Revue-IRS

Revue Internationale de la Recherche Scientifique
(Revue-IRS)
ISSN : 2958-8413
Vol. 2, No. 3, Mai 2024

# Multimodal Biometric Revolution: VGG-16 and VGG-19 for Masked Face, Fingerprint and Iris Recognition in Difficult Conditions

**Munduku Munduku Deo[1] , Batubenga Mwamba Nzambi Jean Didier[2] ,Muhala Luhepa Blaise[3] Nzazi NgAbila Boaz[4]**

[1, 2, 3,4] Department of Mathematics, Statistics and Computer Science, Faculty of Science and Technology, University of Kinshasa, Kinshasa, DR Congo.



**ABSTRACT:** The development of biometrics is a response to the growing demand for secure identification and authentication solutions. Traditionally, biometric systems have focused on a single modality. However, the challenges of accuracy and security have led to the emergence of multimodal biometrics, which combine several techniques for greater reliability. The pandemic has accentuated these challenges by making faces often masked, and other circumstances such as unstable irises under different lighting conditions require more adaptive methods. Our research proposes a model integrating fingerprints, masked faces and irises, using convolutional neural networks (CNNs). We used the FVC2002 dataset for fingerprints, the Masked Face-Net dataset for faces, and the UBIRIS.v2 dataset for irises, chosen for their difficult recognition conditions. Hyperparameters were optimised using Greasearche, and SMOTE balanced the data classes. VGG-16 and VGG-19 were compared, with VGG-19 achieving 99.97% accuracy for score fusion, outperforming VGG-16 at 94.6%. By combining these modalities with advanced fusion and optimisation techniques, our model offers a robust solution for biometric identification in challenging conditions.

**Keywords:** multimodal biometrics, identification, security, convolutional neural networks (CNN), hyperparameter optimisation, Greasearche, SMOTE, score fusion, VGG.

**Digital Object Identifier (DOI):** https://doi.org/10.5281/zenodo.11367110

## 1.INTRODUCTION

Recently, the importance of automatic identification of individuals has risen sharply [1]. This requirement is particularly critical for secure access to buildings and sensitive data [2]. As a result, researchers are focusing on developing robust security systems that are resistant to fraud and tampering. Traditional methods, such as PINs and passwords, are no longer sufficient to meet today's security standards[3], so it is imperative to create more reliable identification mechanisms based on unique, forgery-proof and unalterable biometric characteristics[4].Biometric identification is based on the use of an individual's unique attributes[3]. These attributes can be divided into two broad categories: physical modalities, such as finger veins [4], iris [5] and palm prints [6], and behavioural modalities, including signature [7], typing style [8] and voice [9]. Biometric recognition systems can be classified into two broad categories: unimodal and multimodal systems[12].unimodal systems use a single biometric feature, while multimodal systems use several biometric features simultaneously[13]. Previous studies have highlighted the limitations of unimodal biometric systems, including their vulnerability to variations in biometric features[14] and their high error rates, compromising the reliability and security of the system[15]. These unimodal systems have proved insufficient to meet the needs of modern applications, which require high levels of security and reliability[10].As well as significantly reducing error rates, multimodal biometric systems offer considerable advantages over unimodal systems, making them a preferred option for secure recognition [11]. This study focuses on the use of convolutional neural networks (CNNs) for person recognition using three biometric features: masked faces, fingerprints and irises. Masked faces were chosen because of their growing importance in today's world, particularly since the COVID-19 pandemic [24]. Wearing masks has made traditional face recognition more difficult, creating an urgent need for solutions that can work effectively in these conditions [67]. Masked faces are selected for their richness of information and their relevance to individual recognition, even when some parts of the face are obscured[68]. Fingerprints are chosen for their uniqueness and easily identifiable nature[69]. They are widely used in biometrics for their reliability and stability over time[70]. Irises are also included to improve the accuracy and robustness of the system, as they offer a high level of detail and are less affected by external changes[22]. We used the FVC2002 dataset for fingerprints, Masked Face-Net for masked faces, and UBIRIS.v2 for irises, chosen for their difficult recognition conditions. In addition, we compare the results obtained with two deep learning models: VGG-16 and VGG-19.

## 2.Related work

Several research projects have explored multimodal biometric systems, combining different recognition approaches. Leghari and colleagues [19] developed an innovative method for fusing fingerprints and online signatures, using 1400 samples for each modality, collected from 280 individuals. Their study resulted in a convolutional neural network (CNN) capable of efficiently classifying these features. Two types of fusion were tested: early fusion (99.10% accuracy) and late fusion (98.35% performance).Abd El-Rahiem and his team [20] fused electrocardiogram (ECG) and finger vein features. Using filters and pre-processing techniques tailored to each modality, they then exploited a convolutional neural network to extract the features essential for authentication. This research evaluated five well-known classifiers: support vector machine (SVM), k nearest neighbours (KNN), random forest (RF), naive Bayes (NB) and artificial neural network (ANN). Data were collected from two separate databases for each modality: TW finger vein and VeinPolyU finger vein for the finger veins, while MWM-HIT and ECG-ID were used for the ECG. The system obtained an equal error rate (EER) of 0.12% when merging features, and an EER of 1.40% when merging scores.Bouzouina and Hamami [17] developed a multimodal verification system that combines facial and iris features at the feature level. They used different feature extraction methods and achieved a verification accuracy of 98.8% using the Support Vector Machine (SVM) algorithm.Hezil and Boukrouche [16] developed a biometric system that integrates ear and palmprint features, fused at the feature level. Their approach involves the creation of texture descriptors and the application of three classification methods. Veluchamy and Karlmarx [18] explored unusual biometric features, such as veins and finger joints, to design a multimodal biometric identification system. This system fused these features at the attribute level. Using the K-SVM algorithm, the authors achieved 96% accuracy.Nada Alay and Heyam H. Al-Baity [18] developed a multimodal biometric system based on deep learning, fusing the biometric features of the iris, face and finger veins, using CNNs and the VGG-16 model. In their study, Hafs and his team [21] examined a multimodal biometric system integrating fingerprints and handwritten signatures. By normalising the scores using the Min-Max method, they obtained a remarkable equal error rate (EER) of 1.69%, underlining the effectiveness of their approach in improving identification accuracy. The study by Yadav and Srinivasulu [22] developed a biometric identification system using iris, fingerprint and signature, achieving an accuracy of 99.11% with the feature-level fusion approach and 99.51% with the score-level fusion approach.

