Renforcement . . . 7
1.3.4 Apprentissage Automatique basé sur les méthodes d'ensemble . . . 7
1.4 Apprentissage profond . . . 9
1.4.1 Réseau de neurones convolutif (CNN – Convolutional Neural Network) . . . 9
1.4.2 Réseaux de neurones récurrents (RNN – Recurrent Neural Network) . . . 11
1.5 Conclusion . . . 12

2 Etat de l'art sur l'authentification biométrique basée sur les mouvements oculaires **13**
2.1 Introduction . . . 13
2.2 Historique sur l'authentification . . . 13
2.3 Protocoles d'authentification basés sur les mouvements oculaires . . . 14
2.3.1 Authentification à un facteur . . . 14
2.3.2 Authentification à deux facteurs . . . 19
2.4 Discussion des résultats . . . 24
2.5 Conclusion . . . 25

3 Proposition d'un nouveau jeu de données pour l'authentification biométrique 26
3.1 Introduction . . . 26
3.2 Notions préliminaires . . . 26
3.2.1 Résolution d'image . . . 26
3.2.2 La Définition d'une image . . . 27
3.2.3 Ouverture de la caméra . . . 27
3.2.4 Frame Rate (FPS) . . . 27
3.3 Quelques jeux de données sur les mouvements oculaires . . . 27
3.3.1 Jeux de données basés sur les signaux . . . 28
3.3.2 Jeux de données basés sur les vidéos . . . 30
3.3.3 Jeux de données basés sur les images . . . 32
3.3.4 Discussion . . . 32
3.4 Présentation de EyeReg . . . 33
3.4.1 Participants . . . 33
3.4.2 Matériels utilisés . . . 34
3.4.3 Protocole de collecte du jeu de données vidéos . . . 34
3.4.4 Propriétés de EyeReg version vidéo . . . 35
3.4.5 EyeReg version image . . . 35
3.4.6 Comparaison . . . 37
3.5 Conclusion . . . 38

4 Expérimentation et résultats 39
4.1 Introduction . . . 39
4.2 Environnement expérimentale . . . 39
4.2.1 Matériel . . . 39
4.2.2 Outils et bibliothèques . . . 40
4.3 Résultats obtenus avec notre jeu de données sur le protocole de Goussem et Djallil . 41
4.4 Expérimentations réalisées . . . 43
4.4.1 L'effet de nombre de caractéristiques sur l'authentification . . . 43
4.4.2 Corrélation dans notre jeu de données . . . 46
4.4.3 Augmentation du nombre des points de repère oculaires . . . 47
4.4.4 Proposition de nouvelles caractéristiques pour l'authentification . . . 48
4.4.5 Entrainement avec les réseaux de neurones convolutifs . . . 49
4.5 Conclusion . . . 52

Conclusion et perspectives 53

Bibliographie 55

Table des figures

1.1 Structure générale des machines à Vecteurs de Support. 4
1.2 Structure du modèle KNN. 5
1.3 Illustration du modèle d'un arbre de décision. 6
1.4 Fonctionnement de l'algorithme K-means pour le clustering de données. 7
1.5 Illustration du modèle l'apprentissage par renforcement. 7
1.6 Cadre général de l'ensemble parallèle Bagging. 8
1.7 Cadre général du Boosting. 8
1.8 Cadre général de Stacking. 9
1.9 Architecture d'un CNN 10
1.10 La fonction d'activation ReLU. 11
1.11 Diagramme de RNN [66]. 12

2.1 L'architecture Eye Know You Too (EKYT) [50]. 15
2.2 Processus de système d'identification des utilisateurs [71]. 17
2.3 Architecture de système Blinkey [41]. 20
2.4 Architecture du modèle de Goussem et Djallil [31]. 22
2.5 Phase d'identification du modèle de Goussem et Djallil [31]. 22
2.6 Phase d'inscription du modèle de Goussem et Djallil [31] 23
2.7 Phase d'authentification du modèle de Goussem et Djallil [31]. 23

3.1 Classification des jeux de données étudiés. 28
3.2 Matériels et procédure de collection GazeBase [37]. 29
3.3 L'environnemt virtuel de Sarker et al [71]. 29
3.4 Casque de réalité virtuelle HTC Vive Pro Eye [5]. 30
3.5 Formulaire de consentement. 34
3.6 Protocole de collection. 34
3.7 Aperçus de notre jeu de données vidéos. 35
3.8 Aperçus du premier sous-ensemble de notre jeu de données images. 36
3.9 Aperçus du deuxième sous-ensemble de notre jeu de données images. 37

4.1 Points de repère oculaires. 47
4.2 Les caractéristiques Proposées. 48

4.3 Modèle CNN pour le jeux de données image des deux yeux. 49
4.4 Modèle CNN pour le jeux de données image des yeux séparés. 50

Liste des tableaux

2.1 Tableau comparatif des protocoles d'authentification biométrique basés sur le mouvement oculaire de notre étude. 25

3.1 Tableau comparatif. 37

4.1 Caractéristiques de l'ordinateur portable 40
4.2 Résultats de l'authentification avec le protocole de Goussem et al. [31] pour un seuil de 4. 42
4.3 Résultats de l'authentification avec le protocole de Goussem et al. [31] un seuil de 4.01. 42
4.4 Résultats de l'authentification avec le protocole de Goussem et al. [31] un seuil de 4.1. 42
4.5 Résultats de l'authentification avec 2 caractéristiques et un seuil de 0.038. 44
4.6 Résultats de l'authentification avec 3 caractéristiques et un seuil de 0.99. 44
4.7 Résultats de l'authentification avec 4 caractéristiques et un seuil de 1.1. 45
4.8 Résultats de l'authentification avec 5 caractéristiques et un seuil de 2.7. 45
4.9 Coefficients de corrélation de chaque Caractéristique. 46
4.10 Résultats de l'authentification avec 6 points de repère oculaires. 47
4.11 Résultats de l'authentification avec neuf caractéristiques 49

Liste des abréviations

CNN Convolutional Neural Network

CoV Coefficient of Variation

DL Deep Learning

DT Decision Tree

EKYT Eye Know You Too

EMB Eye Movement Biometrics

FAR False Acceptance Rate

FFT Fast Fourier Transform

FHD Full High Definition

FPS Frames Per Second

FRR False Rejection Rate

HD High Definition

HTC High Tech Computer

IPS Images Par Seconde

KNN K-Nearest Neighbors

LDA Linear Discriminant Analysis

LSTM Long short-term memory

ML Machine Learning

MP Megapixel

NPC Neighboring Pixel Communication

RBF Radial Basis Function

RF Random forest

RFE Recursive Feature Elimination

ReLU Rectified Linear Unit

RNN Recurrent Neural Networks

SVM Support Vector Machine

TEE Taux d'Erreur d'Égalité

TFP True Positive Rate

VR Virtual Reality

VS Code Visual Studio Code

ZJU Zhejiang University

2FA Two-Factor Authentication

3D Three-Dimensional

3FA Three-Factor Authentication

Introduction générale

L'authentification par la biométrie est une approche qui a attiré beaucoup d'attention dans le domaine de la sécurité. Cette combinaison de la biométrie humaine et de la sécurité représente une évolution des méthodes d'authentification traditionnelles telles que les mots de passe. En utilisant les caractéristiques physiologiques ou comportementales uniques des individus, telles que les empreintes digitales, les mouvements des yeux et la reconnaissance vocale, l'authentification biométrique offre une approche robuste et pratique pour vérifier l'identité des individus.

Dans ce mémoire de fin d'étude, nous étudions les systèmes d'authentification biométriques en utilisant les caractéristiques liées aux yeux. L'utilisation des caractéristiques oculaires semblent être un bon choix pour l'authentification des individus dans plusieurs domaines, et plus particulièrement, dans le domaine de la réalité virtuelle où, grâce à un casque de réalité virtuelle, l'utilisateur peut entrer dans un monde artificiel créé par ordinateur, offrant une expérience réaliste et interactive.

Nos contributions dans ce travail sont multiples. Pour les présenter, nous avons divisé le mémoire en 04 chapitres : dans le chapitre 01 nous étudierons les méthodes les plus utilisées dans l'apprentissage automatique et l'apprentissage profond. Le chapitre 02 est un état de l'art sur différents protocoles d'authentification biométrique, en mettant l'accent sur le protocole de Goussem et Djallil [31], qui servira de base pour nos contributions. Dans le chapitre 03 nous présenterons d'abord les jeux de données existants utilisés par les protocoles d'authentification étudiés précédemment, puis nous introduirons notre propre jeu de données pour l'authentification biométrique oculaire et la détection de clignements. Le chapitre 04 exposera une série d'expérimentations réalisées sur le protocole d'authentification de Goussem et Djallil [31], ainsi que les améliorations apportées sur la base des résultats obtenus. Enfin, ce mémoire se conclura par une conclusion générale et des perspectives d'avenir pour notre travail.

Chapitre 1

Apprentissage automatique

1.1 Introduction

L'apprentissage automatique est un domaine de recherche passionnant en informatique. Son objectif principal est de permettre à la machine de penser comme un être humain, c'est-à-dire d'apprendre, de prendre des décisions, etc.

Dans ce chapitre, nous abordons différents types d'apprentissage automatique. Nous commençons par l'apprentissage supervisé, et présenterons trois classificateurs de ce type : les Machines à Vecteurs de Support (SVM), les k-plus proches voisins (KNN) et les arbres de décision (DT). Nous parlerons ensuite de l'apprentissage non supervisé, où les instances ne sont pas étiquetées. Nous étudions pour ce type, l'algorithme de k-means clustering. Nous présenterons aussi l'apprentissage automatique par renforcement et les méthodes d'ensemble. Nous explorons ensuite l'apprentissage profond, un sous domaine de l'apprentissage automatique, en présentant les réseaux de neurones convolutifs et les réseaux de neurones récurrents.

1.2 Apprentissage Automatique

L'apprentissage automatique (Machine Learning) est un sous-domaine de l'intelligence artificielle qui consiste au développement d'algorithmes et de modèles permettant aux systèmes informatiques d'apprendre à partir de données, de faire des prédictions ou de classifier des nouvelles données. Il s'agit d'un domaine en pleine croissance de nombreuses d'applications courantes, telles que le traitement du langage naturel, la reconnaissance d'images, les systèmes de recommandation et les systèmes de sécurité.

L'apprentissage automatique s'inspire de diverses domaines, notamment l'informatique, les statistiques, la biologie et la psychologie. Son objectif principal est d'apprendre aux ordinateurs à trouver automatiquement un modèle basé sur des expériences passées. Cette tâche est réalisée par un classifieur [42], c'est-à-dire un algorithme qui crée un modèle à partir des données observées. Ce modèle est ensuite utilisé pour prédire des résultats ou classer de nouvelles données.

1.3 Types d'apprentissage Automatique

L'élément clé de l'apprentissage automatique est les données. Ces données sont représentées par des ensembles appelés "jeux de données".

Chaque instance dans un jeux de données est représentée en utilisant le même ensemble de caractéristiques (features). Les caractéristiques peuvent être continues, catégoriques ou binaires. Si les instances sont données avec des étiquettes connues (les sorties correctes correspondantes), alors l'apprentissage est appelé supervisé. Sinon, quand les instances ne sont pas étiquetées, l'apprentissage est appelé non supervisé [23].

L'apprentissage peut également être réalisé par renforcement, où l'algorithme s'améliore en interagissant avec son environnement ou en utilisant des méthodes d'ensembles qui visent à améliorer les performances d'un ensemble d'algorithmes afin d'optimiser leurs résultats.

1.3.1 Apprentissage Automatique supervisé

L'apprentissage supervisé est un type d'apprentissage automatique dans lequel l'algorithme est entraîné sur des données étiquetées. Ceci signifie que les données d'entrée ont une sortie (ou une étiquette) correspondante que l'algorithme cherche à prédire. Dans cette catégorie, nous présentons les trois classifieurs les plus utilisés : les machines à vecteurs de support (SVM), la méthode des k plus proches voisins et les arbres de décision.

1.3.1.1 Machines à vecteurs de support (SVM – Support Vector Machine)

SVM est une méthode de classification binaire par apprentissage supervisé. Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprié qui vise à trouver la meilleure frontière possible pour séparer différents classes dans un ensemble de données, tel qu'illustré par la Figure 1.1.

La marge représente l'espace entre la frontière de décision et les points les plus proches de chaque classe. Augmenter la marge peut réduire les erreurs de classification, tandis qu'une marge plus réduite peut entraîner des classifications incorrectes. Ce compromis est contrôlé par le paramètre C de l'algorithme. Une valeur plus petite de C conduit à une marge plus large, ce qui permet une meilleure séparation entre les classes et une meilleure généralisation de la frontière de décision, réduisant ainsi les erreurs de classification. En revanche, une valeur plus grande de C peut entraîner une marge plus étroite, ce qui réduit les erreurs de classification sur l'ensemble d'entraînement, mais pourrait ne pas bien généraliser aux données inconnues, conduisant à des classifications incorrectes. Il est donc essentiel de trouver un bon équilibre en ajustant le paramètre C pour obtenir une marge appropriée et minimiser les erreurs de classification.

Les vecteurs de support sont les points de données qui se trouvent sur la marge. L'algorithme se base sur ces vecteurs de support pour déterminer la frontière de décision.

Les SVM peuvent traiter les données qui peuvent être séparées par une ligne droite en utilisant un noyau linéaire. Cependant, il peut également fonctionner avec des données qui ne sont pas séparables linéairement en appliquant des fonctions de noyau. Ces fonctions transforment les données en espaces de dimensions supérieures, où une séparation par une frontière linéaire devient possible [57].

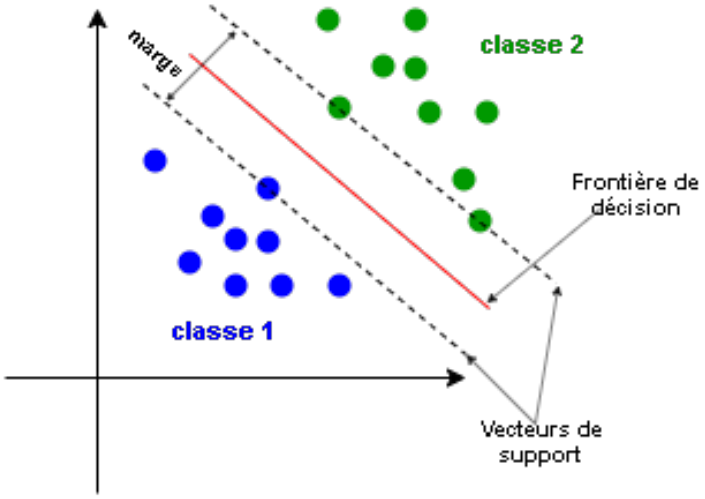


FIGURE 1.1 – Structure générale des machines à Vecteurs de Support.

La Figure 1.1 illustre la structure générale des SVM. La frontière de décision, représentée par la ligne rouge est positionnée de manière centrée entre les vecteurs de support indiqués par les lignes continues. Cela permet de séparer efficacement les deux classes 1 et 2.

1.3.1.2 k plus proche voisins (KNN - K-Nearest Neighbors)

KNN est un algorithme d'apprentissage automatique utilisé pour la régression et la classification. Il analyse les étiquettes des points de données voisins d'un point cible afin de prédire sa classe en se basant sur la distance entre ce point et ses voisins les plus proches, tel qu'illustré par la Figure 1.2. Le paramètre k, qui correspond au nombre de voisins à prendre en compte, est le cœur de cette technique. KNN peut être utilisé avec différents types de données, tels que des données numériques ou catégorielles [59].

Les mesures de distance utilisées dans KNN sont souvent la distance euclidienne et la distance de Manhattan. La distance euclidienne calcule la distance entre deux points a et b dans un espace multidimensionnel de dimension n en utilisant la formule mathématique de la distance euclidienne représenté l'équation (1.1) [21].

$$d_E(a,b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \tag{1.1}$$

La distance de Manhattan, quant à elle, calcule la distance en effectuant la somme des différences absolues entre les coordonnées des points a et b dans chaque dimension (1.2) [21].

$$d_M(a,b) = \sum_{i=1}^{n}|a_i - b_i| \tag{1.2}$$

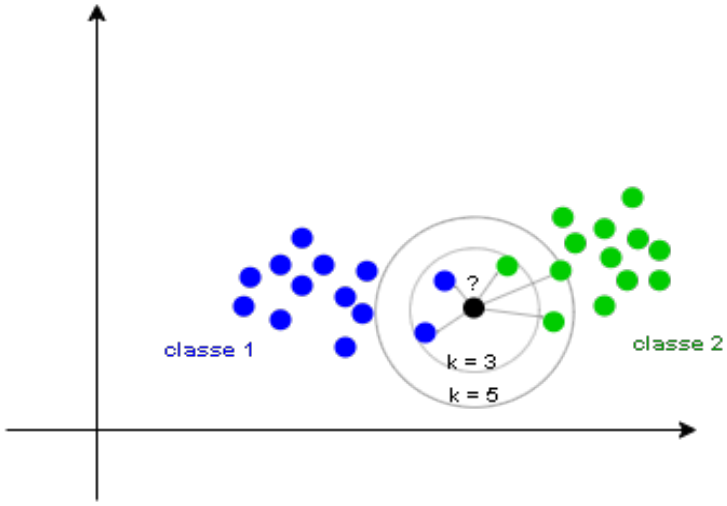


FIGURE 1.2 – Structure du modèle KNN.

La Figure 1.2 illustre la structure générale de l'algorithme KNN. Dans ce cas, la classification du point noir se fait en calculant la distance entre ce point et les points voisins. La décision de classification est prise en fonction du nombre de voisins pris en compte. Pour k égal à 3, le point est classé dans la classe 1, tandis que pour k égal à 5, le point est classé dans la classe 2 .

1.3.1.3 Arbres de décision

Un arbre de décision est un modèle d'apprentissage automatique qui fonctionne selon le principe de partition récursive de l'espace d'entrée et qui définit un modèle local dans chaque région résultante de l'espace d'entrée [32]. Cela peut être représenté par un arbre, avec une feuille par région tel qu'illustré par la Figure 1.3 .

L'arbre est construit en partitionnant de manière récursive les données en fonction des valeurs des différentes caractéristiques. À chaque nœud, l'algorithme sélectionne la meilleure caractéristique qui offre le plus de gain d'information ou la meilleure division. Ce processus se poursuit jusqu'à ce qu'une condition d'arrêt soit atteinte, telle que l'atteinte d'une profondeur maximale ou lorsque la division ultérieure de l'espace d'entrée ne fournit pas d'améliorations significatives.

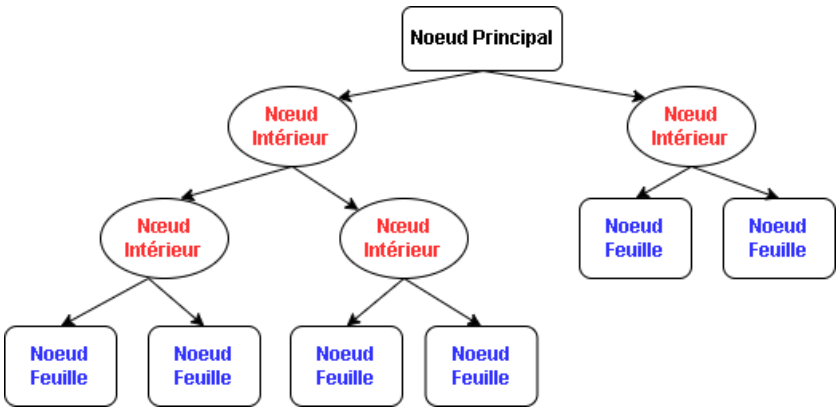


FIGURE 1.3 – Illustration du modèle d'un arbre de décision.

