

Problématique

Avec le développement des systèmes d'information et l'utilisation d'outils informatiques, il y a de plus en plus de :

- ✗ Fuites de données au sein des entreprises.
- ✗ Failles et problèmes de sécurité.

Les méthodes de sécurité classiques ne suffisent plus. Il est par conséquent primordial de mettre en place des mesures de sécurité robustes.

Une solution à ces problèmes est d'adopter une approche biométrique. Ce type de système permet d'identifier une personne en utilisant des caractéristiques biométriques comme l'empreinte digitale, la voix, le visage ou l'iris [3].

Objectifs

- ✓ Développer un système d'identification basé sur une image de l'iris en passant par des techniques d'apprentissage automatique.
- ✓ Mettre le système à disposition à travers une **API REST** destinée aux applications Web et mobiles.

Description du Système

Afin de réaliser le système (Figure 2) :

- **Scikit-Image** : <https://scikit-image.org/>
- **Scikit-Learn** : <https://scikit-learn.org/>

1 Pré-traitement

Les opérations appliquées (Figure 1) :

- **Hough Transform** [1] pour détecter et extraire l'iris dans les images,
- **Daugman's rubber sheet** pour transformer l'image circulaire en une rectangulaire,

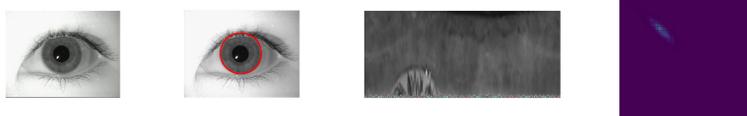


FIGURE 1 – Processus de pré-traitement (a, b et c) et aperçu de **GLCM** (d).

2 Extraction des caractéristiques

La **GLCM** (*Gray-Level Co-occurrence Matrix*) ou matrice de co-occurrences a été proposée par Haralick et al. [2]. Elle permet d'analyser les textures en calculant les répartitions des différents niveaux de gris dans l'image en se basant sur les positions relatives des pixels.

Il est possible de l'utiliser directement et appliquer la *réduction de dimensionnalité* ou bien d'extraire différentes caractéristiques comme le contraste, l'énergie, l'homogénéité, etc.

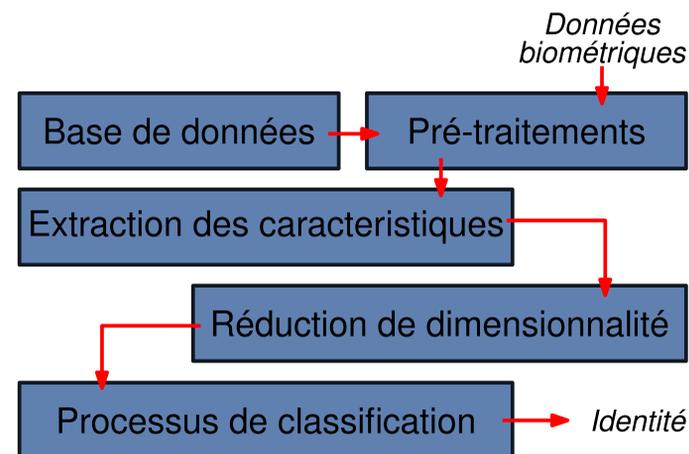


FIGURE 2 – Les différents blocs du système biométrique.

3 Réduction de la dimensionnalité

C'est un processus qui consiste à réduire considérablement la taille du vecteur de caractéristiques dans le but de faciliter le processus d'apprentissage.

La méthode utilisée est la **PCA** (*Principal Component Analysis*) ou l'analyse en composantes principales. Durant le processus de projection et transformation, seule l'information utile est conservée.

4 Processus de classification

Une technique d'apprentissage machine *supervisée* a été utilisée. C'est le **k-NN** (*k-Nearest Neighbors*) ou la méthode des k plus proches voisins.

Le **k-NN** consiste à trouver les k échantillons à partir de la base de données de référence qui se rapprochent le plus de l'échantillon d'entrée. Le critère utilisé est la distance de *Manhattan*. C'est sur la base d'un vote de majorité sur les k échantillons que l'identité est déterminée.

Méthodologie d'évaluation

- La base de données de référence **MMU iris dataset** a été téléchargée à partir de **Kaggle**.
- Caractéristiques biométriques pour 46 personnes, mais uniquement 10 sélectionnées aléatoirement ont été utilisées.
- Durant les évaluations, c'est la *validation croisée* qui a été adoptée. Elle permet de générer $K = 10$ parties *stratifiés* afin d'éviter les biais.

Résultats préliminaires

TABLE 1 – Matrice de confusion pour **GLCM-PCA** et **k-NN**.

Sujets	005	010	013	022	027	030	032	035	038	040
005	6	0	0	0	0	0	1	0	0	3
010	0	9	0	1	0	0	0	0	0	0
013	0	0	8	0	0	1	0	0	0	1
022	0	1	0	8	1	0	0	0	0	0
027	0	1	0	2	6	0	1	0	0	0
030	2	0	1	0	0	4	1	0	0	2
032	0	1	0	0	1	0	7	0	0	1
035	0	1	0	0	1	1	0	7	0	0
038	0	0	0	0	1	0	0	0	9	0
040	0	1	1	0	0	1	0	0	0	7

TABLE 2 – Comparaison des deux méthodes.

Méthode	Taux de reconnaissance
GLCM-PCA et k-NN	71%
GLCM et k-NN	22%

Conclusion

- ✓ Obtention de résultats intéressants avec la représentation **GLCM-PCA** qui est bien meilleure que la simple **GLCM** (+ *caractéristiques*).
- ✓ Améliorer le processus de pré-traitement pour mieux détecter l'iris dans l'image et pour retirer le bruit comme les cils.

Références

- [1] Richard O Duda and Peter E Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1) :11–15, 1972.
- [2] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6) :610–621, 1973.
- [3] Alfi Zuhriya Khoirunnisaa, Lutfi Hakim, and Adhi Dharma Wibawa. The biometrics system based on iris image processing : A review. In *2019 2nd International Conference of Computer and Informatics Engineering (IC2IE)*, pages 164–169. IEEE, 2019.