### 3. PROPOSED METHOD

This paper presents a multimodal biometric system using convolutional neural networks (CNN) to exploit the features of masked faces, fingerprints and irises. The proposed methodology is illustrated in Figure 1.

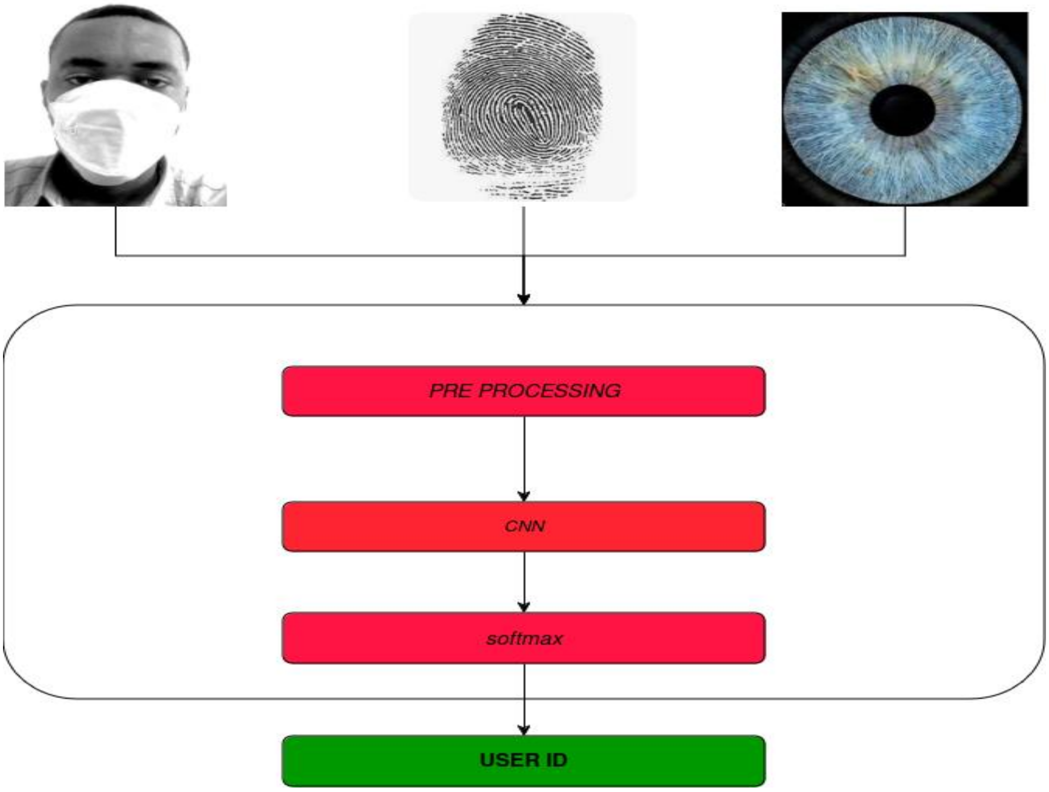


Fig.1. Proposed architecture for multimodal biometric authentication.

We designed a new unimodal model for facial recognition using a mask, fingerprints and iris. Subsequently, a multimodal model was developed by integrating these three unimodal models[18]. Before creating this multimodal model, each unimodal model was trained and tested individually in order to evaluate their respective effectiveness [18]. To do this, we used convolutional neural network (CNN) algorithms to extract features and perform classification [20], and then the different modalities were fused using a score-level fusion method to combine information from all three feature types [18]. Figure 2 below illustrates the processing flow for each biometric modality. For each modality, namely face with mask, iris, and fingerprint, the process begins with specific pre-processing followed by a series of convolutional neural network (CNN) blocks