1.3.2 Apprentissage Automatique non supervisé

L'apprentissage non supervisé vise à extraire les schémas ou les relations significatifs à partir de données non étiquetées. Contrairement à l'apprentissage supervisé où les données étiquetées sont utilisées pour entraîner un modèle, l'apprentissage non supervisé exploite les relations entre les données elles-mêmes. Les algorithmes d'apprentissage non supervisé regroupent les points de données similaires en utilisant différentes techniques telles que le clustering. Ainsi, ces données similaires forment un cluster, qui est considéré comme une nouvelle classe combinant les données. Cette approche est utile lorsque les données étiquetées sont absentes ou indisponibles, ou lorsque nous souhaitons découvrir de nouvelles relations entre les données. Dans cette catégorie, nous présenterons l'algorithme d'apprentissage non supervisé le plus populaire : le K-means clustering.

1.3.2.1 K-means clustering

K-Means clustering est l'un des algorithmes d'apprentissage non supervisés les plus simples qui résolvent le problème bien connu du regroupement (clustering). L'idée principale est de définir k centroïdes , un pour chaque cluster. Ces centroïdes doivent être placés de manière astucieuse car leur emplacement influence le résultat. Ainsi, le meilleur choix est de les placer le plus loin possible les uns des autres, tel qu'illustré par la Figure 1.4.

L'étape suivante consiste à prendre chaque point appartenant à un ensemble de données donné et à l'associer au centroïde le plus proche. Lorsqu'aucun point est libre, la première étape est terminée et un premier regroupement est effectué. À ce stade, il est nécessaire de recalculer k nouveaux centroïdes en tant que centres des clusters résultant de l'étape précédente. Après ces k nouveaux centroïdes, une nouvelle association doit être effectuée entre les mêmes points de données et le nouveau centroïde le plus proche. Une boucle est ainsi générée. À la suite de cette boucle, on peut observer que les k centroïdes changent progressivement de position jusqu'à ce qu'aucun autre changement ne soit effectué. En d'autres termes, les centroïdes ne se déplacent plus [45].

La Figure 1.4 illustre le fonctionnement l'algorithme K-means pour le clustering de données.

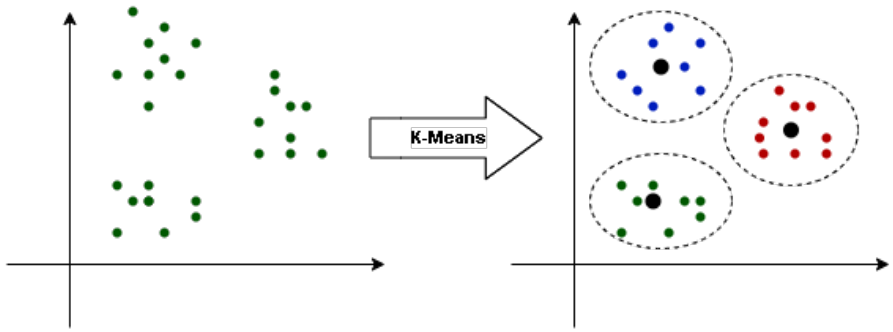


FIGURE 1.4 – Fonctionnement de l'algorithme K-means pour le clustering de données.

1.3.3 Apprentissage Automatique par Renforcement

L'apprentissage par renforcement ou Reinforcement Learning est une méthode d'apprentissage Automatique qui consiste à entraîner des modèles d'intelligence artificielle d'une manière bien spécifique. Un algorithme apprend à prendre des décisions en interagissant avec son environnement et en recevant des rétroactions sur ses actions tel qu'illustré par la Figure 1.5. Ce processus itératif de prise de décision et d'apprentissage est couramment appliqué dans des domaines tels que la robotique, où l'agent apprend à accomplir des tâches complexes, ainsi que dans les jeux, où il peut s'améliorer en jouant contre lui-même ou contre d'autres joueurs virtuels. Cette approche permet de développer des systèmes autonomes capables d'adaptation et d'apprentissage continu en fonction des résultats de leurs interactions avec l'environnement [10].

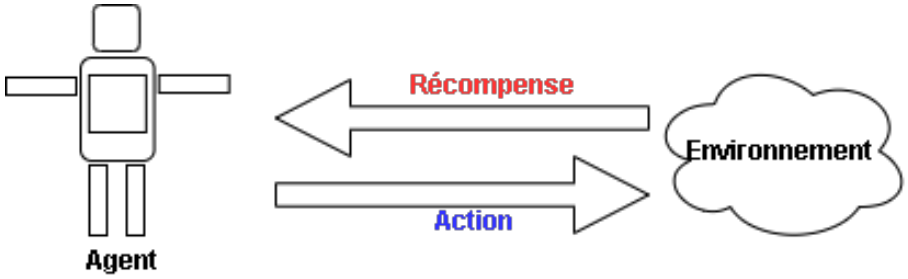


FIGURE 1.5 – Illustration du modèle l'apprentissage par renforcement.

La structure générale de l'apprentissage automatique par renforcement est représentée dans la Figure 1.5 où un agent apprend comment choisir une action dans son espace d'action, dans un environnement particulier, afin de maximiser les récompenses au fil du temps.

1.3.4 Apprentissage Automatique basé sur les méthodes d'ensemble

Les méthodes ensemblistes sont des méthodes qui consistent à combiner plusieurs modèles afin de pouvoir générer un nouveau modèle qui soit potentiellement plus performant mais aussi

plus robuste. Cette combinaison peut s'effectuer généralement de trois façons [16] : avec le Bagging (l'ensemble parallèle), le Boosting (l'ensemble séquentiel) et le stacking (la méthode d'empilement).

1.3.4.1 Bagging

Les méthodes reposant sur le bagging (ou l'ensemble parallèle), telles que les forêts aléatoires [52], consistent à combiner des modèles qui sont appris à partir d'informations différentes et à partir d'échantillons différents. Ces modèles sont ensuite combinés, généralement sommés, afin de créer un seul et unique modèle performant [16] tel qu'illustré sur la Figure 1.6.

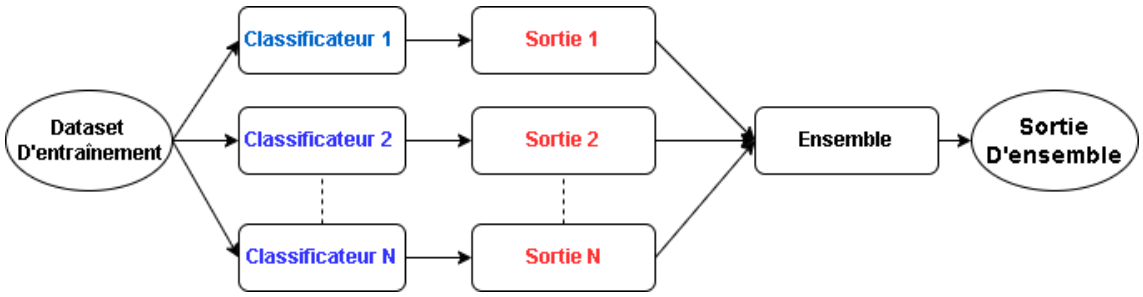


FIGURE 1.6 – Cadre général de l'ensemble parallèle Bagging.

1.3.4.2 Boosting

La méthode de Boosting consiste à combiner plusieurs modèles faibles pour obtenir un prédicteur plus puissant. Ces modèles sont entraînés de manière séquentielle ou un modèle se concentre sur les erreurs faites par le modèle précédent. Cette approche itérative permet d'améliorer progressivement les modèles faibles, ce qui conduit à de meilleures performances globales, tel qu'illustré sur la Figure 1.7 [6]. Le Boosting est une méthode couramment utilisée pour améliorer la précision des algorithmes d'apprentissage automatique dans les tâches de la classification et la régression.

La Figure 1.7 représente le fonctionnement du Boosting.

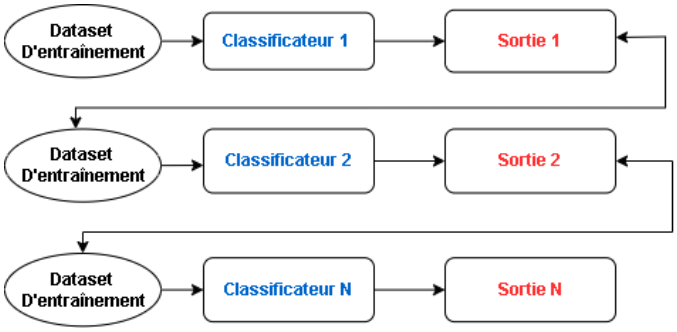


FIGURE 1.7 – Cadre général du Boosting.

1.3.4.3 Stacking

La méthode d'empilement, également appelée Stacked Generalization, est une technique d'ensemble de modèles utilisée pour fusionner les informations de plusieurs modèles prédictifs afin de créer un nouveau modèle (méta-modèle). L'architecture d'un modèle d'empilement comprend deux modèles de base ou plus, appelés modèles de niveau 0, ainsi qu'un méta-modèle qui combine les prédictions des modèles de base, appelé modèle de niveau 1. Les modèles de niveau 0 sont ajustés sur les données d'entraînement et leurs prédictions sont ensuite compilées. Le modèle de niveau 1 tel qu'illustré sur la Figure 1.8, quant à lui, apprend la meilleure façon de combiner les prédictions des modèles de base. Les sorties des modèles de base, utilisées comme entrée pour le méta-modèle, peuvent être des valeurs de probabilité ou des étiquettes de classe dans le cas de la classification [6]. La Figure 1.8 illustre le processus du stacking.

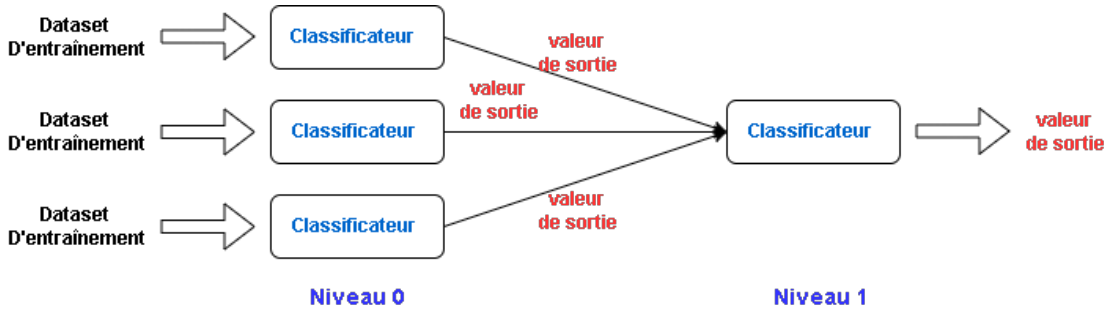


FIGURE 1.8 – Cadre général de Stacking.

1.4 Apprentissage profond

Le deep learning, ou l'apprentissage profond, est un sous-domaine de l'apprentissage automatique qui utilise des réseaux neuronaux profonds pour résoudre des problèmes complexes. Il comprend divers types de réseaux, tels que les réseaux de neurones convolutifs (CNNs) et les réseaux neuronaux récurrents (RNNs). Ces modèles de deep learning sont capables d'apprendre directement à partir des données d'origine, telles que des images et des textes, sans nécessité d'une étape préalable d'ingénierie de caractéristiques manuelle [38].

1.4.1 Réseau de neurones convolutif (CNN – Convolutional Neural Network)

Un réseau de neurones convolutif (CNN), est un type de réseau neuronal profond largement utilisé dans la reconnaissance d'images. Les CNN sont capables d'apprendre à identifier des motifs dans des images en utilisant un processus appelé convolution. La convolution est une opération mathématique qui prend deux fonctions en entrée et produit une troisième fonction en sortie. Dans

le contexte des CNN, la convolution est utilisée pour extraire des caractéristiques à partir d'images [30].

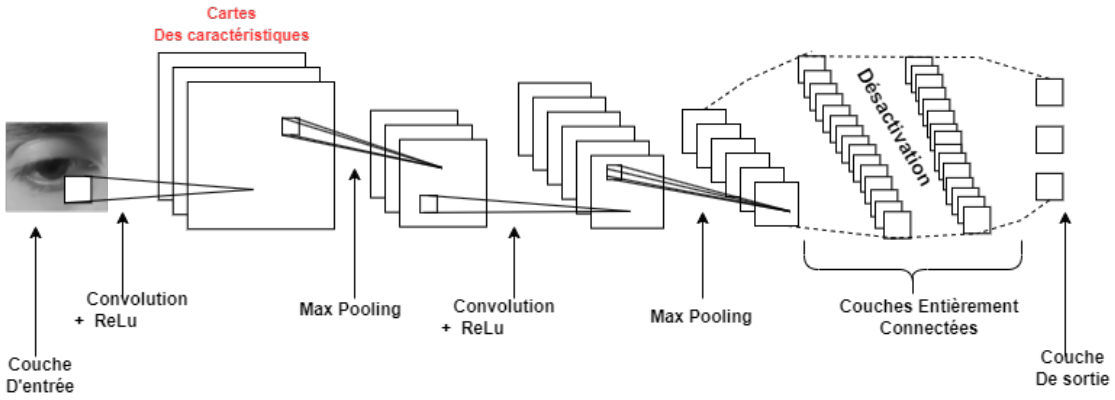


FIGURE 1.9 – Architecture d'un CNN .

Les CNN sont constitués de plusieurs couches, chacune ayant une fonction spécifique. Les couches de CNN montées dans la figure 1.9 fonctionnent comme suite [55] :

Couches de convolution : Les couches de convolution sont les éléments constitutifs principaux des CNN. Elles sont responsables de l'extraction des caractéristiques des images.

L'entrée de la couche de convolution est généralement une image ou une carte de caractéristiques provenant d'une couche précédente, qui peut être considérée comme une fonction. L'opération de convolution consiste à faire glisser une petite matrice appelée noyau (Kernel) ou filtre sur la fonction d'entrée et à effectuer une multiplication entre le noyau et les valeurs correspondantes de la fonction d'entrée. Ce processus est répété sur l'ensemble de la fonction d'entrée, ce qui produit une nouvelle fonction appelée carte de caractéristiques de sortie.

Couches de Pooling : Les couches de pooling sont utilisées pour réduire la taille des cartes de caractéristiques produites par les couches de convolution. Le pooling permet de sélectionner les caractéristiques les plus importantes en prenant soit le maximum (max pooling) en glissant une fenêtre de taille fixe sur la carte de caractéristiques , soit la moyenne (global average pooling) de toutes les valeurs de la carte de caractéristiques.

Couches entièrement connectées : Les couches entièrement connectées sont utilisées pour classifier l'image d'entrée. Elles prennent en entrée les sorties des couches de pooling et les connectent à un certain nombre de neurones de sortie, chacun représentant une classe différente.

Couche de désactivation : La couche désactivation est une technique de régularisation utilisée pour prévenir le surajustement. Elle désactive de manière aléatoire une fraction des unités d'entrée à chaque étape d'entraînement, ce qui aide à réduire les interdépendances entre les neurones. Les couches Dropout peuvent être ajoutées après les couches entièrement connectées ou les couches convolutives.

Couche Flatten : Une couche de flatten est utilisée pour convertir la sortie d'une couche

précédente en un vecteur 1D. Elle est généralement utilisée pour connecter les couches convolutives aux couches entièrement connectées.

Couches d'activation : Les couches d'activation introduisent des non-linéarités dans le réseau en appliquant une fonction d'activation à la sortie d'une couche. Les couches d'activation aident le réseau à apprendre des relations complexes entre les caractéristiques.

La fonction d'activation ReLU (Rectified Linear Unit) est une fonction très couramment utilisée dans les réseaux neuronaux. Elle prend une valeur d'entrée et la renvoie sans modification si elle est positive, zéro si elle est négative.

$$f(x) = \max(0, x) \tag{1.3}$$

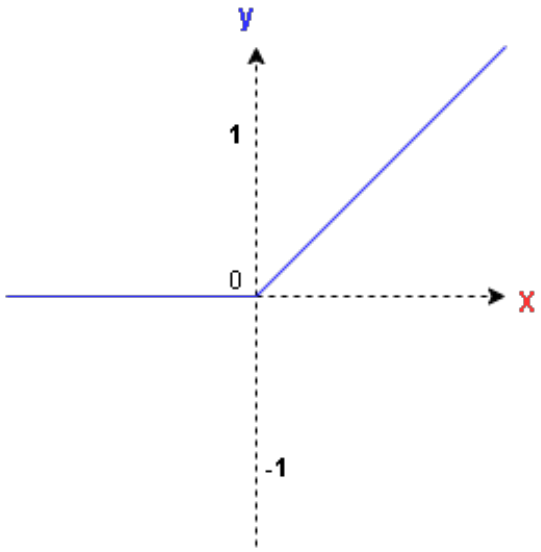


FIGURE 1.10 – La fonction d'activation ReLU.

1.4.2 Réseaux de neurones récurrents (RNN – Recurrent Neural Network)

Un réseau de neurones récurrents (RNN) est un réseau de neurones artificiels présentant des connexions récurrentes afin de se souvenir des événements passés. Ce réseau est constitué de neurones inter-connectés interagissant non-linéairement et pour lesquels il existe au moins un cycle dans la structure. En d'autre termes, c'est un réseau dont les neurones s'envoient des signaux de rétroaction les uns aux autres. Les réseaux neuronaux récurrents conviennent aux données d'entrée de taille variable. Ils sont particulièrement utiles pour l'analyse des séries temporelles. Ils sont utilisés pour la reconnaissance automatique de la parole ou de l'écriture, aussi plus généralement pour la reconnaissance des formes [46].

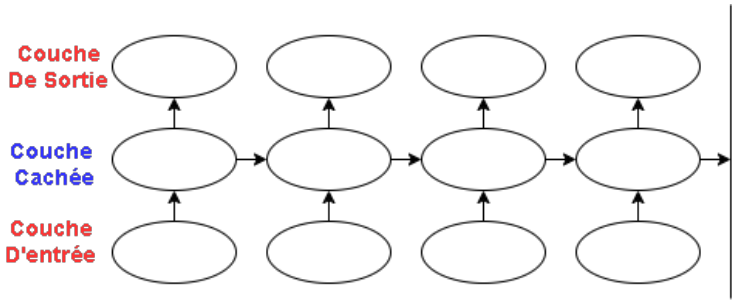


FIGURE 1.11 – Diagramme de RNN [66].

La Figure 1.11 illustre comment les informations circulent dans le RNN entre les différentes couches, allant de la couche d'entrée à la couche de sortie, en passant par la couche cachée (récurrente) qui permet au modèle de saisir et de gérer les dépendances temporelles et séquentielles présentes dans les données.

1.5 Conclusion

Ce chapitre offre un aperçu de différents types d'apprentissage automatique. Nous avons commencé par définir l'apprentissage automatique et ses différents types. Le premier type que nous avons présenté est l'apprentissage supervisé, qui consiste à entraîner des algorithmes sur des données étiquetées. Trois classificateurs de ce type ont été introduits : les Machines à Vecteurs de Support (SVM), les k-plus proches voisins (KNN) et les arbres de décision (DT). Chacun de ces classificateurs offre une approche unique pour résoudre des problèmes de classification en utilisant des méthodes différentes. Ensuite, nous avons abordé l'apprentissage non supervisé avec l'algorithme de k-means clustering. Nous avons également étudié l'apprentissage par renforcement et les méthodes d'ensemble pour améliorer leurs performances et la précision globale de plusieurs modèles. Nous avons présenté pour ce dernier type les trois méthodes les plus utilisées. Nous avons ensuite défini l'apprentissage profond qui est en soi un sous-domaine de l'apprentissage automatique, fondé sur le principe d'une succession de couches permettant d'extraire les caractéristiques des données d'entrée. Les réseaux de neurones convolutifs (CNN) et récurrents (RNN) ont été étudiés.

Dans le prochain chapitre, nous étudierons quelques protocoles d'authentification biométrique basée sur les mouvements des yeux, ainsi que les algorithmes utilisés pour les testes et les résultats obtenus.

Chapitre 2

Etat de l'art sur l'authentification biométrique basée sur les mouvements oculaires

2.1 Introduction

L'authentification biométrique est de nos jours utilisée dans plusieurs institutions publiques et privées. Ce chapitre est consacré à un état de l'art de l'authentification biométrique basée sur les mouvements des yeux. Nous commençons par retracer brièvement l'historique de l'authentification. Ensuite, nous présentons une étude de cas dans laquelle nous nous concentrons sur deux types d'authentification : l'authentification à un facteur et l'authentification à deux facteurs. Enfin, nous discuterons les systèmes d'authentification étudiés et leurs résultats.

2.2 Historique sur l'authentification

L'authentification par ordinateur a connu une évolution significative au fil des décennies, passant des méthodes limitées de mots de passe et de cartes à puce dans les années 60 et 70, aux méthodes biométriques telles que les empreintes digitales et l'iris dans les années 90, et aux méthodes de plusieurs facteurs dans les années 2000 [8].

Les avancées technologiques telles que la reconnaissance d'image et les caméras infrarouges ont permis le développement de nouvelles technologies de suivi oculaire avec une précision et une rapidité accrues.

Des études ont montré dès 2003 [29] que le suivi oculaire pouvait identifier les individus avec fiabilité, et les technologies de suivi oculaire ont depuis connu des améliorations, permettant des applications plus avancées. L'identification par le mouvement oculaire a connu des progrès significatifs en 2010 [83] grâce à l'apprentissage automatique .

En 2012, des chercheurs ont mis au point un système de suivi oculaire basé sur des lunettes, utilisant les mouvements et les clignements des yeux pour identifier l'utilisateur. Récemment, des

chercheurs ont également exploré l'utilisation du suivi oculaire pour l'authentification à distance, permettant à un système de vérifier l'identité d'un utilisateur à partir d'une vidéo enregistrée en direct [40].

2.3 Protocoles d'authentification basés sur les mouvements oculaires

L'authentification fait référence à l'identification d'un utilisateur pour confirmer qu'il est bien la personne qu'il prétend être avant de lui accorder l'accès aux ressources d'information [73]. Nous classifions les systèmes d'authentification basés sur le mouvement oculaire étudiés en deux types : l'authentification à un facteur et l'authentification à deux facteurs. Dans le premier type, le système exige un seul élément d'authentification connu (ou détenu) exclusivement par l'utilisateur [26]. En revanche, dans le deuxième type, le système exige la présence de deux éléments d'authentification afin d'améliorer la sécurité [26].

2.3.1 Authentification à un facteur

Les protocoles d'authentification de cette classe exigent une seule preuve d'identité, un mot de passe par exemple, dans cette étude ce sont les mouvements des yeux. L'extraction des caractéristiques uniques des mouvements oculaires des personnes se fait à l'aide d'outils à haute précision.

2.3.1.1 Eye Know You Too : A DenseNet Architecture for End-to-end Eye Movement Biometrics

Lohr et al. [50] ont introduit une nouvelle architecture de reconnaissance biométrique appelée Eye Know You Too (EKYT) qui fonctionne de manière intégrale. Elle utilise les réseaux de neurones DenseNet pour extraire les caractéristiques des mouvements oculaires et les utiliser afin d'identifier chaque individu de manière unique. Selon les auteurs cette approche est la première à atteindre un niveau de performance d'authentification qui serait acceptable pour une utilisation réelle.

L'architecture Eye Know You Too (EKYT) est une architecture de réseau neuronal qui effectue la transformation suivante :

$$f : R : C \times T \rightarrow R : 128, \tag{2.1}$$

Le modèle prend en entrée deux paramètres : C qui représente le nombre de canaux d'entrée et T qui représente la longueur de la séquence d'entrée. Il génère une sortie qui est un vecteur de 128 caractéristiques apprises extraites des données d'entrée.

L'architecture telle qu'illustrée sur la Figure 2.1 est structurée comme suit :

1. — Les données en entrée passent par un bloc dense composé de 8 couches de convolution unidimensionnelles.
   — Chaque couche de convolution produit 32 cartes de caractéristiques.
   — Les cartes de caractéristiques sont ensuite concaténées avec les cartes de caractéristiques précédentes (reçues de la couche précédente) avant d'être envoyées à la couche de convolution suivante.
   — La sortie du bloc dense cartes de caractéristiques concaténées finales.
2. Une couche de Global Average Pooling (moyenne globale) est appliqué aux cartes de caractéristiques concaténées finales. Il réduit les dimensions spatiales à une seule valeur par carte de caractéristiques.
3. Les cartes de caractéristiques obtenues après le pooling sont aplaties en un vecteur unidimensionnel.
4. Finalement, le vecteur aplati est transmis à une couche entièrement connectée pour produire un embedding de dimension 128. Cet embedding constitue une représentation compacte et dense qui capture les caractéristiques essentielles de la séquence d'entrée.
5. Lorsqu'une classification est requise, une couche entièrement connectée est ajoutée après la couche d'embedding. La couche de classification donne en sortie les logits de classe.
6. Toutes les couches de convolution, global average pooling et la couche de classification sont précédés de la normalisation par lots [15] (Batch Normalization ou BN). Ensuite, la fonction d'activation ReLU est appliquée après la BN

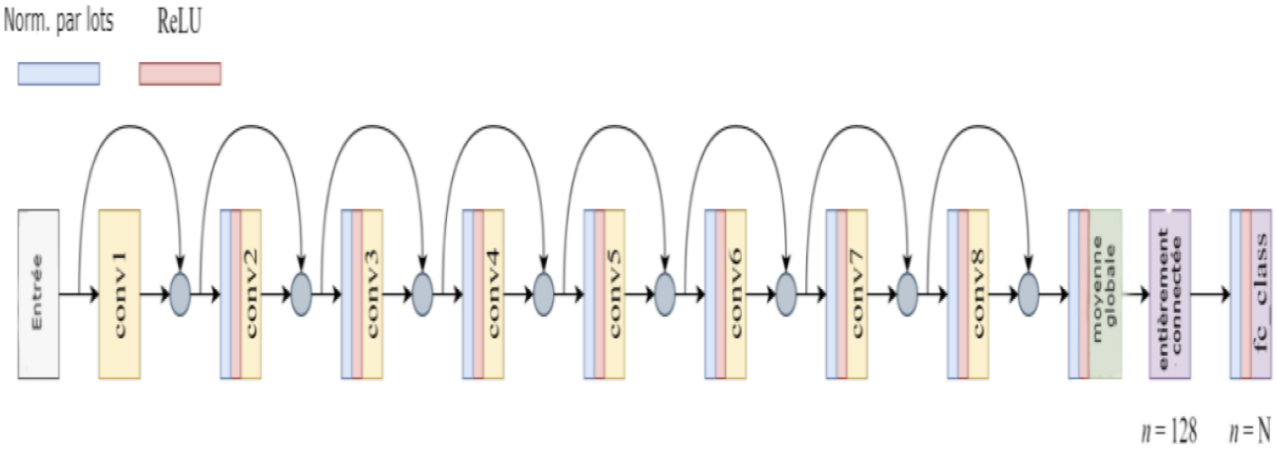


FIGURE 2.1 – L'architecture Eye Know You Too (EKYT) [50].

Pour implémenter cette architecture, ils ont utilisé leur jeu de données Gaze Base [37] qui sera présenté dans le chapitre suivant (voir Section 3.3.1.1). Les auteurs, ont pré-traité les données

en utilisant des techniques telles que la dérivation de Savitzky-Golay [72] et la transformation z-score [7]. Par la suite, ils ont entraîné leur modèle en utilisant une combinaison de deux fonctions de perte : la perte multi-similarité[86], et la perte de cross-entropy catégorielle [91]. La perte multi-similarité est une fonction de perte utilisée dans des tâches qui impliquent l'apprentissage de similarités ou de métriques de distance entre des paires d'échantillons. Elle vise à encourager la similarité entre les paires d'échantillons considérées similaires et la dissimilarité entre les paires considérées différentes. D'autre part, la perte de cross-entropy catégorielle est utilisée dans des tâches de classification multi-classe pour mesurer la dissimilitude entre les probabilités de classe prédites et les vraies étiquettes de classe.

Les résultats obtenus montrent un taux égalité d'erreurs de 3,66% en utilisant seulement 5 secondes de mouvements oculaires, qui est comparable au temps nécessaire pour entrer un code PIN à 4 chiffres ou pour calibrer un dispositif de suivi oculaire pour l'enrôlement et l'authentification. En outre, avec 30 s de données, ils ont obtenu une estimation de 5,08% de FRR, un taux d'erreur d'identification acceptable pour une utilisation en situation réelle.

2.3.1.2 User Identification Utilizing Minimal Eye-Gaze Features in Virtual Reality Applications

Sarker et al. [71] ont exploré les méthodes d'apprentissage automatique et d'apprentissage profond sur les données de regard oculaire pour identifier les utilisateurs avec une précision raisonnable sans aucune tâche d'authentification explicite et avec un nombre minimum de caractéristiques oculaires (figure 2.2).

Les auteurs ont entraîné et testé deux modèles d'apprentissage automatique tel que Random Forest (RF) et k-nearest-neighbors (kNN), ainsi que deux modèles d'apprentissage profond (DL) tel que Convolutional Neural Networks (CNN) et Long Short-Term Memory (LSTM) pour l'identification des utilisateurs.

RF combine les résultats des arbres de décision pour prédire les données finales. Les hyperparamètres de RF ont été optimisés en utilisant "RandomizedSearchCV" [1]. Cette approche a permis de trouver le meilleure combinaison d'hyperparamètre de RF pour améliorer la performance du modèle, avec 200 estimateurs, une profondeur maximale de 460 et la considération maximale de caractéristiques .

Pour le classifieur KNN, les valeurs de k ont été ajustées afin d'éviter le sur-ajustement en tenant compte de la variance des données. Après avoir évalué différentes valeurs de k de 1 à 10, la valeur optimale choisie a été k = 5, en utilisant la métrique de Minkowski [70] par défaut.

Le réseau neuronal convolutif est composé de deux couches Conv1D (ReLU) et de deux couches denses entièrement connectées. Pour réduire la dimension de sortie, un max pooling avec une taille de 2 a été appliqué. Une couche de désactivation (dropout) de 40% a été ajoutée pour prévenir

le sur-ajustement. L'optimiseur Adam [48], avec un taux d'apprentissage de 10^{-3}, a été utilisé en conjonction avec l'entropie croisée catégorique comme fonction de perte.

Pour capturer les caractéristiques spatiales et temporelles des données de regard, les réseaux LSTM (Long Short-Term Memory) ont été utilisés. Les hyperparamètres du modèle LSTM comprenaient une couche de désactivation de 40% ainsi qu'une activation ReLU pour la première couche LSTM et la troisième couche entièrement connectée. La dernière couche dense utilisait une activation softmax pour effectuer la classification des 34 utilisateurs en sortie. Le modèle a été entraîné pendant 50 époques avec une taille de lot (batch size) préétablie.

Ils ont rassemblé un nouvel ensemble de données dans un environnement de réalité virtuelle (voir la fiqure 2.2) pour entraîner et tester les architectures proposées. Les résultats obtenus ont ensuite été comparés avec un ensemble de 12 caractéristiques importantes identifiées par l'algorithme d'élimination récursive de fonctionnalités (Recursive Feature Elimination) [89] et un sous-ensemble de six caractéristiques brutes.

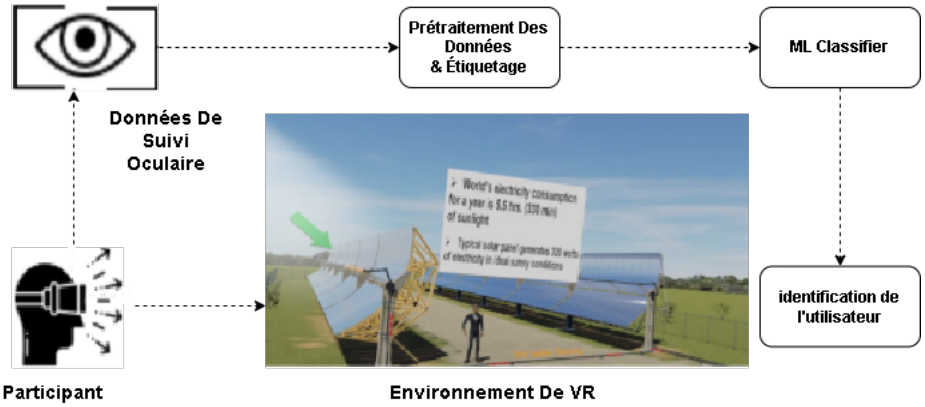


FIGURE 2.2 – Processus de système d'identification des utilisateurs [71].

La figure 2.2 est une illustration associée à l'environnement de réalité virtuelle utilisé pour collecter le nouvel ensemble de données nécessaires à l'entraînement et aux tests des architectures proposées.

Les modèles d'apprentissage automatique et de l'apprentissage profond ont montré une précision de plus de 98% en utilisant uniquement ces six caractéristiques issues de l'ensemble de caractéristiques brutes, mais ils ont identifié des risques de confidentialité liés au partage de données de regard oculaire et recommandent de coder les caractéristiques les plus importantes pour minimiser l'identification des utilisateurs. Ils ont suggéré que cette approche pourrait être appliquée à d'autres applications de RV pour identifier les utilisateurs avec des fonctionnalités minimales basées sur le regard.

2.3.1.3 NeuroBiometric : An Eye Blink Based Biometric Authentication System Using an Event-Based Neuromorphic Vision Sensor

Guang et al. [33] ont proposé une toute nouvelle approche d'authentification par clignements des yeux. Cette approche utilise des signaux transitoires de clignement des yeux qui se produisent lorsque les yeux clignent. Ils peuvent être détectés à l'aide de capteurs de vision neuromorphiques, qui offrent une résolution temporelle de l'ordre de la microseconde. À partir de ces signaux, Guang et al. ont proposé un ensemble de caractéristiques biométriques décrivant le mouvement, la vitesse, l'énergie et le signal de fréquence des clignements des yeux, qui sont utilisées pour identifier et vérifier l'identité de l'utilisateur.

Les auteurs ont rassemblé un jeu de données de signaux de clignement des yeux (voir Section 3.3.1.3). Ensuite, ils ont effectué un traitement de signal afin d'extraire les caractéristiques les plus pertinentes. Le capteur DAVIS346 [87] est utilisé dans cette étude. Ce capteur détecte les variations de l'intensité lumineuse et chaque circuit de pixel du capteur suit les changements de contraste au fil du temps. L'augmentation de l'intensité lumineuse génère un événement ON (ou événement positif), tandis qu'une diminution de l'intensité lumineuse entraîne un événement OFF (ou événement négatif). Un événement est déclenché lorsque le contraste temporel dépasse une valeur de seuil. La densité d'événements est le nombre d'événements dans un intervalle de temps (une fenêtre de 10 ms). Ils ont également utilisé l'algorithme de communication de pixels voisins (Neighboring Pixel Communication) [39] pour filtrer le signal. À la fin du traitement, ils ont extrait 204 caractéristiques.

Pour déterminer les caractéristiques les plus significatives, ils ont appliqué l'élimination récursive de fonctionnalités sur l'ensemble des 204 caractéristiques. Le sous-ensemble résultant est réduit encore par deux approches différentes, l'une basée sur les coefficients de corrélation de Pearson [11] et l'autre utilisant le coefficient de variation (CoV) [54].

Finalement, deux modèles ont été entraînés pour l'authentification : le SVM (Machine à Vecteurs de Support) et le Bagging LDA (Analyse Discriminante Linéaire en Ensachage).

Pour l'authentification à l'aide du SVM, une évaluation a été réalisée en utilisant quatre types de noyaux différents à savoir le noyau linéaire, polynomial, RBF (Radial Basis Function) et sigmoïde. Le noyau linéaire a montré les meilleures performances parmi les modèles non-ensemble. Il a donc été choisi pour les expériences ultérieures.

Quant à l'authentification avec le Bagging LDA, c'est un algorithme qui utilise un ensemble de modèles LDA comme classifieurs de base pour créer un classifieur en ensemble. Chaque classifieur de base ajuste des sous-ensembles aléatoires des données en utilisant la méthode d'ensachage (bootstrap aggregating) [47] et agrège leurs prédictions pour obtenir la prédiction finale. Lors de l'échantillonnage bootstrap, 80% des échantillons et 100% des caractéristiques de l'ensemble

d'entraînement total sont sélectionnés. Ensuite, k modèles LDA sont ajustés en utilisant la décomposition en valeurs singulières. L'identité finale de chaque échantillon est déterminée par agrégation par vote. Bien qu'augmenter la valeur de k puisse améliorer la capacité d'ajustement du modèle, cela entraîne également une augmentation du temps de calcul. Durant l'étude, ils ont constaté que k=30 offrait un bon compromis entre vitesse de calcul et précision. Le modèle d'ensemble a atteint une précision de 0,948 tandis que le modèle non-ensemble a obtenu une précision de 0,925, avec un taux de faux positif d'environ 0.002.

2.3.2 Authentification à deux facteurs

La classe d'authentification à deux facteurs regroupe les protocoles qui nécessitent deux identifiants pour effectuer la tâche d'authentification. Cette mesure améliore la sécurité du protocole et protège mieux les utilisateurs contre l'usurpation d'identité.

2.3.2.1 BlinKey : A Two-Factor User Authentication Method for Virtual Reality Devices

Houadi et al. [41] ont proposé un nouveau système d'authentification à deux facteurs appelé "BlinKey" pour les dispositifs de réalité virtuelle équipés de traceurs oculaires. Lors de l'authentification, un utilisateur utilise une séquence de clignements comme mot de passe unique. Ce mot de passe est combiné avec les variations de la taille de la pupille entre les clignements consécutifs. Le système proposé repose sur deux types de caractéristiques distincts : les caractéristiques basées sur la connaissance et les caractéristiques basées sur la biométrie.

— **Caractéristiques basées sur la connaissance :** c'est les caractéristiques issues de la série de clignements.

1. **Instant de clignement :** La séquence de clignements peut être identifiée de manière unique par un ensemble horodaté de débuts et de fins de clignements.
2. **Intervalle entre clignements :** La durée entre les débuts de deux clignements consécutifs.
3. **Intervalles relatifs :** C'est le rapport d'un intervalle entre un clignement donné et son précédent.