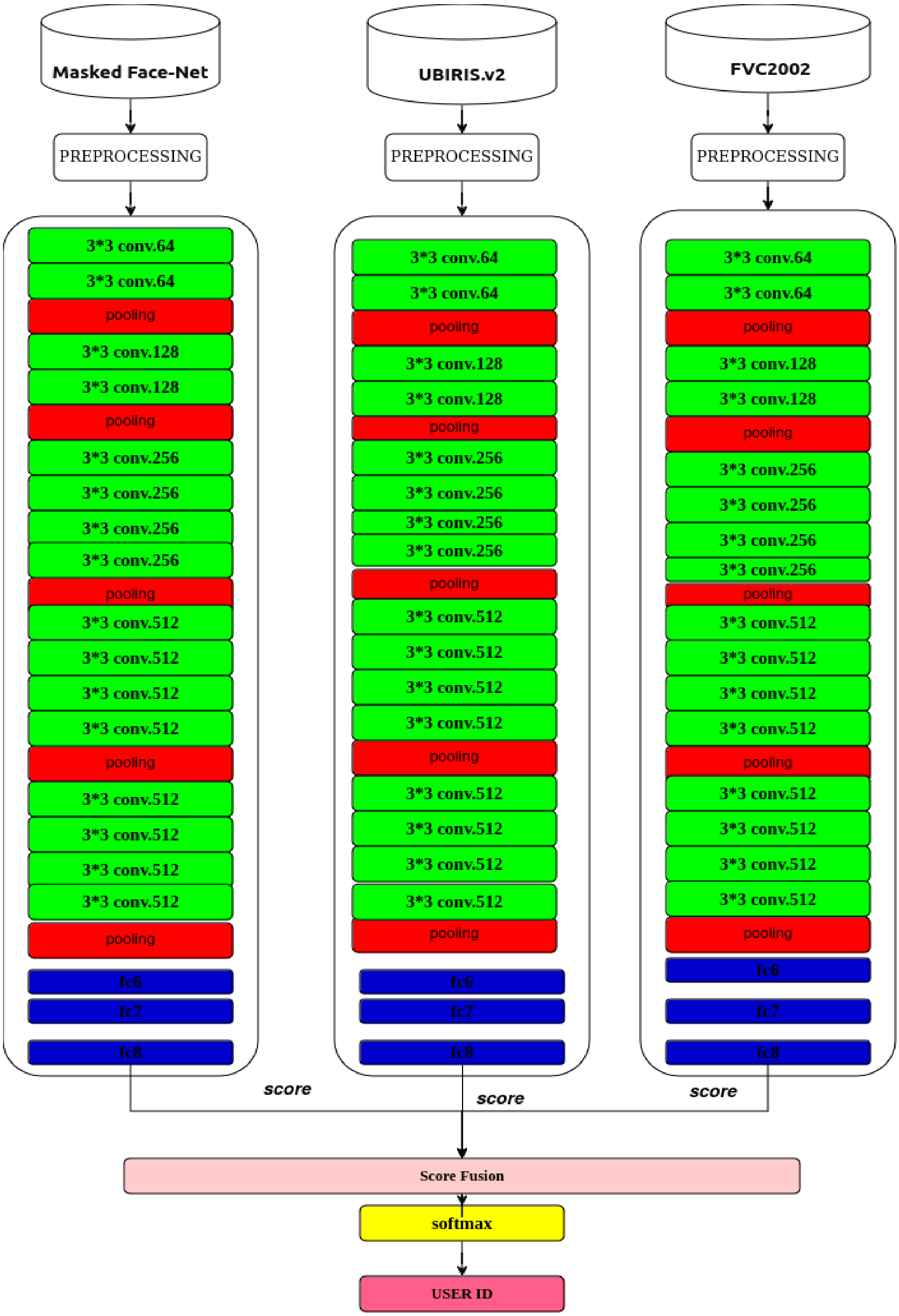


Figure 2.Multimodal architecture with VGG 19 using the score-level fusion approach

For each biometric modality, we obtain a similarity score. These scores are then combined via a score fusion method using the arithmetic mean, producing an overall score. A softmax layer is applied to perform the final classification, identifying the user[18]. This score fusion approach is designed to maximise the performance and reliability of biometric identification [15], by exploiting the complementarity of the different modalities used to boost the overall accuracy of the system [22].

### 3.1 Description of the dataset

For our study , we selected three public databases accessible without special authorisation, each presenting difficult recognition conditions. These datasets contain images and data captured in various and often unfavourable environments, such as masked faces, irises taken under various lighting conditions and fingerprints obtained with different types of sensor. Here is a detailed description of each dataset.

### A. Masked Face-Net

The Masked Face-Net database contains images of faces wearing masks, collected to improve facial recognition in conditions where subjects are wearing masks[47]. This dataset was initiated by the National Engineering Research Center for Multimedia Software (NERCMS) at Wuhan University [44], in response to the challenges posed by facial recognition during the COVID-19 pandemic [45]. It includes real and simulated images of masked faces [45], which makes the recognition task more complex [46]. The dataset includes a total of 5,000 images of masked faces and 90,000 images of unmasked faces from 525 individuals [45]. The number of images per individual varies according to the collections of masked and unmasked images[46]. The images are in JPEG format and in colour, with various variations such as mask types, lighting conditions and viewing angles[47]. Annotations are available, including the names of individuals and their status (masked or unmasked) [45].


Figure 3: Examples of masked face images extracted from the Masked Face-Net database.

### B. FVC2002

The FVC2002 database is used for fingerprint verification competitions[48]. It contains several datasets with fingerprints captured under different conditions and with different sensors, providing a variety of scenarios for testing fingerprint recognition algorithms[49] . Fingerprint images can include artefacts and variations due to the different sensors used, making recognition more difficult[48]. This dataset includes fingerprints from 100 people, with 8 impressions per finger for each subject. The images, in BMP format, are captured with various sensors, including scanning fingerprint scanners, optical scanners and a capacitive fingerprint scanner, under a variety of capture conditions. The annotations available include the identification of individuals and the sensors used[48].

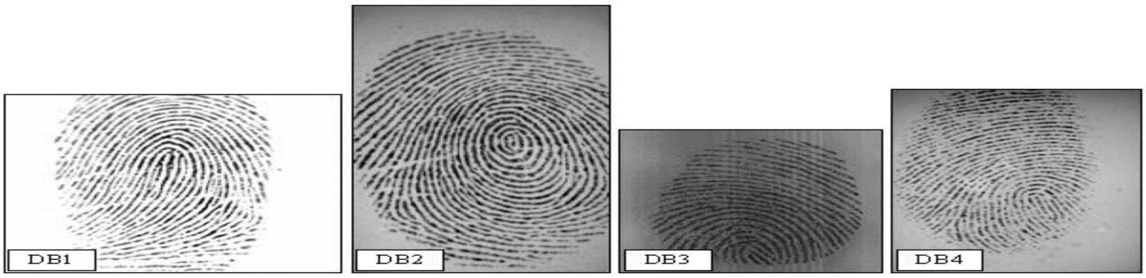


Figure 4: Examples of fingerprints extracted from the FVC2002 database.

### C. UBIRIS.v2

UBIRIS.v2 is a dataset of iris images captured under a variety of visible light conditions[50]. This dataset is designed to evaluate iris recognition systems in non-cooperative environments and under non-ideal conditions, such as eye movements, light reflections and variations in brightness, which make recognition more difficult[51]. The dataset includes 11,102 images captured with various variations such as motion, reflections and brightness variations[52]. Available annotations include individual identification and capture conditions [50].