— **Caractéristiques basées sur la biométrie :**

1. **Coefficients de Fourier :** Ils ont appliqué la transformée de Fourier rapide (Fast Fourier Transform) sur des échantillons du domaine temporel pour extraire les informations de la variation de la taille de la pupille.
2. **Caractéristiques statistiques :** En plus des coefficients de Fourier, ils ont sélectionné un ensemble de caractéristiques dans le domaine du temps et de la fréquence. Cet en-

semble comprend les caractéristiques suivantes : Maximum, Minimum, Moyenne, Médiane, quadratique moyenne , Écart-type , Déviation absolue moyenne , Kurtosis, inclinaison, Écart inter-quartile, Rugosité, Netteté, Croisements de la moyenne , Amplitude de Willison, Changement de signe de la pente.

Le système BlinKey comprend deux phases : la phase d'inscription et la phase de connexion. Les deux phases sont suivies d'un traitement des données.

- Dans une scène virtuelle en pop-up, l'utilisateur est invité à cligner des yeux selon un schéma qu'il a lui-même créé, cela est une entrée "blinkey".
- Une fois la procédure d'authentification activée, le traceur oculaire enregistre en continu les signaux de taille de pupille en temps réel de l'utilisateur et les transmet au serveur.
- Le signal passe d'abord par le module de détection de début/fin pour segmenter l'ensemble du "blinkey".
- Le signal brut est ensuite pré-traité à l'aide de deux processus :
  - Le premier est le dé-bruitage, qui permet d'extraire les clignements volontaires et les adaptations de la pupille.
  - Le deuxième processus est la décomposition, qui permet d'extraire du signal débruité le rythme des clignements de l'utilisateur et les segments qui portent les caractéristiques basées sur les connaissances et les caractéristiques biométriques d'un "blinkey".
- Ils sont ensuite alimentées dans le module d'extraction de caractéristiques pour extraire des caractéristiques basées sur les connaissances et biométriques.
- Enfin, le classifieur KNN décide si le "blinkey" donné est légitime ou non.

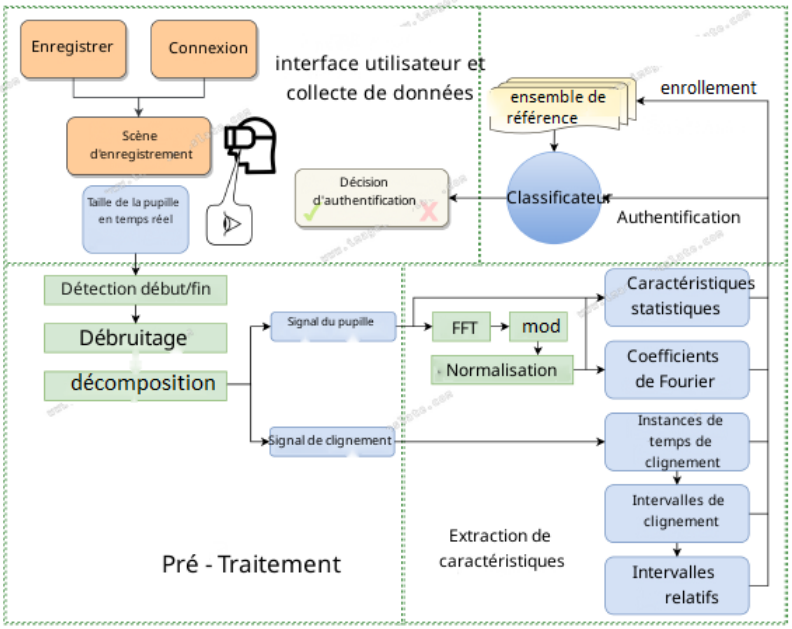


FIGURE 2.3 – Architecture de système Blinkey [41].

Pour l'implémentation de système, les auteurs ont développé le prototype de BlinKey sur un

dispositif de tête HTC Vive Pro, connecté à un serveur local exécutant SteamVR pour prendre en charge l'environnement de réalité virtuelle. Un traceur oculaire Pupil Labs est installé dans l'appareil RV pour enregistrer la taille de la pupille en temps réel. Le taux d'échantillonnage est fixé à 200 Hz, ce qui signifie que des échantillons de la taille de la pupille sont collectés toutes les 5 millisecondes. Les données collectées sont transmises au serveur via l'interface de programmation d'application ZeroMQ. Toutes les fonctions, telles que la détection de début/fin, le pré-traitement, l'extraction de caractéristiques et la classification, sont mises en œuvre dans Unity, un moteur multiplateforme pour les jeux en réalité virtuelle. L'architecture de Blinkey est illustrée sur la Figure 2.3.

Le système a atteint un taux d'égalité d'erreur (TEE) de 4,0% avec seulement 6 échantillons d'entraînement par participant.

2.3.2.2 Biométrie et Réalité Virtuelle de Goussem et Djallil

Goussem Ayoub et Djallil Massinissa [31] se sont inspiré du modèle de Blinkey [41] pour proposer un nouveau système d'authentification barométrique basé sur le comportement des yeux.

Dans leur étude, ils ont utilisés deux types de caractéristiques :

— **Caractéristiques basées sur la connaissance :** ce sont les informations que l'utilisateur introduits lors de l'authentification qui sont un nom d'utilisateur unique et une série de clignements propre à chaque utilisateur.

— **Caractéristiques basées sur la biométrie :** se sont des caractéristiques extraites lors de traitement de signal des yeux. Les caractéristiques utilisés sont :

1. **Quotient minimum :** La valeur minimale du rapport entre la distance verticale (distance entre le point le plus haut de la paupière supérieure et le point le plus bas de la paupière inférieure) et la distance horizontale (entre le point le plus à gauche et le point le plus à droite).
2. **Quotient maximum :** La valeur maximale du rapport entre la distance verticale (distance entre le point le plus haut de la paupière supérieure et le point le plus bas de la paupière inférieure) et la distance horizontale (entre le point le plus à gauche et le point le plus à droite).
3. **Quotient moyen :** La valeur moyenne du rapport entre la distance verticale (distance entre le point le plus haut de la paupière supérieure et le point le plus bas de la paupière inférieure) et la distance horizontale (entre le point le plus à gauche et le point le plus à droite).
4. **FFT maximum :** Le pic maximum retourné par la fonction FFT.

5. **Taux oeil ouvert :** Le taux des yeux ouverts.
6. **Taux oeil fermé :** Le taux des yeux fermés.

L'architecture de leur modèle est composé de 03 phases : la phase d'identification, la phase d'inscription et la phase d'authentification (Figure 2.4). Le protocole nécessite un casque VR avec une caméra attachée pour capturer les clignements.

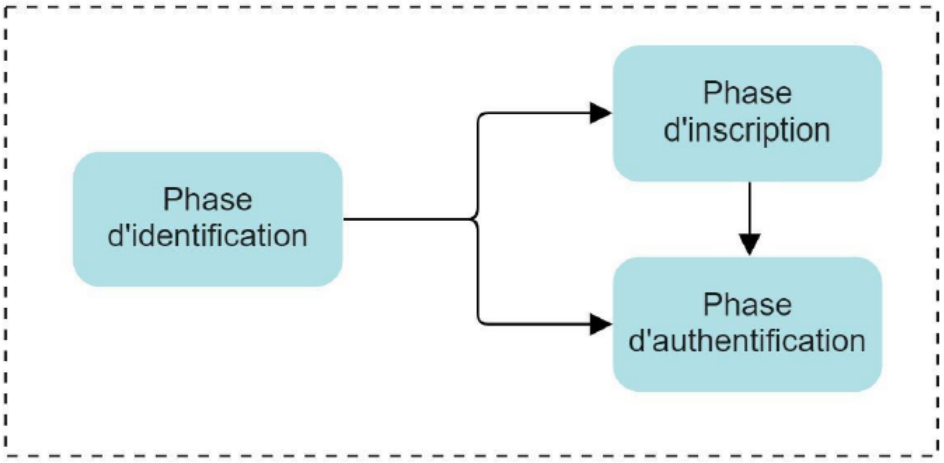


FIGURE 2.4 – Architecture du modèle de Goussem et Djallil [31].

Au cours de la première phase illustrée par la Figure 2.5, l'utilisateur entre son nom d'utilisateur dans l'interface d'identification. Ensuite, en fonction de sa situation dans le système (inscrit ou non), il peut choisir entre l'inscription ou l'authentification.

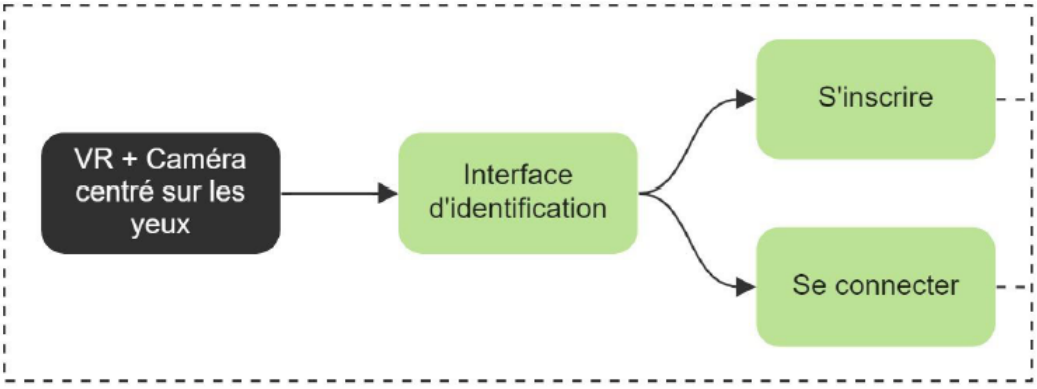
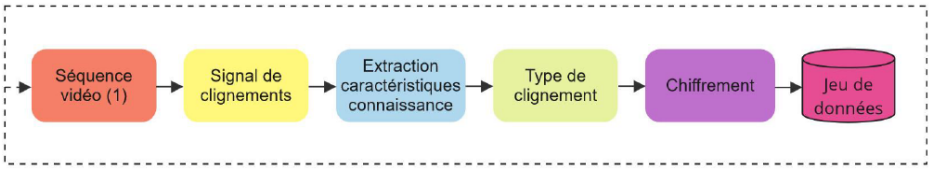
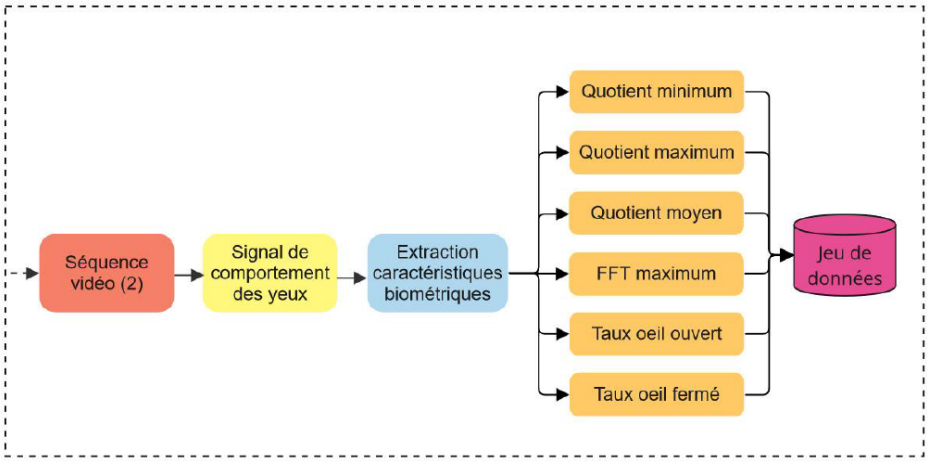


FIGURE 2.5 – Phase d'identification du modèle de Goussem et Djallil [31].

La deuxième phase se divise en deux parties : une partie consciente (Figure 2.6a) où l'utilisateur choisit sa séquence de clignements et une partie biométrique (Figure 2.6b) où le système collecte les caractéristiques biométriques de l'utilisateur pour ensuite les enregistrer dans le jeu de données.



(a) partie consciente



(b) partie biométrique

FIGURE 2.6 – Phase d'inscription du modèle de Goussem et Djallil [31]

La phase d'authentification est elle même divisée en deux parties : une partie consciente où l'utilisateur entre la séquence de clignements qu'il a choisie précédemment, et une partie biométrique qui fonctionne comme celle de la phase d'inscription (2.6b). À la fin de cette phase, le système décide d'authentifier ou de rejeter l'utilisateur après un processus de classification à l'aide d'un classifieur KNN. Les caractéristiques extraites dans la partie biométrique sont comparées à celles du jeu de données en appliquant la distance euclidienne avec une valeur de seuil fixe.

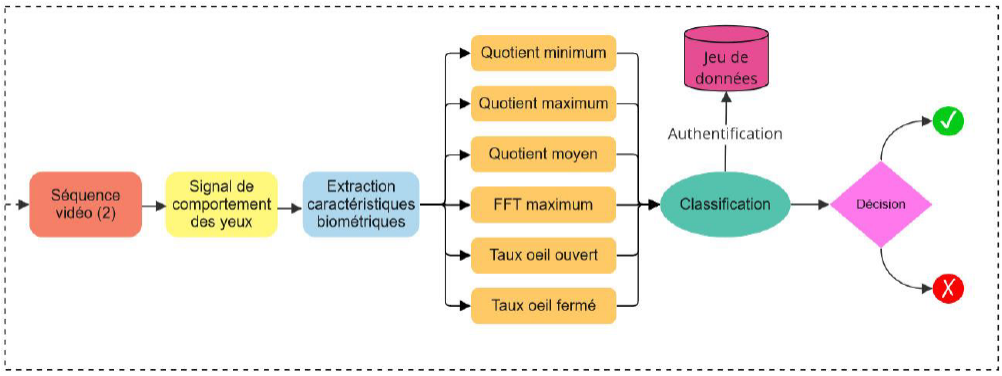


FIGURE 2.7 – Phase d'authentification du modèle de Goussem et Djallil [31].

Afin d'évaluer leur modèle, les auteurs ont collecté un ensemble de données (présenté dans 3.3.2.1). Ils ont effectué la classification avec différentes valeurs de K (nombre de voisins) et des valeurs de seuil comprises entre 4.5, 5, 5.5 et 6.

Les meilleurs résultats ont été obtenus pour l'œil gauche atteignant un taux d'égalité d'erreurs (TEE) de 35%, avec un k=2 et un seuil égale à 5. Leur système a atteint une précision maximale de 65%.

2.4 Discussion des résultats

Dans notre étude de cas, nous avons examiné deux classes d'authentification biométrique basée sur le mouvement oculaire : l'authentification à un facteur et l'authentification à deux facteurs.

Dans la première classe, nous avons étudié la proposition de Lohr et al. [50], une architecture EMB (Eye Movement Biometrics) basée sur les réseaux de neurones DenseNet. L'objectif était d'évaluer la performance d'authentification de l'EMB avec des données de haute qualité pour une utilisation pratique.

Sarker et al. [71] ont exploré la possibilité d'authentifier les utilisateurs en utilisant un nombre réduit de caractéristiques biométriques des mouvements des yeux.

Guang et al. [33] ont utilisé les capteurs de vision neuromorphiques pour proposer une nouvelle méthode d'authentification basée sur les caractéristiques des clignements des yeux. Les capteurs de vision neuromorphiques ont la particularité de détecter les changements locaux au niveau des pixels causés par les clignements des yeux, offrant ainsi des avantages tels qu'une réponse ultra rapide. Les résultats expérimentaux ont démontré que cette méthode permet d'identifier et de vérifier les sujets avec une grande précision.

Parmi les systèmes d'authentification à deux facteurs, nous avons analysé Blinkey, un système d'authentification à deux facteurs dans les environnements de réalité virtuelle proposé par Huadi et al. [41]. Les utilisateurs possèdent un code unique (une série de clignements) pour s'authentifier dans le système, à partir de laquelle un ensemble de caractéristiques est extrait pour juger de l'authenticité de l'utilisateur. Ces caractéristiques sont catégorisées en deux types : les caractéristiques basées sur la connaissance et les caractéristiques basées sur la biométrie.

Goussem Ayoub et Djallil Massinissa [31] ont proposé un système d'authentification biométrique dans la réalité virtuelle inspiré de Blinkey. Ce système est fondé sur les caractéristiques de connaissance (nom d'utilisateur + séquence de clignements) et les caractéristiques biométriques (extraites du signal de clignements).

Les protocoles 1FA ont obtenu de meilleurs résultats en termes de TEE que les protocoles 2FA. Le TEE le plus bas (3,66%) a été observé dans le protocole de Lohr et al. [50] qui ont utilisé tandis que Huadi et al. [41] ont obtenu un TEE de 4,0% avec KNN. Sarker et al. [71] quant à eux, ont obtenus la précision la plus élevée (98%) en utilisant KNN. Ils ont aussi utilisé CNN, RF et LSTM et ont obtenus une précision presque similaire. Goussem Ayoub et Djallil Massinissa [31] ont utilisé KNN et ont obtenu une précision de 65 %.

Il est important de noter que la comparaison des performances entre les deux types de proto-

coles peut ne pas être définitive, car différentes études ont utilisé des algorithmes différents (KNN, CNN , SVM, etc.) et des ensembles de données différents. La comparaison serait plus solide et significative si les mêmes ensembles de données étaient utilisés de manière cohérente dans différentes études et avec les différents algorithmes utilisés. Par conséquent, une étude plus approfondie de l'ensemble de données est nécessaire pour une comparaison plus significative et fiable entre les différents protocoles et classes d'authentification.

Classe d'authentification	**Article**	**Algorithmes**	**Type de dataset**	**Résultats**
Authentification à un facteur	Lohr et al. [50]	CNN	Signals	TEE=3,66% pour 5s TEE= 5,08% pour 30s
	Sarker et al. [71]	RF KNN LSTM CNN	Signals	précision=98% en utilisant uniquement six caractéristiques
	Guang et al. [33]	SVM Bagging LDA	Signals	précision=0,948 pour le modèle d'ensemble précision=0,925 pour le modèle non-ensemble TFP=0.002
Authentification à deux facteurs	Huadi et al. [41]	KNN	Vidéos	TEE=4,0% avec 6 échantillons/participant
	Goussem Ayoub, Djallil Massinissa [31]	KNN	Vidéos	TEE =35% avec K=2 et seuil=5 précision = 65%.

TABLEAU 2.1 – Tableau comparatif des protocoles d'authentification biométrique basés sur le mouvement oculaire de notre étude.

2.5 Conclusion

Dans ce chapitre, nous avons exposé un état de l'art sur l'authentification biométrique basée sur les mouvements des yeux. Nous avons commencé par un historique de l'authentification biométrique. Ensuite, nous avons réalisé une étude de cas sur les protocoles d'authentification récents, en examinant deux types d'authentification : l'authentification à un facteur et l'authentification à deux facteurs. Enfin, nous avons clôturé le chapitre par une discussion sur ces protocoles et les résultats obtenus. Dans le chapitre suivant, nous procéderons à l'analyse des jeux de données de mouvements oculaires utilisés dans ces travaux, ainsi que d'autres jeux de données utilisés pour l'élaboration des algorithmes de détection de clignements et de mouvements oculaires. De plus, nous proposerons également un nouveau jeu de données.