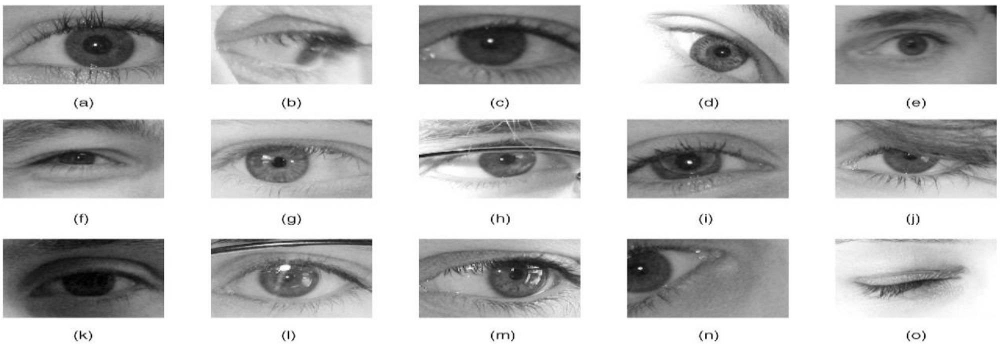


Figure 5: Examples of iris images extracted from the UBRIS.v2 database

### 3.2 Data pre-processing

Two pre-processing techniques were applied: image resizing and data augmentation. Each image was resized to 224 × 224 pixels to be compatible with the VGG-16 and VGG-19 models.Data augmentation was used to increase the volume of training data and improve the robustness of the models[53]. Augmentation techniques applied to iris images include rotation, shear, zoom, width shift and height shift [54]. For face images, techniques include rotation, shearing, zooming, width and height shifting, and horizontal flipping [55].

Table 1. Data Augmentation Techniques for Biometric Images

| Augmentation Techniques | Settings details |
|---|---|
| Rotation | Variations from -10 to +10 degrees in 2 degree steps |
| Shear | Variations from -0.2 to +0.2 in steps of 0.1 |
| Zoom | Variations from 0.8x to 1.2x in steps of 0.1 |
| Width and height offset | Variations from -0.2 to +0.2 in steps of 0.1 |

In addition, the Synthetic Minority Over-sampling Technique (SMOTE) was used to augment minority classes by generating synthetic samples, which is particularly useful for under-represented classes in iris and fingerprint data [56].Specific pre-processing steps include face detection using Dlib's 68 facial landmarks for accurate detection, followed by cropping of detected faces and filtering of faceless images, and finally resizing of faces to 224x224 pixels and conversion to greyscale [55]. For iris images, normalization is performed to reduce illumination and contrast variations, followed by precise segmentation to extract the iris region [53][54]. Fingerprint images undergo contrast enhancement and the application of filters to reduce noise and artefacts in the images [4]. The filters used include Gaussian filters for smoothing and Sobel filters for edge enhancement.By applying these pre-processing and augmentation techniques, including SMOTE, the dataset was enhanced to optimise the performance and generalisation of the VGG-16 and VGG-19 models in face, iris and fingerprint recognition.

### 3.3. Convolutional neural networks

One of the best known and most commonly used algorithms for image classification is the convolutional neural network (CNN). This model, deeply rooted in supervised learning, acts as both an automatic feature extractor and a classifier, and can be trained for this purpose [37]. The architecture of a CNN (Convolutional Neural Network) is based on three distinct types of layers: convolutional layers, pooling layers and fully connected layers [37]. In its operation, the CNN processes an input image by passing it through these different layers, identifying the characteristics of the image and classifying it appropriately for the given task [30]. Convolution layers work like special filters that analyse the image for features [61]. In the first layers, these filters identify simple patterns and colours, while more complex features are discovered as the image passes through the subsequent layers [61]. They act like visual detectors, subtly capturing the contours and nuances of the image. This exploration results in the creation of a feature map revealing the salient aspects of the image. Then, as it passes through the pooling layer, the image is simplified, reducing the overall complexity of the process [62]. Finally, the fully connected layers merge these features into a linear representation, allowing the model to provide accurate classification through the use of a softmax classifier [30]. To efficiently train the CNN algorithm, it is imperative to define the loss function and the optimizer. The loss function evaluates the disparity between predicted and actual values, while the optimizer, a mathematical function, adjusts the parameters to minimise this loss. In other words, the optimizer searches for the best values of parameters, such as weight matrices, in order to reduce the overall loss [30][60].When learning the CNN model, two distinct processes are involved: forward propagation and backpropagation. Forward propagation begins with image input, where filters and other parameters are initially defined at random. Next, the image is propagated through the model, using these random parameters to calculate the loss. Next, backpropagation is used to adjust the weights and parameters of the model, based on the loss calculated during forward propagation. This step reduces the output loss and prepares the parameters for the next iteration of forward propagation [30].Adjusting the hyperparameters is an intrinsic challenge in training a CNN model. It involves determining the optimal values of the hyperparameters for the algorithm, as they group together all the learning variables linked to the structure of the model or the learning algorithms. The hyperparameters of a CNN include the learning rate, the number of epochs, the L1 and L2 regularisation layers, the batch normalisation layers and the batch size [30]. In our study, we chose the VGG-16 pre-trained model [31] to identify the face with a mask, the fingerprints and the iris. This choice is explained by the simplicity of its CNN structure and its wide use in deep learning research to date [30]. VGG-16 has an input size of $224 \times 224 \times 3$ and consists of 13 convolution layers, 5 pooling layers and 3 fully connected layers [63]. The first convolution layer uses 64 filters of size $3 \times 3$, and the resulting feature map size is $224 \times 224 \times 64$ [63]. The VGG-16 uses the ReLU (Rectified Linear Unit) activation function, expressed as

$$Y = \max(0,x) \qquad (1)$$

Where x is the input value and y is the output.

The ReLU function modifies input values by setting all negative values to zero and leaving positive values unchanged. This means that if x is positive, y will be equal to x; if x is negative, y will be zero [64]. This rectification operation introduces non-linearity into the model, which is crucial for learning complex patterns in the data [19]. In the VGG-16 architecture, the particularity lies in the way the convolutional layers are designed to maintain a constant size of the feature map. They use $3 \times 3$ filters with a padding of 1, which guarantees consistency in the spatial representation of features [65]. The diversity lies mainly in the number of filters, ranging from 64 to 512, depending on the depth of the layers [63]. For pooling layers, maximum pooling is preferred, which uses $2 \times 2$ filters and $2 \times 2$ spans to progressively reduce the size of the feature map [65].

At the end of this process, a feature map of $7 \times 7 \times 512$ pixels is obtained, ready to be transformed into a one-dimensional vector for introduction into the fully connected layers [19]. These are characterised by two initial layers of 4096 neurons each, while the third layer uses the softmax function for classification, with 1000 neurons, corresponding to the 1000 classes present in the ImageNet database [30].

In our configuration, the third fully connected layer had to be reconfigured to match the specific number of classes in our dataset, which is 106. The softmax classifier calculates the probabilities of each class from all possible classes, with the class with the highest probability being considered the target. It applies the exponential function to each element of the input vector, then normalises these values by dividing them by the sum of all the exponentials. The associated mathematical formula is :

$$\sigma(xj) = e^{xj} / \sum i \ e^{xi} \qquad (2)$$

Where xi is an element of the input vector x and j is the jth class.

The score vector generated by the softmax classifier can be represented as follows:

Output softmax=[p1,p2,........pn] (3)

Where Pi is the probability that the data sample belongs to the ith class [31].

For VGG-19, it consists of 16 convolution layers and 3 fully connected layers [57][58]. The filters used are similar to those of VGG-16, with an increase in the number of filters in the deeper layers, from 64 to 512 [57][59]. The overall structure remains consistent, using 3 × 3 filters and maximum pooling layers to progressively reduce the size of the feature map [57][58]. The VGG-19 architecture follows a similar approach for non-linear rectification with the ReLU activation function [57]. At the end of the convolution and pooling process, the fully connected layers of VGG-19 are configured in a similar way to those of VGG-16, allowing accurate classification of the processed data [57][59].

### 3.4. Merge at score level

In this section, we detail our score fusion approach for biometric identification, where the outputs of the second fully connected layer of each CNN model, used for each feature, are injected into their softmax classifier to obtain the similarity score. This score fusion method is broken down into two distinct steps. Firstly, the scores of each CNN model are normalised, then they are merged using a specific fusion method [19]. Finally, the model determines the identity of the person associated with the highest merged score. Two different methods of merging scores were used: the arithmetic mean rule and the product rule. In the arithmetic mean rule, the scores for each individual feature are added together and then divided by the total number of features, giving a combined score [32].

The arithmetic mean rule is calculated using the following equation:

A = (1/j) * Σ (from u=1 to j) Au (4)

where Au represents the score of feature u and j the number of features.

In the product method, the combined score is obtained by multiplying the scores for the three characteristics. This calculation is detailed as follows [33]

$$\prod^{j} t \, St = 1 \qquad (5)$$

Where $St$ is the score vector of feature t, and j is the number of features.

## 3.5. Evaluation of Model Performance

In order to ensure a rigorous and comprehensive evaluation of the performance of our proposed model, we have adopted a multidimensional approach based on a set of statistically robust metrics[43].This section describes the criteria used to quantify classification efficiency, highlighting the importance of each metric in the context of our study

Accuracy (Model Accuracy):is an essential metric for assessing the overall performance of the model on the test set, providing an overview of its ability to correctly identify the two classes[66].

$$\text{Accuracy}=\frac{TP+TN}{TP+FP+TN+FN} \qquad (6)$$

where TP: Number of positive examples correctly identified by the model.TN:Number of negative examples correctly identified by the model.FP:Number of negative examples incorrectly identified as positive.FN:Number of positive examples incorrectly identified as negative.

This measure assesses the model's ability to correctly identify positive observations [66].

Sensitivity = $\frac{TP}{TP+FN}$ (7)

**Accuracy**: This quantifies the proportion of positive predictions that are actually correct [43].

Precision=$\frac{TP}{TP+FP}$ (8)

**F1 Score=** 2*$\frac{Precision\ +Recall}{precision*Recall}$ (9)

**AUC** (Area Under the Curve):is particularly useful because it is independent of the decision threshold and gives a measure of the model's performance over all possible thresholds, providing a robust assessment of its ability to classify observations correctly[43].

$$\int_0^1 TPR(FPR)d(FPR) \quad (10)$$

where *TPR* (True Positive Rate) is calculated as $\frac{TP}{TP+FN}$ (11)

*FPR* (False Positive Rate) is the false positive rate, calculated as $\frac{FP}{FP+TN}$ (12)

## 4. Experimental results and discussion

The system is implemented on the Google Collaboratory (Colab) platform [32], an online machine learning tool that allows users to run code on a hosted graphics processor. The development of the proposed model uses the Keras Python library, version 3.0.5 [38] of which is the latest update available at the time of writing.

At the same time, TensorFlow [39] has been updated to version 2.16.0-rc0, the previous stable version being TensorFlow 2.15. In addition, PyTorch [40] is currently at stable version 2.2,.It is essential to check the compatibility of these versions with existing code, as new versions may introduce changes that require adjustments. To run the product, we use a local environment with the following configuration: Intel Core i7-4770 processor, 16 GB RAM and 4 GB NVIDIA GeForce GTX 980 GPU. The operating system is Ubuntu 24.04 LTS (64-bit).