Chapitre 3

Proposition d'un nouveau jeu de données pour l'authentification biométrique

3.1 Introduction

L'authentification biométrique basée sur les caractéristiques oculaires est devenue une méthode largement adoptée dans les systèmes d'authentification récents. Jusqu'à présent, divers jeux de données ont été utilisés dans la littérature. Dans ce chapitre, nous présenterons quelques jeux de données utilisés par les protocoles d'authentification présentés dans le chapitre précédent. Nous discuterons ensuite leur utilités et limitations. Enfin, nous présenteront deux jeux de données pour l'authentification biométrique basée sur les mouvements des yeux. Le premier jeu de données est constitué de vidéos capturant la partie supérieure du visage de 22 participants. Le deuxième jeu de données est constitués d'images extraites à partir des vidéos du premier jeu de données. Ce deuxième jeu de données est divisées en deux parties, un premier ensemble d'images contenant les deux yeux de chaque participant, et un deuxième ensemble d'images contenant les deux yeux séparés de chaque participant.

3.2 Notions préliminaires

Avant de présenter les jeux de données, il est nécessaire de décrire certaines caractéristiques technologiques telles que la résolution d'image et l'ouverture de la caméra.

3.2.1 Résolution d'image

La résolution d'image fait référence au nombre de pixels présents dans une image [82]. Elle est généralement mesurée en largeur et en hauteur, par exemple, 1920x1080 pixels qui représente une définition Full HD. Dans notre étude, une résolution d'image élevée permet de capturer plus de détails et de précision dans les mouvements des yeux.

3.2.2 La Définition d'une image

1. **FHD** : Full High Definition (pleine haute définition), fait référence à une résolution de 1920x1080 pixels. C'est une norme de qualité d'image élevée utilisée dans les dispositifs d'affichage tels que les téléviseurs, les moniteurs et les caméras [88].
2. **HD** : High Definition (haute définition), cela fait référence à une résolution de 1280x720 pixels. C'est également une norme de qualité d'image élevée, mais légèrement inférieure à la résolution FHD [88].

3.2.3 Ouverture de la caméra

L'ouverture de la caméra désigne le diamètre de l'ouverture de la caméra [64], exprimé en valeurs de f-stop. Elle contrôle la quantité de lumière qui atteint le capteur de la caméra. Une ouverture plus grande permet de capturer plus de lumière, ce qui peut être avantageux dans des conditions de faible luminosité. Une bonne ouverture de caméra peut garantir des enregistrements clairs et précis des mouvements des yeux.

3.2.4 Frame Rate (FPS)

Un Frame Rate est mesuré en FPS (Frames Per Second). C'est le nombre d'images par seconde (IPS) d'une vidéo ; il indique combien d'images sont affichées chaque seconde pour créer une vidéo fluide [85]. Un FPS plus élevé donne une apparence plus fluide aux mouvements des yeux.

3.3 Quelques jeux de données sur les mouvements oculaires

Dans cette section, nous examinerons les jeux de données de mouvements oculaires que nous avons divisés en 3 catégories : les jeux de données basés sur les signaux, les jeux de données basés sur les vidéos et les jeux de données basés sur les images. Ces catégories sont illustrées dans la Figure 3.1.

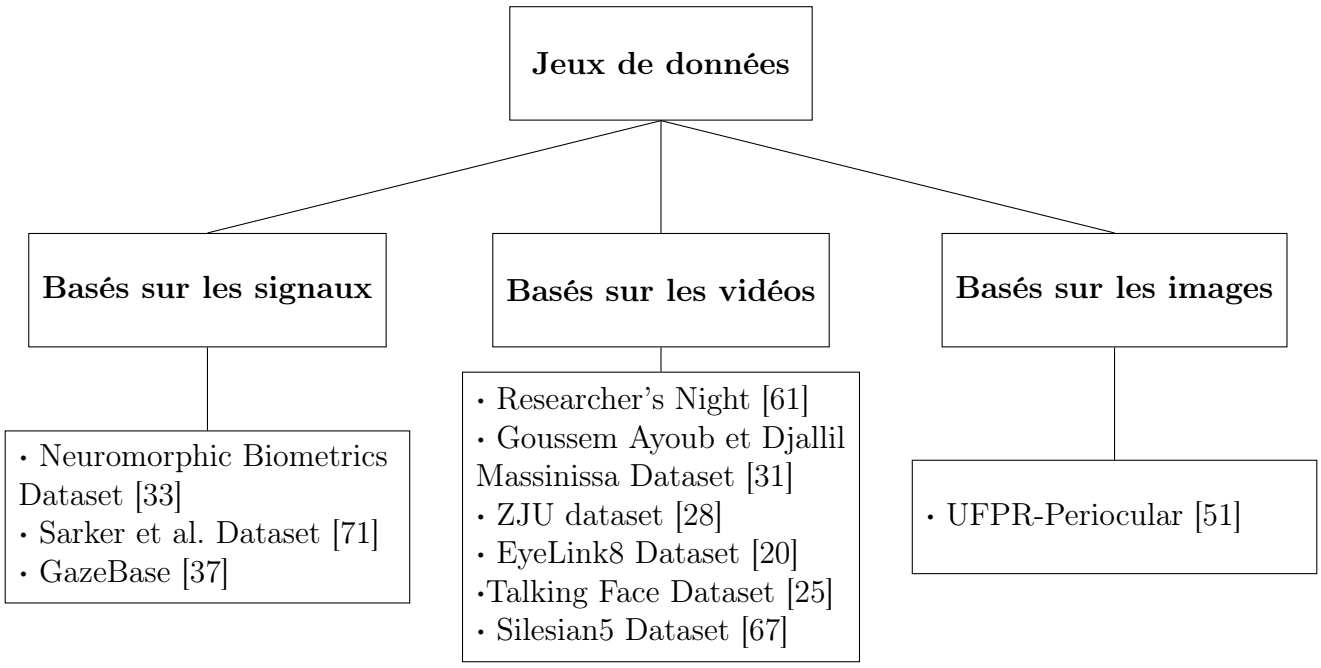


FIGURE 3.1 – Classification des jeux de données étudiés.

3.3.1 Jeux de données basés sur les signaux

Cette classe comprend les jeux de données contenant des signaux de mouvements oculaires enregistrés à l'aide de dispositifs de haute précision tels que des traceurs oculaires et des capteurs neuromorphiques.

3.3.1.1 GazeBase

Lohr et al [37] ont proposé le jeu de données publique GazeBase qui est composé de 12 334 enregistrements de mouvements oculaires de l'œil gauche. Ce jeu de données a été capturé auprès de 322 participants d'âge universitaire. Les signaux capturés comprennent les mouvements de fixation, les saccades horizontales, les saccades obliques aléatoires, les mouvements oculaires lors de la lecture, la vision libre de vidéos cinématographiques et les mouvements oculaires lors de jeux basés sur le regard. Ces signaux permettent d'analyser les mouvements des yeux tels que les fixations, les saccades et les mouvements de poursuite, ainsi que la zone de la pupille.

Ils ont effectué, pour chaque enregistrement, une séquence de sept tâches réparties en deux sessions consécutives, sous la supervision d'un superviseur expérimenté. Les enregistrements ont été effectués dans un environnement calme et sans fenêtres. Les participants assis devant un moniteur d'affichage ont été bien stabilisé à l'aide d'un appui-tête de menton et de front afin de maintenir une position de tête et de corps stable et éviter de cligner des yeux excessivement. Toutes les données ont été collectées à l'aide d'un traceur oculaire EyeLink 1000 (voir Figure 3.2) à une fréquence d'échantillonnage de 1 000 Hz, avec un protocole de calibration et de validation réalisé

avant chaque tâche pour garantir la qualité des données. La collecte de ce jeu de données s'est faite en une période de 37 mois.

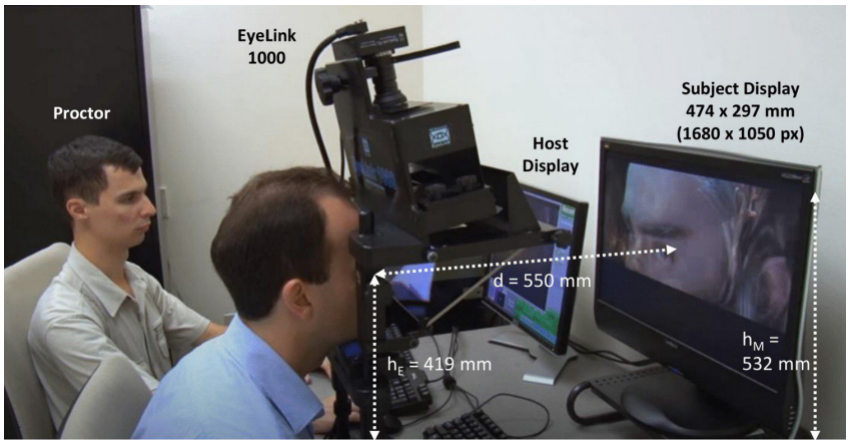


FIGURE 3.2 – Matériels et procédure de collection GazeBase [37].

3.3.1.2 Sarker et al. Dataset

Le jeu de données créé par Sarker et al. [71] concerne 34 participants (25 hommes et 9 femmes). Pour la collecte, les auteurs ont créé un environnement virtuel qui consiste en une visite éducative d'un centre énergétique virtuel. Un avatar explique les objets de la scène à l'aide de sons, d'animations visuelles et de diapositives de texte. La figure 3.3 illustre des captures de cet environnement.

FIGURE 3.3 – L'environnemt virtuel de Sarker et al [71].

Un participant porte un casque de réalité virtuelle HTC Vive Pro Eye (Voir Figure 3.4) à une fréquence de 120 Hz avec un traceur oculaire calibré. L'environnement virtuel a été réalisé avec Unity 3D. Ainsi, les données brutes du regard qui ont été enregistrées comprennent des horodatages, le diamètre des yeux, l'ouverture des yeux, la largeur des yeux, la position du regard et la direction du regard.

FIGURE 3.4 – Casque de réalité virtuelle HTC Vive Pro Eye [5].

L'expérience a été divisée en 4 sessions, chacune portant sur un concept différent. Les 4 sessions durent environ 10 minutes et la durée totale de collecte est de 20 à 25 minutes pour chaque participant. Au total, ce jeu de données comporte 34 enregistrements.

3.3.1.3 Neurmorphic Biometrics Dataset

Guang et al. [33] ont rassemblé un jeu de données de signaux de clignement des yeux enregistrés par le capteur DAVIS346, qui a une résolution de 346 * 260 pixels et une résolution temporelle de 1 microseconde. Ils ont rassemblé 45 volontaires (23 hommes et 22 femmes) se trouvant dans un état psychologique et physiologique normal. Les sujets ont été invités à s'asseoir devant le capteur et à cligner des yeux de manière spontanée. Chaque sujet a effectué au moins 4 sessions de clignement des yeux d'une durée totale de 480 secondes. Une pause de 60 secondes après chaque enregistrement de 120 secondes a été prévue pour éviter les clignements des yeux inhabituels dus à la fatigue et à la distraction éventuelles. La collecte des données a été réalisée dans un environnement intérieur avec une lumière naturelle d'une intensité d'environ 1200 LUX pendant la journée.

3.3.2 Jeux de données basés sur les vidéos

Cette classe regroupe les jeux de données vidéo qui offrent aux chercheurs la possibilité d'extraire manuellement les mouvements des yeux et de mener des études axées sur la région oculaire. Ce type de jeux de données permet une analyse approfondie des mouvements oculaires et offre aux chercheurs la flexibilité d'explorer différents aspects du comportement oculaire et ainsi de la partie précoculaire des yeux.

3.3.2.1 Goussem et Djallil Dataset

Ce jeu de données a été conçu par Goussem Ayoub et Djallil Massinissa pour implémenter leur modèle étudié dans la section 2.3.2.2. Le jeu de données contient 96 vidéos de 16 participants

montrant leur vue frontale. Les participants ont regardé une vidéo de 20 secondes pendant l'enregistrement de leurs mouvements oculaires par une caméra placée en face. Ensuite, la procédure a été répétée six fois pour chaque participant, ce qui a donné un total de 96 vidéos.

3.3.2.2 Researcher's Night Datatset

Le jeu de données Researcher's Night [61] a été collecté lors d'un événement appelé "Researcher's Night" en 2014. Il se compose de 107 vidéos enregistrées en intérieur avec une caméra d'ordinateur, où chaque participant est invité à lire un article sur l'écran ou à cligner des yeux pendant l'enregistrement.

Le jeu de données contient deux sous-ensembles : Researcher's Night 15 et Researcher's Night 30, capturés respectivement à 15 et 30 FPS.

3.3.2.3 ZJU dataset

Le jeu de données ZJU [28] implique 20 participants, dont 13 hommes et 7 femmes, certains portant des lunettes. Les vidéos ont été enregistrées à l'intérieur avec une caméra web de 320x240 pixels à 30 FPS montrant chaque participant dans une vue frontale et une vue vers le haut. Au total, il y a 80 vidéos (4 vidéos par participant) dans le jeu de données.

3.3.2.4 EyeLink8 Dataset

Le jeu de données EyeLink8 [20] contient 8 vidéos de 4 participants (7 hommes et une femme), dont un porte des lunettes. Les vidéos sont enregistrées à l'intérieur avec une résolution de 640x480. Les participants sont assis devant la caméra et agissent de manière naturelle durant l'enregistrement.

3.3.2.5 Talking Face Dataset

Talking Face Dataset [25] n'a pas été créé spécifiquement pour la détection et l'évaluation des clignements des yeux. Cependant, il a été largement utilisé pour ces tâches malgré sa taille réduite. Le jeu de données contient 4 vidéos de 50 secondes chacune montrant une personne (homme) face à la caméra, avec une résolution de 720 * 576 pixels à 25FPS.

3.3.2.6 Silesian5 Dataset

Silesian5 est un sous ensemble d'un jeu de données appelé Silesian Deception Database [67],un jeu de données qui est composé de 101 enregistrements vidéo capturant le visage du sujet pendant qu'il dit la vérité et qu'il ment dans un laboratoire avec une caméra Basler à 100 FPS. Le sous-ensemble Silesian5 contient 5 vidéos enregistrées de 5 participants.

3.3.3 Jeux de données basés sur les images

Cette dernière classe est réservée aux jeux de données de type images, qui offrent une variété d'images de région oculaire avec différentes conditions telles que la luminosité, la qualité, etc.

3.3.3.1 UFPR-Periocular dataset

Luiz et al. ont proposé un nouveau jeu de données UFPR-Periocular [51] comprenant des images collectées de la région oculaire de 1122 participants. Pour la collecte, les auteurs ont développé une application mobile qui a permis aux participants de collecter leurs images à l'aide de leurs smartphones. Le jeu de données se compose d'images prises à partir de 1 122 sujets, pour un total de 16 830 images des deux yeux. Les images ont été prises à partir de 196 appareils mobiles différents, les cinq modèles les plus utilisés sont Apple iPhone 8, Apple iPhone 9, Xiaomi Mi 8 Lite, Apple iPhone 7 et Samsung Galaxy J7 Prime. Les images ont été prises lors de 3 sessions, avec 5 images par session pour un intervalle minimum de 8 heures entre les sessions.

Ensuite, les auteurs ont découpé les images des régions oculaires en 2, une image par oeil, en attribuant une classe unique à chaque oeil. Ils ont manuellement annoté les coins des yeux avec quatre points sur chaque image (les coins intérieurs et extérieurs de l'oeil). En utilisant le point central de chaque oeil, l'image a été tournée et mise à l'échelle pour normaliser la position de l'oeil avec une résolution de (512 * 256) pixels. Les images ont ensuite été divisées en deux parties à (256 * 256 pixels) pour séparer l'oeil gauche et droit créant un nombre total de 33 660 images à partir de 2 244 classes. La variation intra et inter-classe dans ce jeu de données est due à la lumière, l'occlusion, la réflexion spéculaire, le flou, le flou de mouvement, le port de lunettes, un mauvais alignement, une fixation, un maquillage et des expressions faciales.

3.3.4 Discussion

Dans cette étude de cas, nous avons analysé trois types de jeux de données existants. Ces jeux de données ont été collectés soit pour l'authentification biométrique basée sur les mouvements oculaires, ou pour des études sur la détection des clignements.

Après avoir examiné ces jeux de données, nous avons remarqué plusieurs limitations.

Les jeux de données basés sur les signaux [71, 33, 37] ne fournissent pas le type de signal que nous recherchons dans notre étude. Par exemple, le jeu de données Gasebase[37] ne fournit pas la distance verticale (entre le point le plus haut de la paupière supérieure et le point le plus bas de la paupière inférieure) et la distance horizontale (les points les plus extrêmes à droite et à gauche de l'œil) utilisée dans le travail de Goussem et Djallil [31].

En ce qui concerne le jeux de données de type image UFPR-Periocular [51], il contient un très grand ensemble d'images de la région oculaire des yeux des participants. Cependant, il ne

comprend pas un ensemble d'images successives représentant un clignement , ce qui est intéressant pour notre étude.

Enfin, Pour les jeux de données basés sur les vidéos [28, 67, 25, 20, 61], nous remarquons qu'ils ne se concentrent pas sur la région d'intérêt (les yeux) mais elles fournissent une vue frontale du visage. Ceci peut rendre la détection des clignements et des mouvements des yeux plus difficile et moins précise en raison de la distance entre la caméra et les yeux. De plus, la qualité des enregistrements est souvent faible comme c'est le cas pour ZJU [28]. Aussi, soit la taille des jeux de données est limitée (Silesian5 [67], Talking Face [25]), soit les jeux de données ne contiennent pas suffisamment de vidéos pour chaque participant (EyeLink8 [20], Researcher's Night [61]). Le jeu de données de Goussem et Djallil pourrait constituer une exception à ces limitations car il propose plus de vidéos par participant, avec une taille moyenne (96 vidéos). Cependant, ce jeu de données n'est pas accessible publiquement.

Après avoir examiné ces jeux de données, nous avons constaté qu'ils présentent souvent des limitations en termes de qualité des enregistrements, de nombre de vidéos par participant et de distance entre la caméra et les yeux. C'est pourquoi nous avons jugés nécessaire de collecter notre propre jeu de données.

3.4 Présentation de EyeReg

Dans cette partie, nous visons à résoudre les limitations discutés dans la section précédente, en présentant un nouveau jeu de données que nous avons appelé EyeReg à grande échelle conçu pour les études de clignements et de mouvements des yeux. Notre jeu de données vise à fournir aux chercheurs des enregistrements de haute qualité en se concentrant sur la région des yeux. Ceci facilite la détection précise des clignements et des mouvements des yeux. En présentant ce nouveau jeu de données, nous espérons contribuer à l'avancement de la recherche dans le domaine de la biométrie oculaire.

Dans ce qui suit, nous décrirons le protocole de collecte de notre premier jeu de données comportant des vidéos ainsi que les participants et le matériel utilisé. Ensuite nous décrirons comment nous avons déduit un deuxième jeu de données de type images.

3.4.1 Participants

Notre jeu de données regroupe 22 participants de différentes tranches d'âge (entre 16 et 42 ans) et couleur des yeux (deux en total vert et marron). Les participants ont signé un formulaire (Voir Figure 3.5) de consentement éclairé autorisant la collecte de vidéos qui montrent la partie supérieurs de leur visage.

Id : . . . . . .

CNI/PC : . . . . . .

Je, soussigné(e), ... déclare autoriser la collecte et la publication de vidéos montrant la partie supérieure de mon visage sans révéler mon identité par NEMEUR Chemseddine et AMEUR Zineddine dans le cadre de la construction d'un dataset plus large qui sera utilisé dans leur mémoire de Master 2 (2022/2023) et futures recherches scientifiques.

☐ Homme ☐ Femme Âge : . . . . . .

☐ Allergies : . . . . . .

☐ Porte des lunettes

☐ Autre maladie : . . . . . .