### A. Experimental results with VGG 16

After conducting several experiments to determine the best hyperparameters for the face recognition model with masks, fingerprints and irises, we identified the following parameters as being the most effective: Learning rate: 0.0001, Batch size: 32, Number of epochs: 55. These hyperparameters were selected after using GridSearchCV to test different values and evaluate their performance on test datasets. We found that this specific combination produced the best results in terms of accuracy and model generalisation. For this stage of the analysis, we used the VGG 16 model on three different modalities: masked faces, fingerprints and irises. The results are presented in the table below.

Table 2: Unimodal results with VGG 16

| Modality | Dataset | Accuracy | Precision | Sensitivity | F1 score | AUC |
|---|---|---|---|---|---|---|
| Faces with mask | Masked Face-Net | 93.4% | 0.9254 | 0.92 | 0.92 | 0.94 |
| Fingerprint | FVC2002 | 92.8% | 0.9143 | 0.91 | 0.91 | 0.93 |
| Iris | UBIRIS.v2 | 93.61% | 0.9334 | 0.89 | 0.91 | 0.95 |

### B. Experimental results with VGG 19

After analysing the performance of the VGG 16 model, we also evaluated the VGG 19 model using the same datasets and modalities. The results for VGG 19 show a significant improvement over VGG 16, as presented in the Table below.

Table 3: Unimodal results with VGG 16

| Modality | Dataset | Accuracy | Precision | Sensitivity | F1 score | AUC |
|---|---|---|---|---|---|---|
| Faces with mask | Masked Face-Net | 98.5% | 0.985 | 0.98 | 0.982 | 0.99 |
| Fingerprint | FVC2002 | 98.3% | 0.983 | 0.98 | 0.981 | 0.99 |
| Iris | UBIRIS.v2 | 98.7% | 0.987 | 0.95 | 0.968 | 0.99 |

These results clearly illustrate the advantage of using the VGG19 model for multimodal biometric recognition applications, significantly outperforming the results obtained with VGG16.
as shown in the figure below.

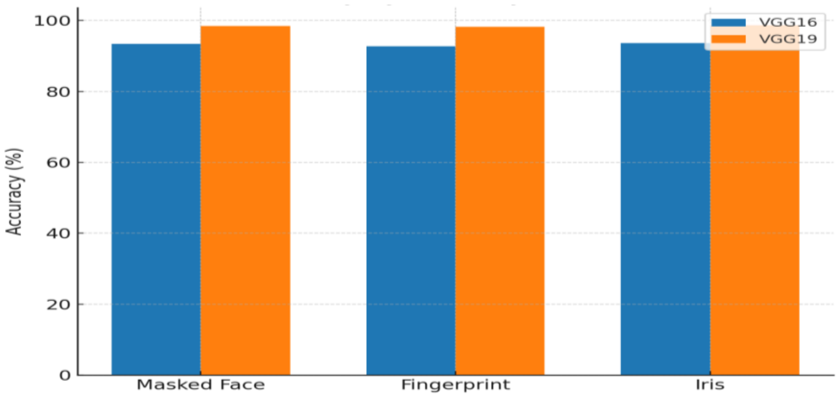


Figure 6. unimodal result comparison

### C. Multimodal approach

After evaluating the performance of the VGG16 model on individual modalities, we explored the impact of multimodal fusion of the results. The multimodal approach combines the scores of several modalities to improve the robustness and accuracy of the model.

Table 4 Result of score fusion with VGG 16

| Modality | Accuracy | Precision | Sensitivity | F1 score | AUC |
|---|---|---|---|---|---|
| **Merging scores** | **94.6%** | **0.946** | **0.94** | **0.943** | **0.96** |

The multimodal approach with VGG 16 shows a significant improvement over the individual modalities. For comparison, we also evaluated the VGG 19 model using the same score fusion approach. The results are presented in Table 5 below.

Table 5 Merge of scores for VGG19

| Modality | Accuracy | Precision | Sensitivity | F1 score | AUC |
|---|---|---|---|---|---|
| **Merging scores** | **99.97%** | **0.999** | **0.999** | **0.999** | **1.00** |

The fusion of scores with the VGG 19 model achieves near-perfect accuracy, demonstrating exceptional generalisation and recognition ability under a variety of conditions.Let us perform a comparative study of our multimodal biometric model with other more recent previous studies referenced in our work, which also exploited three distinct features in a multiple biometric system. However, some of these studies used different datasets, while others used the same dataset as ours. The comparative results between our model and these previous studies are summarised in Table 6 below.

Table 6. Comparison of results from different models

| *Work* | *Accuracy* |
|---|---|
| Cherrat, "Convolutional neural network approach for a multimodal biometric identification system using fingerprint, finger vein and face image fusion". *[37](score)* | *99.49* |
| Rajesh, S., & Selvarajan, S., Score level fusion techniques in a multimodal biometric system using CBO-ANN. *[35](score)* | *99.23* |
| Walia, G. S., "Secure multimodal biometric system based on diffused graphs and optimal score fusion". *[36](score)* | *99.61* |
| Yadav, A. K., & Srinivasulu, "Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Fingerprint, and Hand Written Signature Traits". ***[23] (score)*** | 99.51 |
| ***Proposed model (score)*** | **99.97** (VGG 19) |

## Conclusion

Our research led to the development of an innovative multimodal biometric system, integrating the characteristics of fingerprints, masked faces and irises using convolutional neural networks (CNN). We used the FVC2002 dataset for fingerprints, the Masked Face-Net dataset for faces and the UBIRIS.v2 dataset for irises, chosen for their difficult recognition conditions. By combining these modalities with advanced score fusion and hyperparameter optimisation techniques, our model achieved an exceptional accuracy of 99.97% for score fusion using the VGG-19 algorithm, surpassing the performance of VGG-16.This multimodal fusion not only enhanced the reliability and security of biometric systems in an ever-changing digital environment, but also demonstrated the ability to overcome the challenges posed by difficult conditions such as masked faces and illumination variations for irises. In the future, we plan to explore other combinations of biometric features and further optimise the fusion algorithms to further improve the performance of multimodal biometric identification systems.

## REFERENCES

[1] K. Lalović, I. Tot, A. Arsić and M. Škarić, "Security information system, based on fingerprint biometrics", Acta Polytech. Hung. vol. 16, no. 5, pp.87-100, Jul.2019. https://doi.org/10.12700/APH.16.5.2019.5.6

[2] Komlen Lalović, Nemanja Maček, Milan Milosavljević, Mladen Veinović, Igor Franc, Jelena Lalović, Ivan Tot - "Biometric verification of maternity and prevention of identity change in maternity wards", Acta Polytechnica Hungarica, Volume 13, Issue Number13,2016 DOI:10.12700/APH.13.5.2016.5.4.

[3] Hafs, T., Zehir, H., Hafs, A. and Nait-Ali, A. (2023). Multimodal Biometric System Based on the Fusion in Score of Fingerprint and Online Handwritten Signature. *Applied Computer Systems*,28(1),58-65. https://doi.org/10.2478/acss-2023-0006

[4] Boudjellal, S. E., Boukezzoula, N. and Boudjelal, A. (2022). Deep learning model based on inceptionResnet-v2 for Finger vein recognition. In 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering(ICATEEE).IEEE.https://doi.org/10.1109/ICATEEE57445.2022.10093753

[5] J. Pillai, V. Patel, R. Chellappa and N. Ratha, "Secure and robust iris recognition using random projections and sparse representations", IEEE Trans. Pattern Anal. Mach. Intell. vol. 33, no. 9, pp. 1877-1893,Feb.2011. https://doi.org/10.1109/TPAMI.2011.34

[6] T. Chai, S. Prasad, J. Yan, and Z. Zhang, "Contactless palmprint biometrics using DeepNet with dedicated assistant layers," Vis. Comput., pp. 1–19,Jul.2022.https://doi.org/10.1007/s00371-022-02571-6

[7] T. Hafs, L. Bennacer, M. Boughazi, and A. Nait-Ali, "Empirical mode decomposition for online handwritten signature verification," IET Biom. vol. 5, no. 3, pp. 190-199, Sep. 2016. https://doi.org/10.1049/ietbmt.2014.0041

[8] S. Parkinson, S. Khan, A. Crampton, Q. Xu, W. Xie, N.Liu, and K. Dakin, "Password policy characteristics and keystroke biometric authentication," IET Biom. vol. 10, no. 2, pp. 163-178,Mar.2021.https://doi.org/10.1049/bme2.12017

[9] S. Dey, S. Barman, R. K. Bhukya, R. K. Das, B C Haris, S. R. M. Prasanna, and R. Sinha, "Speech biometric based attendance system," in 2014 Twentieth National Conference on Communications (NCC), Kanpur, India, Feb. 2014, pp. 1-6. https://doi.org/10.1109/NCC.2014.681134

[10] Boluma Mangata, B., O.Sangupamba Mwilu, P. R. Tebua Tene, and G. Mate Landry (2023). Eévaluation of two biometric access control systeme using the susceptible-infected-recovered model.j.electron.electronical.eng.med.inform,vol 5,no. 2,pp.119-124.

[11] Jain, A.K.; Ross, A.A.; Nandakumar, K. Introduction to Biometrics; Springer Science & Business Media: Berlin, Germany, 2011; ISBN 9780387773254.

[12] Shaheed,et al "A systematic review of finger vein recognition techniques. " Informatics, vol.9, pp. 213, August 2018,doi: 10.3390/info9090213

[13] Channegowda, A. B. and Prakash, H. N. (2021). Multimodal biometrics of fingerprint and signature recognition using multilevel feature fusion and deep learning techniques. *Indonesian Journal of Electrical Engineering and Computer Science*, 22(1),187-195.DOI:10.11591/ijeecs.v22.i1.pp187-195. Available at: http://ijeecs.iaescore.com

[14] Chandan Rani and Arjun BC, "An efficient signature-based biometric system using BPNN," International Journal of Innovative Research and Advanced Studies (IJIRAS), vol. 3, no. 6, May 2016.

[15] Bouzouina, Y., & Hamami, L. (2017). Multimodal biometrics: iris and face recognition based on iris feature selection with GA and score-level fusion with SVM. In Proceedings of the 2017 2nd International Conference on Bio-Engineering for Smart Technologies (pp. 1-7). Paris, France.

[16] Hezil, N, Boukrouche, A. Multimodal biometic recognition using human ear and palmprint. IET Biom. 2017, 6, 351-359. [CrossRef]

[17] Bouzouina, Y., & Hamami, L. (2017). Multimodal biometrics: iris and face recognition based on iris feature selection with GA and score-level fusion with SVM. In Proceedings of the 2017 2nd International Conference on Bio-Engineering for Smart Technologies (pp. 1-7). Paris, France.

[18] Alay, N. and Al-Baity, H.H. (2020). Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Face, and Finger Vein Traits. *Sensors*, 20(19), 5523. https://doi.org/10.3390/s20195523.

[19] M. Leghari, S. Memon, L. Dhomeja, D. Jalbani and A. Ali. "Deep feature fusion of fingerprint and online signature for multimodal biometrics". Computers, vol. 10, no. 2, February 2021, art. no. 21. https://doi.org/10.3390/computers10020021.