Fait à..... Le ...../...../.....

Signature

FIGURE 3.5 – Formulaire de consentement.

3.4.2 Matériels utilisés

Les vidéos ont été enregistrée à l'aide de deux caméras de téléphone d'une résolution de 13MP et des ouverture de f/1,9 et f/2,0 respectivement. Le téléphone est fixé à hauteur des yeux en utilisant un support de 50 centimètres tel qu'illustré sur la Figure 3.6.

FIGURE 3.6 – Protocole de collection.

3.4.3 Protocole de collecte du jeu de données vidéos

La collecte des vidéos de notre jeu de données s'est effectuée dans un environnement intérieur, avec une lumière naturelle dans la journée et lumière artificielle durant la nuit. Nous avons effectué 10 sessions d'enregistrement par participant. Pour chaque session, le participant est invité à regarder une vidéo sur un écran d'ordinateur disposé en face de lui. Nous avons choisit des vidéos de genres variés, comprenant des scènes d'action, des scènes de joie et des scènes humoristiques afin de capturer au mieux les différents comportements des yeux.

Il a été demandé aux participants d'interagir de manière spontanée et naturelle avec les vidéos regardées sur l'écran d'ordinateur, pendant qu'on lance l'enregistrement à l'aide de nos caméras pour une durée moyenne de 60 secondes. La Figure 3.9 montre des aperçus de vidéos résultantes.

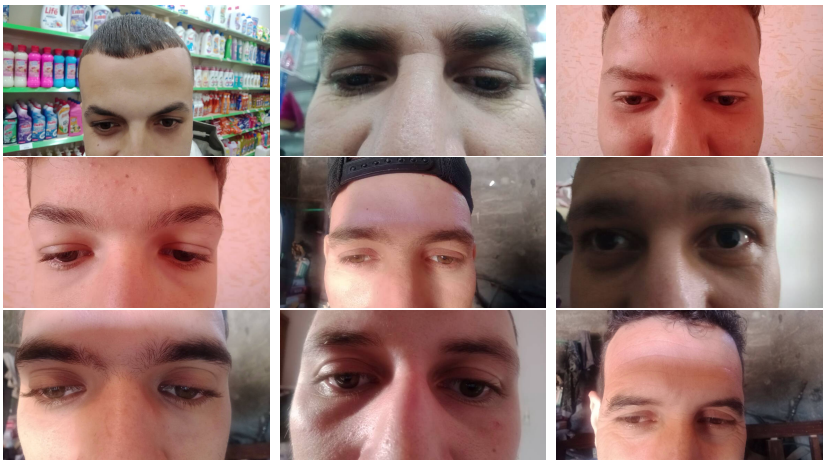


FIGURE 3.7 – Aperçus de notre jeu de données vidéos.

3.4.4 Propriétés de EyeReg version vidéo

Notre jeu de données possède au total 220 vidéos de la partie supérieure du visage de 22 participants, avec 10 enregistrements par participant. Ces vidéos ont été capturées en intérieur à une résolution de 1920*1080 pixels à 30 FPS. Les niveaux de lumière varient de forts à faibles. En outre, il existe des vidéos présentant des perturbations temporaires dues à des détournements de visages par rapport à la caméra, ou au flou de mise au point en raison de la courte distance entre la caméra et le participant.

Les vidéos sont structurées selon le format suivant : "user{i}_cap{j}.mp4". Cela signifie que chaque vidéo est identifiée par l'identifiant du participant "i" et le numéro de la vidéo "j".

3.4.5 EyeReg version image

À partir des enregistrements vidéos collectés, nous avons créé deux sous-ensembles de jeux de données de type images.

Pour chacune des 220 vidéos, nous avons extrait un intervalle de 10 secondes où le participant cligne ces yeux au moins une fois, afin de capturer l'état des yeux ainsi que les moments de clignement. Ensuite, deux types d'images sont extraites à partir de ces clips : des images combinant les deux yeux et des images des deux yeux séparés. Chaque type forme un ensemble de données distinct.

Le premier ensemble de jeu de données contient 69 505 images des deux yeux ensembles, avec

une résolution de 616 × 408 pixels. Nous avons extrait en moyenne 300 images par vidéo, ce qui donne en moyenne un total de 3000 images par participant.

Les images sont organisées en fonction de l'identifiant de l'utilisateur. Chaque utilisateur dispose d'un dossier distinct portant son identifiant, et les images appartenant à cet utilisateur sont stockées à l'intérieur de ce dossier.

Les images elles-mêmes sont nommées selon le format "eye_sequencenumber". Ici, "sequencenumber" représente une valeur numérique comprise entre 1 et 9999, indiquant le numéro de séquence de l'image.

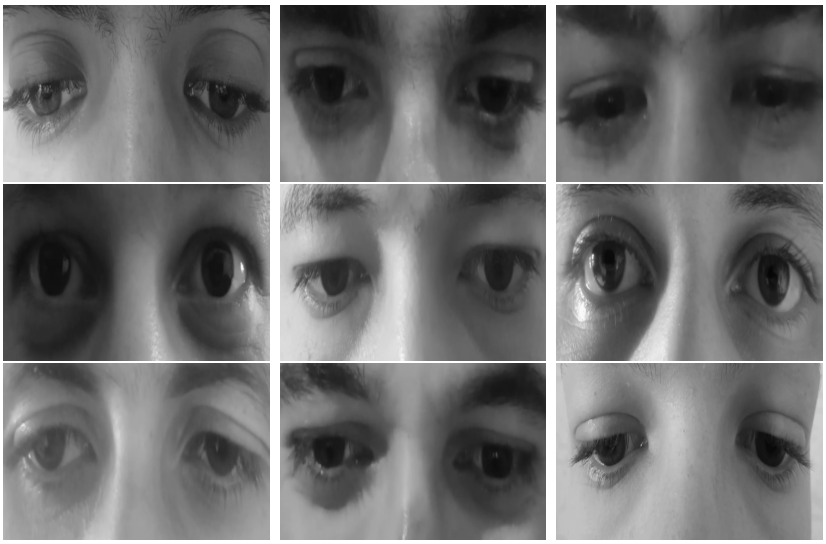


FIGURE 3.8 – Aperçus du premier sous-ensemble de notre jeu de données images.

Le deuxième ensemble comprend 138 889 images des yeux droits et gauches séparés, avec une résolution de 192 × 168 pixels. En effet, nous avons extrait en moyenne 600 images par vidéo (300 par œil), ce qui équivaut à un total de 60000 images par participant en moyenne, comprenant les images pour chaque œil.

Pour cet ensemble, la structure est similaire à celle précédemment décrite, à l'exception de la structure des images à l'intérieur de chaque dossier. Les noms des images sont structurés comme suit : "left_eye_sequencenumber" et "right_eye_sequencenumber".

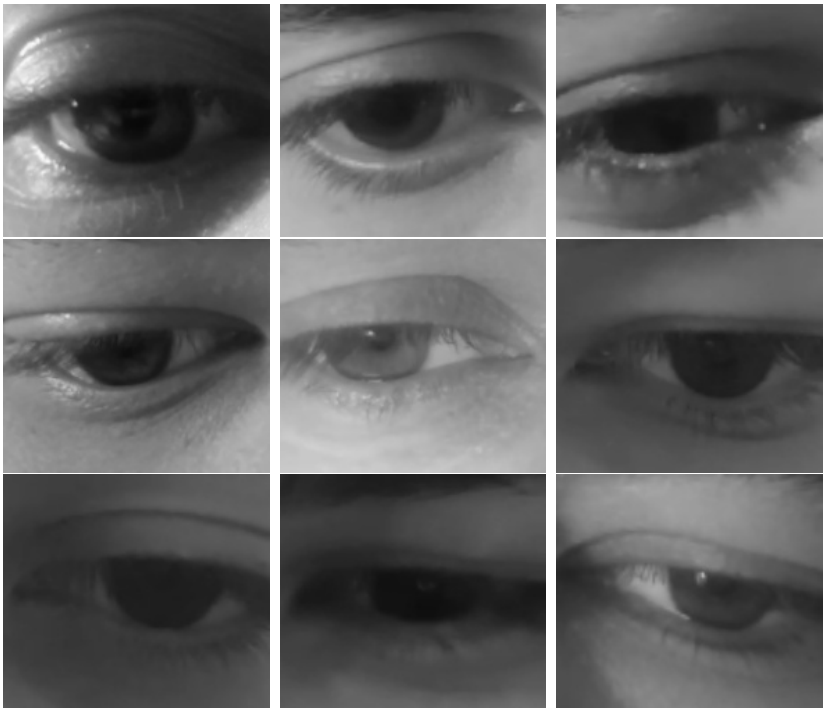


FIGURE 3.9 – Aperçus du deuxième sous-ensemble de notre jeu de données images.

3.4.6 Comparaison

Le tableau 3.1 résume les différences entre les jeux de données vidéos et images étudiés et ceux que nous proposons.

Jeux de données	Participants	Taille	Durée	Environnement	Résolution	Licence
Researcher's Night [61]	107	107 vidéos	/	Intérieur	640*480	À la demande
ZJU Dataset [28]	20	80 vidéos	/	Intérieur	320*240	À la demande
EyeLink8 Dataset [20]	4	8 vidéos	/	Intérieur	640*480	Publique
Talking Face [25]	1	4 vidéos	50s	Intérieur	720*576	Publique
Goussem et Djallil [31]	16	96 vidéos	20s	/	/	Inaccessible
Silesian5 Dataset [67]	5	5 vidéos	/	Laboratoire	640*480	À la demande
EyeReg vidéo	22	220 vidéos	60s	Intérieur	1920×1080	Publique
UFPR-Periocular1 [51]	1122	16 830 images	/	Variable	512*256	Publique
UFPR-Periocular2 [51]	1122	33 660 images	/	Variable	256*256	Publique
EyeReg Images1	22	69 505 images	/	Intérieur	616*408	Publique
EyeReg Images2	22	138 889 images	/	Intérieur	192*168	Publique

TABLEAU 3.1 – Tableau comparatif.

Nous constatons que notre jeu de données vidéos présente le plus grand nombre de vidéos

(220) ainsi que la plus haute résolution (1920×1080). En plus, nos vidéos sont plus longues (60 secondes) que celles des autres jeux de données.

Notre jeu de données d'images 1 & 2 est plus étendu en termes du nombre d'images, malgré le nombre restreint de participants par rapport aux jeux de données UFPR-Periocular 1 et 2. En ce qui concerne la résolution, nous pouvons dire que notre jeu de données d'images des yeux combinées a la résolution la plus élevée, le contraire pour le jeu de données d'images séparées ou UFPR-Periocular a la plus haute résolution.

Un avantage de nos jeux de données d'images est la présence d'images de clignement, ce qui diffère des jeux de données UFPR-Periocular. Cela peut être bénéfique lors de l'élaboration et du test de protocoles d'authentification, comme dans notre étude.

Certains jeux de données sont disponibles à la demande, seul le jeu de données de Goussem et Djallil n'est pas accessible. En revanche, nos jeux de données sont rendus publiques pour les chercheurs.

3.5 Conclusion

Dans ce chapitre, nous avons étudié plusieurs jeux de données utilisés dans le domaine de l'authentification biométrique basée sur les mouvements des yeux. Nous avons classé ces jeux de données en trois catégories selon le type de la donnée : des signaux, des vidéos et des images. En outre, nous avons discuté les limitations de chacune de ces catégories de jeux de données. Pour contribuer dans ce domaine, nous avons proposé un nouveau jeu de données pour l'authentification biométrique basée sur les mouvements des yeux. Notre jeu de données se compose de vidéos capturant la région supérieure du visage de 22 participants. De ces vidéos, nous avons extrait des images, créant ainsi deux sous-ensembles d'un deuxième jeu de données : l'un contenant des images des deux yeux ensemble et un autre sous-ensemble contenant des images des yeux séparés.

Le chapitre suivant présente les expérimentations réalisées sur le protocole d'authentification biométrique proposé par Gousem et Djallil [31]. Nous examinerons les résultats obtenus en expérimentant différentes techniques pour améliorer le protocole.

Chapitre 4

Expérimentation et résultats

4.1 Introduction

Dans ce dernier chapitre, nous conduisons une série d'expériments visant à explorer les possibilités d'amélioration du protocole d'authentification de Goussem et Djallil [31], et présentons les résultats obtenus. Nous commençons par une description de l'environnement expérimental, comprenant le matériel utilisé ainsi que les outils et les bibliothèques employés. Ensuite, nous effectuerons un processus d'authentification en utilisant notre jeu de données et leur protocole comme première tentative. Les expériences comprendront les objectifs suivants : l'authentification en utilisant un nombre variable de caractéristiques et l'exploration de la corrélation entre les jeux de données de l'œil droit et de l'œil gauche. Par la suite, nous proposerons un ensemble supplémentaire de caractéristiques pour améliorer le protocole. En ce qui concerne les jeux de données d'images, nous entraînerons et testerons deux modèles de Réseaux de Neurones Convolutionnels (CNN).

4.2 Environnement expérimentale

Nous présentons, dans ce qui suit, les matériaux et outils utilisés dans l'étude.

4.2.1 Matériel

Nous avons utilisé un ordinateur portable Asus pour cette expérience. Ses caractéristiques sont détaillées sur le tableau 4.1.

Matériel	Caractéristiques
Asus VivoBook Laptop	Processeur AMD Rysen 5 5600 Series . Ram 8GB DDR4 512GB 3.0 SSD NVIDIA GEFORCE RTX 4GB

Tableau 4.1 – Caractéristiques de l'ordinateur portable

4.2.2 Outils et bibliothèques

Pour la partie logicielle, nous avons utilisé plusieurs bibliothèque de Python. Nous les présenterons dans ce qui suit.

4.2.2.1 Python

Python est un langage de programmation de haut niveau connu pour sa simplicité et sa lisibilité. Il a été créé par Guido van Rossum et initialement publié en 1991. Python prend en charge différents styles de programmation et dispose d'une vaste collection d'outils et de bibliothèques préconstruits. Il est largement utilisé dans le domaine de l'apprentissage automatique et de la science des données [79].

4.2.2.2 Numpy

NumPy est une bibliothèque fondamentale pour le calcul scientifique en Python. Elle offre des structures de données puissantes, telles que des tableaux multidimensionnels, ainsi qu'une vaste bibliothèque de fonctions mathématiques pour effectuer des calculs efficaces sur ces tableaux. NumPy est largement utilisé pour les calculs numériques, la manipulation des données et les opérations d'algèbre linéaire [76].

4.2.2.3 OpenCV

OpenCV est une bibliothèque populaire de vision par ordinateur en open source qui offre une large gamme d'outils et d'algorithmes pour l'analyse d'images et de vidéos, notamment la détection d'objets, la reconnaissance d'images et l'étalonnage de caméras. OpenCV propose une interface en C++, ainsi que des liens pour différents langages de programmation, dont Python. OpenCV est largement utilisé dans la recherche en vision par ordinateur et les applications associées [77].

4.2.2.4 Pandas

Pandas est une puissante bibliothèque de manipulation et d'analyse de données en Python. Elle fournit des structures de données faciles à utiliser, telles que les DataFrames, qui permettent une manipulation efficace des données structurées. Pandas offre des fonctionnalités de nettoyage des données, de fusion, de remodelage, de découpage, de filtrage et d'analyse statistique. Elle est largement utilisée dans les domaines de la science des données et de l'analyse de données [78].

4.2.2.5 Keras

Keras est une bibliothèque de deep learning de haut niveau qui offre une interface utilisateur conviviale et intuitive pour la construction et l'entraînement de réseaux neuronaux. Elle est construite sur d'autres frameworks de deep learning tels que TensorFlow et peut utiliser différentes technologies sous-jacentes comme moteur de calcul. Keras simplifie le processus de conception et d'entraînement des modèles de deep learning, ce qui la rend accessible tant aux débutants qu'aux praticiens expérimentés [74].

4.3 Résultats obtenus avec notre jeu de données sur le protocole de Goussem et Djallil

Nous avons implémenté le protocole de Goussem Ayoub et Djallil Massinissa [31] présenté dans la Section 2.3.2.2 avec notre jeu de données. Comme discuté précédemment, ils ont utilisé le classifieur KNN avec différentes valeurs de k, ainsi que des valeurs de seuils différentes (dans ce cas, la distance entre la cible et les autres échantillons du jeu de données). Dans ce qui suit , on a choisit un nombre de voisins qui varie de 1 à 7, et les seuils utilisés sont 4,4.1,4.01. Les meilleurs résultats de notre étude ont été obtenus avec ces paramètres. Ces résultats, présentés dans les tables 4.2, 4.4 et 4.3, sont exprimés en termes de taux de faux rejets (FRR – False Rejection Rate), de taux de fausses acceptations (FAR – False Acceptation Rate) et du taux d'égalité d'erreurs (EER - Equal Error Rate).

Seuil = 4

k	/	1	2	3	4	5	6	7
Œil Droit	**FRR (%)**	33.33	33.33	33.33	33.33	33.33	33.33	33.33
	FAR (%)	14.29	14.29	14.29	14.29	14.29	14.29	14.29
	TEE (%)	23.80	23.80	23.80	23.80	23.80	23.80	23.80
Œil Gauche	**FRR (%)**	66.66	66.66	66.66	66.66	66.66	66.66	66.66
	FAR (%)	14.29	14.29	14.29	14.29	14.29	14.29	14.29
	TEE (%)	40.47	40.47	40.47	40.47	40.47	40.47	40.47

TABLEAU 4.2 – Résultats de l'authentification avec le protocole de Goussem et al. [31] pour un seuil de 4.

Seuil = 4.01

k	/	1	2	3	4	5	6	7
Œil Droit	**FRR (%)**	33.33	33.33	33.33	33.33	33.33	33.33	33.33
	FAR (%)	14.29	14.29	14.29	14.29	14.29	14.29	14.29
	TEE (%)	23.80	23.80	23.80	23.80	23.80	23.80	23.80
Œil Gauche	**FRR (%)**	66.00	66.00	66.00	66.00	66.00	66.00	66.00
	FAR (%)	14.29	14.29	14.29	14.29	14.29	14.29	14.29
	TEE (%)	37.14	37.14	37.14	37.14	37.14	37.14	37.14

TABLEAU 4.3 – Résultats de l'authentification avec le protocole de Goussem et al. [31] un seuil de 4.01.

Seuil = 4.1

k	/	1	2	3	4	5	6	7
Œil Droit	**FRR (%)**	33.33	33.33	33.33	33.33	33.33	33.33	33.33
	FAR (%)	28.57	28.57	28.57	28.57	28.57	28.57	28.57
	TEE (%)	30.95	30.95	30.95	30.95	30.95	30.95	30.95
Œil Gauche	**FRR (%)**	53.30	53.30	53.30	53.30	53.30	53.30	53.3
	FAR (%)	28.57	28.57	28.57	28.57	28.57	28.57	28.57
	TEE (%)	40.95	40.95	40.95	40.95	40.95	40.95	40.95

TABLEAU 4.4 – Résultats de l'authentification avec le protocole de Goussem et al. [31] un seuil de 4.1.

Les tableaux montrent que les meilleurs résultats ont été observés avec une valeur de seuil égale à 4, en réalisant un taux FRR = 33.33% et un taux FAR = 14.29% avec l'oeil de droit.

Nous rappelons que les meilleurs résultats obtenus par Goussem et Djallil [31] avec leur jeu de données étaient un taux de FRR = 50% et FAR = 20% avec un seuil égal à 5 pour l'oeil gauche.

Nous pouvons constater que les résultats de l'authentification ne sont pas affectés par le nombre de voisins k, seul le seuil a un impact. Par conséquent, nous utiliserons k = 1 pour le reste des expérimentations.

4.4 Expérimentations réalisées

Dans cette section, nous présenterons une série d'expérimentations réalisées selon le plan suivant :

— Nous étudions l'effet de nombre de caractéristiques sur le protocole de Goussem et Djallil [31].
— Nous étudions la corrélation entre les caractéristique des deux yeux dans notre jeu de données.
— Nous proposons l'utilisation de nouvelles caractéristiques.
— Nous proposons un modèle de réseaux de neurones convolutifs pour chacun des sous-ensembles de jeux de données d'images, et discutions des résultats obtenus.

4.4.1 L'effet de nombre de caractéristiques sur l'authentification

Nous étudions l'authentification en utilisant différents nombres de caractéristiques biométriques, en commençant par 2 caractéristiques jusqu'à 5 caractéristiques. Dans chaque cas, seule la valeur du seuil qui donne les meilleurs résultats seront présentés. L'objectif est de comprendre l'impact du nombre de caractéristiques sur le protocole d'authentification.

4.4.1.1 Authentification avec 2 caractéristiques

Les caractéristiques choisies sont les suivantes : le quotient moyen et le FFT maximum. Le tableau 4.5 présente les meilleurs résultats qui ont été obtenus avec un seuil de 0.038.

k	/	1
Œil Droit	**FRR** (%)	80.00
	FAR (%)	42.28
	TEE (%)	**61.14**
Œil Gauche	**FRR** (%)	66.00
	FAR (%)	42.28
	TEE (%)	51.14

TABLEAU 4.5 – Résultats de l'authentification avec 2 caractéristiques et un seuil de 0.038.

4.4.1.2 Authentification avec 3 caractéristiques

Les caractéristiques choisies sont le quotient minimum, le quotient maximum et le quotient moyen. La tableau 4.6 présente les meilleurs résultats obtenus avec un seuil de 0.99.

k	/	1
Œil Droit	**FRR** (%)	60.00
	FAR (%)	28.57
	TEE (%)	**44.28**
Œil Gauche	**FRR** (%)	26.67
	FAR (%)	57.14
	TEE (%)	41.90

TABLEAU 4.6 – Résultats de l'authentification avec 3 caractéristiques et un seuil de 0.99.

4.4.1.3 Authentification avec 4 caractéristiques

Les caractéristiques choisies sont les suivantes : le quotient minimum, le quotient maximum, le quotient moyen et le FFT maximum. Le tableau 4.7 présente les meilleurs résultats avec un seuil de 1.1.

k	/	1
Œil Droit	**FRR (%)**	66.66
	FAR (%)	28.57
	TEE (%)	**47.61**
Œil Gauche	**FRR (%)**	40.00
	FAR (%)	57.14
	TEE (%)	48.57

TABLEAU 4.7 – Résultats de l'authentification avec 4 caractéristiques et un seuil de 1.1.

4.4.1.4 Authentification avec 5 caractéristiques

Les caractéristiques choisies sont les suivantes : le quotient minimum, le quotient maximum, le quotient moyen, le FFT maximum et le relatif minimum. Le tableau 4.8 présente les meilleurs résultats obtenus avec un seuil de 2.7.

k	/	1
Œil Droit	**FRR (%)**	53.33
	FAR (%)	28.57
	TEE (%)	**40.95**
Œil Gauche	**FRR (%)**	46.66
	FAR (%)	28.57
	TEE (%)	37.61

TABLEAU 4.8 – Résultats de l'authentification avec 5 caractéristiques et un seuil de 2.7.

4.4.1.5 Discussion de résultats

Nous avons examiné l'impact du nombre de caractéristiques sur les résultats de l'authentification. L'étude a pris en compte différents nombres de caractéristiques, allant de 2 à 5, et a sélectionné la valeur de seuil qui donnait les meilleurs résultats dans chaque cas. L'objectif était d'évaluer l'influence du nombre de caractéristiques sur le protocole d'authentification.

Les résultats indiquent que l'ajout de plus de caractéristiques a un impact positif sur les résultats de l'authentification, en termes de taux de faux rejet (FRR) et de taux de faux acceptation (FAR). En augmentant le nombre de caractéristiques, le système peut efficacement différencier les utilisateurs légitimes des imposteurs, ce qui se traduit par des taux de FRR et de FAR plus bas. On peut conclure que l'augmentation du nombre de caractéristiques dans le processus d'authenti-

fication conduit à de meilleurs résultats en termes de FRR et de FAR, ce qui améliore globalement les performances et l'efficacité du protocole d'authentification.

4.4.2 Corrélation dans notre jeu de données

La corrélation est une mesure statistique qui évalue le degré de relation linéaire entre deux variables [32]. Le coefficient de corrélation est un nombre compris entre -1 et 1, où 0 indique l'absence de corrélation. Un coefficient de corrélation positif indique une corrélation positive, ce qui signifie que les variables évoluent généralement dans la même direction. Un coefficient de corrélation négatif indique une corrélation négative, ce qui signifie que les variables évoluent généralement dans des directions opposées.

Nous calculons le coefficient de corrélation des caractéristique issues de notre jeu de données de l'œil gauche et celui de l'œil droit afin de déterminer la relation linéaire entre ces deux ensembles. Sachant que les deux jeux de données ont les mêmes caractéristiques, le coefficient de corrélation nous permettra de quantifier la direction de cette relation.

Les résultats de la fonction de corrélation "corrwith()" de la bibliothèque pandas sont présentés dans le tableau 4.9 .

Caractéristique	**Coefficient de corrélation**
Quotient minimum	0.963865
Quotient maximum	0.806778
Quotient moyen	0.932717
FFT maximum	0.932695
Relatif minimum	0.949679
Relatif maximum	0.949472

Tableau 4.9 – Coefficients de corrélation de chaque Caractéristique.

4.4.2.1 Discussion des résultats

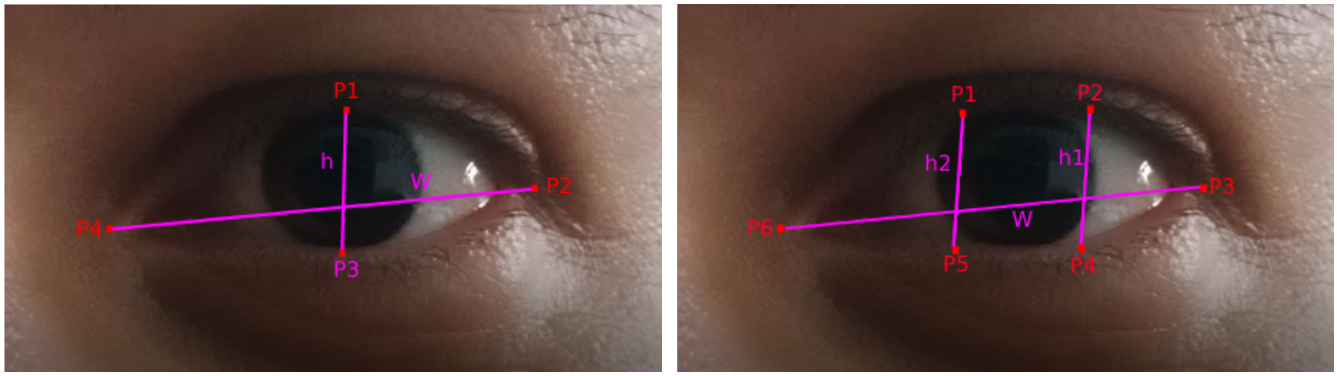
Nous avons étudié la corrélation entre les jeux de données de l'œil gauche et de l'œil droit afin de déduire la possibilité de se limiter à un seul ensemble pour le processus d'authentification.

D'après les résultats présentés dans le tableau 4.9, nous pouvons conclure qu'il existe une forte corrélation positive entre les jeux de données de l'œil gauche et de l'œil droit pour toutes les caractéristiques analysées. Les coefficients de corrélation sont proches de 1, ce qui indique une forte relation linéaire. Par conséquent, il est raisonnable de supposer que l'utilisation d'un seul oeil pour l'authentification avec les deux yeux peut être efficace, car il existe une forte relation entre les caractéristiques des deux yeux.

4.4.3 Augmentation du nombre des points de repère oculaires

Nous étudions à présent les résultats de l'authentification en augmentant le nombre de points de repère oculaires. Dans leur protocole original [31], Goussem et Djallil ont utilisé 4 points de repère oculaires indiqués sur la Figure 4.1a. Nous essayerons d'utiliser le 6 points de repère oculaires illustrés sur la figure 4.1b. Dans ce cas, nous calculons le rapport de signal à l'aide de l'équation 4.1.

$$\text{Rapport} = \frac{h1 + h2}{2 * W} \tag{4.1}$$



(a) Les 4 points de repère oculaires. (b) Les 6 points de repère oculaires.

FIGURE 4.1 – Points de repère oculaires.

Les résultats de l'authentification avec 6 points de repères un seuil de 3.75 sont présentés sur le Tableau 4.10.

k	/	1
Œil Droit	**FRR (%)**	66.66
	FAR (%)	14.29
	TEE (%)	**40.47**
Œil Gauche	**FRR (%)**	66.00
	FAR (%)	28.57
	TEE (%)	44.28

TABLEAU 4.10 – Résultats de l'authentification avec 6 points de repère oculaires.

4.4.3.1 Discussions

Nous remarquons que résultats obtenus ne sont pas meilleurs. En effet, avec 4 points de repère nous avons obtenus un meilleur taux FRR de 33.33% pour l'oeil droite contre un taux FRR = 66.66%.

L'augmentation du nombre de points de repère oculaires n'a pas d'effet positif sur le protocle d'authentification .

4.4.4 Proposition de nouvelles caractéristiques pour l'authentification

Étant donné que l'effet de l'augmentation des caractéristiques a été démontré comme positif dans les résultats de l'authentification (Voir Section 4.4.1) nous avons ajouté 03 caractéristiques au protocole :

— **Hauteur** : la hauteur de l'œil.
— **Largeur** : la largeur de l'œil.
— **Moyenne** : le rapport de la hauteur sur la largeur de l'œil.

Ces caractéristiques sont statiques pour chaque utilisateur et sont illustrées sur la Figure 4.2.

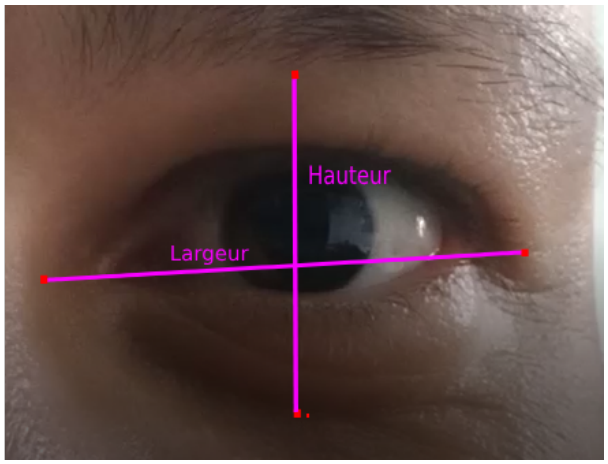


FIGURE 4.2 – Les caractéristiques Proposées.

Les résultats de l'authentification biométrique en utilisant les neuf caractéristiques pour un seuil de 10 sont représentés dans le tableau 4.11.

k	/	1
Œil Droit	**FRR (%)**	26.67
	FAR (%)	14.29
	TEE (%)	**20.47**
Œil Gauche	**FRR (%)**	33.33
	FAR (%)	28.57
	TEE (%)	30.95

TABLEAU 4.11 – Résultats de l'authentification avec neuf caractéristiques

4.4.4.1 Discussion des résultats

Le tableau 4.11 montre que les résultats sont relativement meilleurs que ceux du protocole original en termes de taux de faux rejet , avec 26.67 % comparé à 33,33 % dans le protocole original pour l'œil droit, et 33,33 % comparé à 60 % pour l'œil gauche.

4.4.5 Entrainement avec les réseaux de neurones convolutifs

Nous avons entraîné deux modèles de réseaux de neurones convolutifs (CNN) pour effectuer l'authentification à l'aide de deux sous-ensembles jeux de données d'images que nous avons extraites du jeu de données vidéos.

4.4.5.1 Modèle CNN pour le jeux de données image des deux yeux

Ce modèle est composé de deux ensembles de couches de convolution et de max pooling suivi par deux couches entièrement connectées.

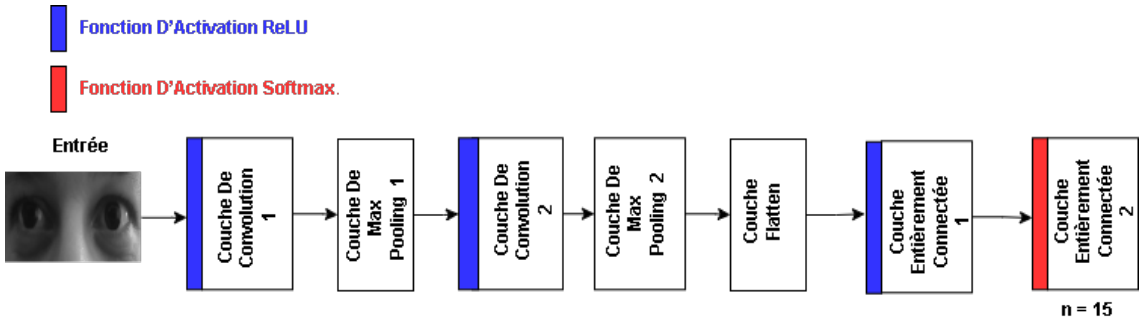


FIGURE 4.3 – Modèle CNN pour le jeux de données image des deux yeux.

La Figure 4.3 illustre l'architecture de notre modèle qui est décrit comme suit :

1. La 1ère couche de convolution applique 16 filtres, chacun de taille 3x3 pixels, avec la fonction d'activation ReLU. Elle prend en entrée une image avec une dimension de type (hauteur, largeur, nombre de canaux).
2. La couche de max pooling qui suit réduit les dimensions spatiales de l'entrée en prenant la valeur maximale dans chaque région de taille 2x2.
3. La deuxième couche de convolution applique 32 filtres de même taille que précédemment.
4. Une autre couche max pooling est appliquée pour réduire encore les dimensions.
5. Une couche Flatten convertit les cartes de caractéristiques 2D en un vecteur 1D, préparant les données pour les couches entièrement connectées ultérieures.
6. La première couche entièrement connectée est composée de 64 neurones et utilise la fonction d'activation ReLU.
7. Enfin, la dernière couche entièrement connectée produit les probabilités de classe en utilisant la fonction d'activation softmax, avec un nombre de neurones égal au nombre de classes dans la tâche de classification (15 classes).

Lors de l'entraînement, le modèle a été alimenté par des lots de 40 images avec un taux d'apprentissage de 0,001. Nous avons dû réduire la résolution des images à 154*102 pixels en raison de la limitation de la mémoire disponible pour gérer l'entraînement avec des grandes tailles de données.

Les résultats d'authentification avec un seuil de 0,999 sont les suivants :

— Taux de Faux Rejet (FRR) : 26,67 %. 4 tentatives d'authentification légitime sur 15 ont été incorrectement rejetées.
— Taux de Fausse Acceptation (FAR) : 14,29 %. 1 tentative d'authentification d'un imposteur sur 7 ont été incorrectement acceptées.

4.4.5.2 Modèle CNN pour le jeux de données image des yeux séparés

Ce modèle a été entraîné sur les images des yeux séparés.

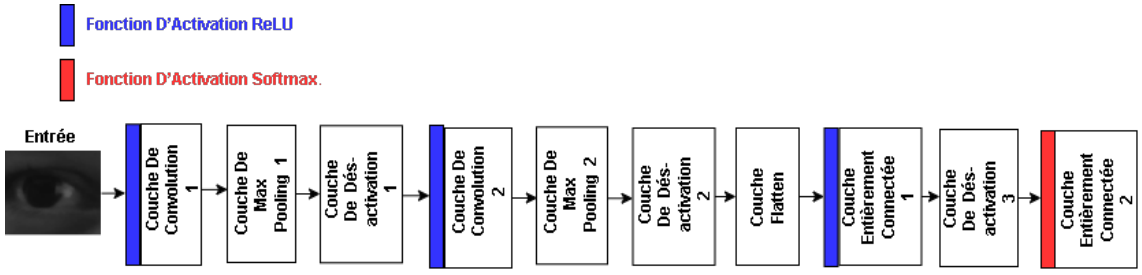


FIGURE 4.4 – Modèle CNN pour le jeux de données image des yeux séparés.

La Figure 4.4 illustre l'architecture de notre modèle qui est décrit comme suit :

1. Le modèle commence par une couche de convolution en entrée avec 32 filtres de taille 3x3 et une activation ReLU. Les dimensions des images d'entrée sont réduites à (128*112).
2. Cette 1ère couche est suivie d'une 1ère couche de max pooling qui réduit les dimensions spatiales en prenant les valeurs maximales dans les régions de taille 2x2.
3. Pour éviter le surapprentissage, une couche de désactivation est ajoutée avec un taux de 25%.
4. Ensuite, une 2e couche de convolution suit juste après avec 64 filtres de taille 3x3 et une activation ReLU.
5. Une 2e couche de max pooling est appliquée encore une fois pour réduire davantage les dimensions spatiales.
6. Une 2e couche de désactivation est ajoutée avec un taux de 25%.
7. La couche Flatten est utilisée pour convertir les cartes de caractéristiques 2D en un vecteur 1D.
8. une 1ère couche entièrement connectée comprend 128 neurones avec une activation ReLU.
9. Une 3e couche de désactivation est ajoutée avec un taux de 50% pour améliorer la régularisation.
10. Enfin, une 2e couche entièrement connectée est utilisée pour produire les probabilités de classe en utilisant l'activation softmax. Le nombre de neurones dans cette couche correspond au nombre de classes dans la tâche de classification.

Les résultats d'authentification avec un seuil de 0,92 sont les suivants :

— Œil droite : taux de fausse acceptation est 14.29%, taux de faux rejet est 26.67%
— Œil gauche : taux de fausse acceptation est 28.57 taux de faux rejet est 33.33%

4.4.5.3 Discussion

L'architecture CNN entraînée avec des images des deux yeux combinés a démontré de meilleurs résultats par rapport à l'architecture entraînée avec les deux yeux séparés. Pour le modèle des deux yeux séparés, les meilleurs résultats ont été observés dans l'œil droit avec un taux de fausse acceptation de seulement 14,29 % et un taux de faux rejet de 26,67 %, des résultats similaires à ceux du modèle des deux yeux combinés. Cependant, l'œil gauche a donné des résultats moins précis avec un taux de fausse acceptation de 28,57% et un taux de faux rejet de 33,33 %. Cela pourrait être dû à un surajustement. En général, l'utilisation du premier modèle (les deux yeux combinés) s'est avérée plus précise que le deuxième modèle.

Le modèle KNN utilisé dans le protocole de Goussem et al. a montré des résultats similaires aux CNN en termes de FAR avec une valeur de 14,29% et un FRR de 33,33% en utilisant notre jeu de données. Cependant, en intégrant des caractéristiques supplémentaires dans le protocole, nous avons pu améliorer les performances de knn, atteignant un taux de FRR de 26,67% ,le taux obtenue à l'aide du modèle CNN.

4.5 Conclusion

Dans ce chapitre, nous avons mené une série d'expériences visant à améliorer le protocole d'authentification de Djallil et Goussem [31]. Nous avons démontré qu'un jeu de données plus large et plus de caractéristiques permet d'améliorer les taux FRR et FAR du protocole d'authentification. Nous avons aussi démontrer la possibilité d'utiliser un seul œil au lieu de deux pour effectuer le processus d'authentification. Nous avons aussi proposé d'ajouter plus de caractéristiques pour l'authentification. Par conséquent, il est juste de dire que l'utilisation des résultats du modèle proposé de 9 caractéristiques de l'œil droit constitue une amélioration pour le protocole d'authentification de Goussem et Djallil [31]. Les modèles CNN ont également donné des résultats similaires en termes de taux FAR, avec seulement 14,29%. Globalement, en utilisant notre jeu de données, l'authentification a donné lieu à 4 rejets erronés sur 15 tentatives et 1 acceptation erronée sur 7 tentatives.

Conclusion et perspectives

Dans ce projet de fin d'études, nous avons exploré le domaine de l'authentification biométrique basée sur le mouvement des yeux. Plus précisément, nous nous sommes basés sur le travail de Goussem et Djallil, qui ont proposé un système d'authentification basé sur les mouvements oculaires pour les applications de réalité virtuelle. Il nous a été nécessaire au départ d'étudier d'abord les méthodes d'apprentissage automatique et de l'apprentissage profond afin de pouvoir contribuer dans ce domaine.

Notre contribution dans ce domaine s'articule en plusieurs points. Nous avons d'abord présenté un état de l'art sur les protocoles d'authentification biométrique basée sur le mouvement oculaire. Nous avons ensuite présenté un 2e état de l'art sur les jeux de données existants pour l'authentification oculaire.

Nous avons aussi contribué avec la création de 2 jeux de données plus importants et spécialement conçus pour l'étude de l'authentification utilisant les données oculaires et la détection des clignements des yeux. Le 1er ensemble de données de type vidéo a été utilisé pour extraire le 2e jeu de données divisé en deux sous-ensembles d'images de la région oculaire : l'un regroupant les deux yeux et l'autre avec des images séparées pour chaque œil.

Enfin, nous avons mené des expériences en utilisant notre jeu de données de vidéos et le système d'authentification proposé par Goussem et Djallil. Pour les ensembles de données images, nous avons entraîné deux modèles CNN afin de tenter l'authentification. Les résultats de nos expériences ont montré des résultats prometteurs, nous avons obtenu un taux de fausses acceptation significativement faible : 14.29 % et un taux de faux rejet de 26.57 % en utilisant leur protocole. Avec les modèles d'apprentissage profond, nous avons pu obtenir de meilleurs résultats en termes de taux de faux rejet, atteignant seulement 26.67 %.

En tenant compte des limites du matériel utilisé lors de la collecte des jeux de données et durant les expériences menées, les résultats pourraient être plus précis en utilisant un matériel de meilleure qualité. Des études futures pourraient explorer l'amélioration de notre jeux de données en utilisant une caméra haute définition avec un taux de rafraîchissement plus élevé. Cela pourrait être bénéfique lors de l'extraction du signal des mouvements oculaires en fournissant à l'algorithme

un flux d'images par seconde plus élevé. De plus, l'utilisation d'un ordinateur plus puissant pourrait entraîner de meilleurs résultats pour les modèles d'apprentissage profond.

Bibliographie

[1] sklearn.model_selection.RandomizedSearchCV. https://scikitlearn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html#sklearn.model_selection. BSD License.

[2] deep neural networks. https://www.kdnuggets.com/2020/02/deep-neural-networks.html, 2020.

[3] http://hadoop.apache.org, (Consulté le 10 Janvier 2016).

[4] Kingma.D Adam.j. A method for stochastic optimization. *arXiv*, page 1412.6980, arXiv 2014.

[5] Adrian.W. Htc vive pro vs htc vive pro eye vs htc vive pro 2. "https://www.pocket-lint.com/fr", 8 avr 2023.

[6] Rania.k Ammar.M. *A comprehensive review on ensemble deep learning : Opportunities and challenges*. sciencedirect, Cairo, Egypt, February 2023.

[7] Andrade.C. Z scores, standard scores, and composite test scores explained. *Indian Journal of Psychological Medicine*, 43(6) :555–557, 2021.

[8] Alwyn.G Andrew.T, David.L. Biohashing : two factor authentication featuring fingerprint data and tokenised random number. 37(11) :2245–2255, November 2004.

[9] Avinash.N. Machine learning geek. May 12, 2023.

[10] Bastien.L. Reinforcement learning : qu'est-ce que l'apprentissage par renforcement ?, juin 2021. [Online].

[11] Huang.Y-Cohen.Il Benesty.J, Jingdong.C. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*, volume 2 of *Springer Topics in Signal Processing*. Springer, Berlin, Heidelberg, 2009.

[12] Berge.C. *Graphes et hypergraphes*. Dunod, Paris, 2nd edition, 1973.

[13] Adam.S Bernard.S, Heutte.L. Influence of hyperparameters on random forest accuracy. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 5519 :171–180, 2009.

[14] Isabelle.M-Vapnik Vladimir.N Bernhard.E, Guyon. A training algorithm for optimal margin classifiers. pages 144–152, 1992.

[15] Selman.B-Weinberger.K.Q Bjorck.J, Gomes.C. Understanding batch normalization. *Cornell University*, 2018.

[16] Breiman.L. *Random forests.Machine Learning.* 2001.

[17] Slimani.H-Faget.Z Brighen.A, Bellatreche.L. An economical query cost model in the cloud. In B. Hong et al., editor, *DASFAA Workshops 2013*, volume LNCS 7827, pages 16–30. Springer-Verlag, Berlin Heidelberg, 2013.

[18] Elmoataz.A Buyssens.P. Réseaux de neurones convolutionnels multi-échelle pour la classification cellulaire. *Clermont-Ferrand*, Jun 2016.

[19] Chuprina.R. Ai and machine learning in manufacturing : The complete guide. `https://spd.group/machine-learning/ai-and-ml-in-manufacturing-industry/`, (Consulté le 20 Nouvembre 2022).

[20] EyeLink8 Dataset. `https://www.blinkingmatters.com/files/upload/research/eyeblink8.zip`.

[21] Hugo.L Delalleau.O. Algorithme des k plus proches voisins. January 2007.

[22] Hacene Djadel. Utilisation des methodes support vecteur machine (svm) dans l'analyse des bases de donnes. *Universite de Bejaia*, 2006.

[23] Maglogiannis.I et al. *Emerging Artificial Intelligence Applications in Computer Engineering*, volume 186. IOS Press, Amsterdam, 1st edition, 2007.

[24] Rai.R et al. Machine learning in manufacturing and industry4.0 applications. *International Journal of Production Research.*, 59 :4773–4778, 2021.

[25] Talking face. `http://www-prima.inrialpes.fr/FGnet/data/01-TalkingFace/talking_face.html`.

[26] Steven Feltner. Single-factor authentication (sfa) vs. multi-factor authentication (mfa). *delinea.com*, 2023.

[27] Bouchouicha.M Ginoux.J Fnaiech.F Moreau.E Frizzi.S, Kaabi.R. Détection de la fumée et du feu par réseau de neurones convolutifs. *Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle,*, Jul 2017.

[28] Zhaohui.W Shihong.L Gang.P, Lin.S. Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In *2007 IEEE 11th international conference on computer vision*, pages 1–8. IEEE, 2007.

[29] Wichansky Anna.M Goldberg, Joseph.H. Eye tracking in usability evaluation : A practitioner's guide. *The Mind's Eyes : Cognitive and Applied Aspects of Eye Movements*, 2002.

[30] Courville.A Goodfellow.I, Bengio.Y. *Deep Learning.* MIT Press, 2016. `http://www.deeplearningbook.org`.

[31] Djallil.M Goussem.A. Authentification biométrique dans la réalité virtuelle. Master's thesis, Université Abderrahmane Mira, 2022.

[32] Frederick.J Gravetter and Larry.B Wallnau. *Statistics for the Behavioral Sciences*. Cengage/Wadsworth, Boston, 10th edition, 2015.

[33] Xiaoding.Y Zhijun.L Zichen.L Alois.K Guang.C, Fa.W. Neurobiometric : An eye blink based biometric authentication system using an event-based neuromorphic vision sensor. *IEEE/CAA Journal of Automatica Sinica*, 8(1) :206–218, 2021.

[34] Zhiguo.C Lubin.M Zhiwen.F Joey.Z Tianyi Junsong.Y Guilei.H, Yang.X. Towards real-time eyeblink detection in the wild : Dataset, theory and practices. *IEEE Transactions on Information Forensics and Security*, 15 :2194–2208, 2020.

[35] Andreas.S Xiaowei.X Hans-Peter.K, Jörg.S. Institute for computer science, university of munich.

[36] Harifi.K. Bien comprendre l'algorithme des k plus proches voisins (fonctionnement et implémentation sur r et python), September 21 2019.

[37] Evgeny.A Oleg.K Henry.G, Dillon.L. Gazebase, a large-scale, multi-stimulus, longitudinal eye movement dataset. *Scientific Data*, 8(1) :184, Jul 2021.

[38] Lang.Bo Hongyu.L. Machine learning and deep learning methods for intrusion detection systems : A survey. *Applied Sciences*, 9(20) :4396, 2019.

[39] Bender.M Fekete.P Joseph.M Hsiang.T, Arkin.E. Algorithms for rapidly dispersing robot swarms in unknown environments. *arXiv preprint arXiv :2002.XXXXX*, 2020.

[40] Mingyan.X Srinivasan.M Ming.L Huadi.z, Wenqiang.J. Blinkey : A two-factor user authentication method for virtual reality devices. *ACM Interact*, page 164–193, 2020.

[41] Mingyan.X Srinivasan.M Ming.L Huadi.z, Wenqiang.J. Blinkey : A two-factor user authentication method for virtual reality devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(4), dec 2020.

[42] Zhu.Y Iqbal.M. Supervised machine learning approaches : A survey. *ICTACT journal on soft computing*, 05 :946–952, April 2015.

[43] Jason.W. Decision trees. Jan 1.2021.

[44] Kaushik.R. 10 interesting use cases for the k-means algorithm. Mar.27.2018.

[45] Makwana.P Kodinariya.T. Review on determining number of cluster in k-means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6), November 2013.

[46] Jain.LC Larry.R. *recurrent neural networks*. Design and Applications, 2001.

[47] Wang.R Lee.T, Ullah.A. Bootstrap aggregating and random forest. In Peter Fuleky, editor, *Macroeconomic Forecasting in the Era of Big Data*, volume 52 of *Advanced Studies in Theoretical and Applied Econometrics*, chapter 13. Springer, Cham, 2020.

[48] Fontaine.A Lupera.P Llugsi.R, Yacoubi.S. Comparison between adam, adamax and adam w optimizers to implement a weather forecast based on neural networks for the andean city of quito. In *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, pages 1–6, 2021.

[49] Komogortsev.O Lohr.D, J.Aziz.S. Eye movement biometrics using a new dataset collected in virtual reality. *In ACM Symposium on Eye Tracking Research and Applicationst*, 2020.

[50] Li.J Liu.Z Lohr.M. Chen.Y, Chen.J. Eye know you too : A densenet architecture for end-to-end eye movement biometrics. *IEEE Transactions on Information Forensics and Security*, Jan 2022.

[51] Diego.L Lucas.S Alceu.Jr David.M Luiz.Z, Rayson.L. A new periocular dataset collected by mobile devices in unconstrained scenarios. *Scientific Reports*, 12, 10 2022.

[52] Lumivero. ForÊts alÉatoires de classification et de rÉgression. Copyright ©2023 Lumivero, 2023. `https://www.xlstat.com/fr/solutions/fonctionnalites/forets-aleatoires-de-classification-et-de-regression`.

[53] Matthieu.G. Optimisation des chaînes de production dans l'industrie sidérurgique : une approche statistique de l'apprentissage par renforcement. nov 2009.

[54] Weiss.E Milich.L. Gac ndvi interannual coefficient of variation (cov) images : Ground truth sampling of the sahel along north-south transects. *Views*, 67 :235–260, 2010. CrossRef citations to date : 0, Altmetric : 0.

[55] Mayank Mishra. Convolutional neural networks explained. *Towards Data Science*, August 2020.

[56] Fomani.B Mohamadally.H. Svm : Machines à vecteurs de support ou séparateurs à vastes marges. *BD Web, ISTY3*, 2006.

[57] Fomani.B Mohamadally.H. SVM : Machines à Vecteurs de Support ou Séparateurs à Vastes Marges. *BD Web*, ISTY3, January 2006.

[58] Necib.S. Fusion de face 3d couleur, profondeur et profil pour srv3d. Master's thesis, Université de Mohamed khaider, Biskra, 2013.

[59] Nelson.D. Qu'est-ce qu'un knn (k-nearest neighbors) ? *Unite.AI*, August 2020.

[60] netapp.com. What is machine learning ? `https://www.netapp.com/artificial-intelligence/what-is-machine-learning/`, (Consulté le 19 Nouvembre 2022).

[61] Researcher's Night. "`https://marie-sklodowska-curie-actions.ec.europa.eu/event/2022-european-researchers-night`.

[62] Data Transition Numérique. K-means : fonctionnement et utilisation dans un projet de clustering, 2021.

[63] Ouazine.K. *Alliances in graphs : properties and application for reducing saturation and congestion in VANETs networks*. PhD thesis in Computer Science, University of Bejaia, October 2018.

[64] Photography Life. Understanding camera aperture : What is it and how to use it ? `https://photographylife.com/what-is-aperture-in-photography`.

[65] PRABHU.R. Understanding of convolutional neural network (cnn)– deep learning. 2020.

[66] Roberto.I PratikS. Main types of neural networks and their applications — tutorial. Last updated March 17, 2022, https://pub.towardsai.net/main-types-of-neural-networks-and-its-applications-tutorial-734480d7ec8e.

[67] Smolka.B Radlak.K, Bozek.M. Silesian deception database : Presentation and analysis. page 29–35, New York, NY, USA, 2015. Association for Computing Machinery.

[68] Rakotomalala.R. Arbres de décision. *Revue Modulad*, 33 :163–187, 2005.

[69] RAOUNAK.L. *La d´etection de la col'ere chez le conducteur en utilisant le deep learning.* 2020.

[70] Rauscher.E.A. The minkowski metric for a multidimensional geometry. *Lawrence Berkeley National Laboratory*, 1973. Recent Work.

[71] Christoph.W.B Sarker.M.A, Arun.K.K. User identification utilizing minimal eye-gaze features in virtual reality applications. *Virtual Worlds*, 2022.

[72] Golay.M.J.E Savitzky.A. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8) :1627–1639, 1964.

[73] Lisa Schwarz. What is authentication ? definition. https://www.netsuite.com/portal/resource/articles/erp/authentication.shtml?fbclid=IwAR2kFUWWgD_aBiYGglcRkLW-YVrJoaFnUYrNq4I-QUbu6FW_zCTNPEINw0o, April 22, 2022.

[74] site web de keras. "https://keras.io/".

[75] site web de mediapipe. "https://github.com/google/mediapipe".

[76] site web de Numpy.com. "https://numpy.org".

[77] site web de opencv. "https://opencv.org/".

[78] site web de pandas. "https://pandas.pydata.org/".

[79] site web de python. "https://www.python.org/.

[80] Kheddouci.H Slimani.H. Saturated boundary k-alliances in graphs. *Discrete Applied Mathematics*, 185 :192–207, 2015.

[81] Xu.Z Srimani.P.K. Distributed protocols for defensive and offensive alliances in network graphs using self-stabilization. In *Proceedings of the International Conference on Computing : Theory and Applications*, page 2731, Kolkata, India, March 2007.

[82] TechTerms. What is resolution ? https://techterms.com/definition/resolution.

[83] Roman.B Tomi.K, Filip.S. Towards task-independent person authentication using eye movement signals. *Association for Computing Machinery New York, NY, United States*, page 187–190, 2010.

[84] A Vapnik.V, Lerner. Pattern recognition using generalized portrait method. *Automation and Remote Control*, 24, 1963.

[85] Videomaker. What is frame rate (fps) and why does it matter ? `https://www.videomaker.com/article/f2/17580-what-is-frame-rate-fps-and-why-does-it-matter`.

[86] Huang.W Dong.D Scott.M.R Wang.X, Han.X. Multi-similarity loss with general pair weighting for deep metric learning. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5017–5025, 2019.

[87] Davis346 website. `https://inivation.com/wpcontent/uploads/2021/08/2021-08-iniVation-devices-Specifications.pdf`, 2023.

[88] Wikipedia. High-definition television. `https://en.wikipedia.org/wiki/High-definition_television`.

[89] Jong-Cheol.J Xue-wen.C. Enhanced recursive feature elimination. In *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, pages 429–435, 2007.

[90] ZAMOUCHE.D. Intelligence artificielle et sécurité des réseaux. 2022.

[91] Sabuncu Zhang.Z. Generalized cross entropy loss for training deep neural networks with noisy labels. *Electrical and Computer Engineering, Meinig School of Biomedical Engineering, Cornell University*, 2021.

RÉSUMÉ

Dans ce document, nous avons abordé le développement et l'évaluation des systèmes d'authentification biométrique en utilisant les mouvements oculaires. Notre objectif premier est de fournir un vaste ensemble de données spécifiquement conçu pour le développement et le test de systèmes d'authentification biométrique exploitant les caractéristiques oculaires comme modalité d'authentification. L'ensemble de données comprend deux sous-ensembles : un jeu de données de type vidéos et un jeu de données de type images. Notre jeu de données se base sur la qualité des données qui est le facteur le plus important dans l'élaboration d'algorithmes d'apprentissage automatique, en particulier les protocoles d'authentification, car elle influence la précision et la capacité du modèle à classifier correctement les utilisateurs. Notre deuxième objectif est d'améliorer les protocole d'authentification existants, en particulier celui de Goussem et Djallil. Pour cela, nous avons mené une série expériences en utilisant notre jeu de données vidéos et le classificateur kNN afin de déterminer les points à améliorer. De plus, nous avons entraîné deux architectures d'apprentissage en profondeur avec CNN pour authentifier les utilisateurs en utilisant l'ensemble de données de type image. Nos résultats démontrent un taux de fausses acceptations significativement faible de 14.29 % pour à la fois leur protocole et nos architectures CNN proposées.

Mots-clés : Authentification biométrie oculaire, jeux de données vidéos, jeux de données images, kNN, CNN.

ABSTRACT

In this document, we have addressed the development and evaluation of biometric authentication systems using eye movements. Our primary objective is to provide a large dataset specifically designed for the development and testing of biometric authentication systems that exploit ocular biometrics as an authentication modality. Our dataset is divided int two subsets : a video-type dataset and an image-type dataset. The quality of our data is most important factor for machine learning algorithms, particularly authentication protocols, as it influences the precision and model capacity to correctly classify users. Our second objetive is to use the authentication protocol proposed by Goussem and Djallil in previous studies as a basis for experiments using our dataset and the kNN algorithm. Additionally, we trained two CNN architectures to authenticate users using the image-type dataset. Our results demonstrate a significantly low false acceptance rate of 14.29% for both thier protocol and our proposed CNN architectures.

Keywords : Eye-base biometric authentication, Video-based dataset, Image-based dataset, kNN, CNN.