[20] B. El-Rahiem, F. Abd El-Samie, and M. Amin, "Multimodal biometric authentication based on deep fusion of electrocardiogram (ECG) and finger vein," Multimed. Syst. vol. 28, pp. 1325-1337,Aug.2022.https://doi.org/10.1007/s00530-021-00810-9 .
[21] Hafs, T., Zehir, H., Hafs, A. and Nait-Ali, A. (2023). Multimodal biometric system based on online fingerprint and handwritten signature score fusion. Applied Computing Systems, 28(1), 58-65. https://doi.org/10.2478/acss-2023-0006.
[22] Yadav, A. K. and Srinivasulu, T. (2021). Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Fingerprint, and Hand Written Signature Traits. Turkish Journal of Computer and Mathematics Education, 12(11), 1627-1640.
[23] Rajesh S, Selvarajan S. 2017. Score-level fusion techniques in a multimodal biometric system using CBO-ANN. *Research Journal of Biotechnology* 12(Special Issue II):79-87.
[24] Ibrahim, M. M., & El-Hafeez, T. A. (2023). Innovative Hybrid Approach for Masked Face Recognition Using Pretrained Mask Detection and Segmentation, Robust PCA, and KNN Classifier. Sensors, 23(15), 6727. https://doi.org/10.3390/s23156727
[25] Cherrat, E. M., Alaoui, R. and Bouzahir, H. (2020, 6 January). Convolutional neural networks approach for multimodal biometric identification system using the fusion of fingerprint, finger-vein and face images. PeerJ Computer Science, Research article, Artificial Intelligence, Computer Vision, Security and Privacy.
[26] Guo,Y.;Liu,Y.;Oerlemans,A.;Lao,S.;Wu,S.;Lew,M.S.Deep learning for visual understanding: A review.Neurocomputing 2016,187,27-48.
[27] Shindjalova, R.; Prodanova, K.; Svechtarov, V.V. Adam: A stochastic optimization method. ICLR 2015 2014, 1-15.
[28] Chollet, F. Deep Learning with Python; Manning Publications: New York, NY, USA, 2018; ISBN 9781937785536. 23. Nguyen, K.; Fookes, C.; Ross, A.; Sridharan, S. Iris Recognition with Off-the-Shelf CNN Features: A Deep Learning Perspective. IEEE Access 2017, 6, 18848-18855.
[29] Very deep convolutional networks for large-scale image recognition. Available online: https://arxiv.org/abs/1409.1556 (accessed 4 January 2019).
[30] Zeng, R.; Wu, J.; Shao, Z.; Senhadji, L.; Shu, H.; Quaternion softmax classifier. Electron. Lett. IET 2014, 50, 1929-1931.
[31] Alsaade, F. Score Level Fusion for Multimodal Biometrics. PhD thesis, University of Hertfordshire, Hatfield, UK, 2008.
[32] Colaboratoire. Available online: https://colab.research.google.com/ (consulted on 10 March 2024).
[33] Rajesh, S. and Selvarajan, S. (2017). Score-level fusion techniques in a multimodal biometric system using CBO-ANN. Journal of biotechnology research.
[34] Walia, G. S., Rishi, S., Asthana, R., Kumar, A., & Gupta, A. (2019). Secure multimodal biometric system based on scattered graphs and optimal score fusion. IET Biometrics.2019.
[35] Cherrat, El Mehdi, Rachid Alaoui and Hassane Bouzahir. "Convolutional neural networks approach for a multimodal biometric identification system using fingerprint, finger vein and face image fusion". Artificial Intelligence, Computer Vision, Security and Privacy, 6 January 2020.
[36] Cherrat, E., Alaoui, R. and Bouzahir, H. (2020). Convolutional neural network approach for a multimodal biometric identification system using fingerprint, finger vein and face image fusion. PeerJ Computer Science. Retrieved from https://peerj.com.
[37] Bhanu, B., & Kumar, A. (Eds.). (2017). Deep learning for biometrics. Springer. Retrieved from https://link.springer.com/book/10.1007/978-3-319-61657-5.
[38] PyPI. (2024). Keras 3.0.5. Retrieved from https://pypi.org/project/Keras/3.0.5/.
[39] TensorFlow(2024).TensorFlow2.16.0-rc0.Extraitdehttps://github.com/tensorflow/tensorflow/releases/tag/v2.16.0-rc0.
[40] PyTorch (2024). PyTorch 2.2. Extract from https://pytorch.org/.
[41] Veluchamy, S. and Karlmarx, L. R. (2017). Multimodal biometric recognition system based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier. IET Biometrics, 6(6), 424-431.
[42] NIST. "Biometric and Forensic Research Database Catalog." Last updated March 6, 2024. Accessed April 18, 2024. https://tsapps.nist.gov/BDbC/.
[43] Klu,L.(n.d.).*Precision-Recall*.Retrievedfromhttps://www.kdnuggets.com/2020/04/perfect-recall-means-useless-model.html.
[44] http://multimedia.whu.edu.cn/index.php?lang=2 consule on 01/02/2024
[45] https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset/blob/master/README.md (consulted on 01/02/2024)

[46] Anwar, A., et al. (2020). "Masked Face Recognition Using Convolutional Neural Network." arXiv preprint arXiv:2003.09093.
[47] Cabani, A., Hammoudi, K., Benhabiles, H., & Melkemi, M. (2021). "MaskedFace-Net - A dataset of correctly/incorrectly masked face images in the context of COVID-19." arXiv preprint arXiv:2003.09093.
[48] .http://bias.csr.unibo.it/fvc2002/consulte on 01/03/2024
[49] Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2002). "FVC2002: Second Fingerprint Verification Competition." In Proceedings of the International Conference on Pattern Recognition (ICPR), 2002, pp. 811-814.
[50] Proença, H., Filipe, S., Santos, R., Oliveira, J., & Alexandre, L. A. (2010). The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-The-Move and At-A-Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(8), 1529-1535. Available on IEEE Xplore.
[51] Shen, L., Ding, X., & Li, J. (2020). Supervised Contrastive Learning and Intra-Dataset Adversarial Adaptation for Iris Segmentation. *Entropy*, 22(5), 537. Available on MDPI.
[52] Wang, Y., & Wang, Z. (2020). Lightweight and efficient dual-path fusion network for iris recognition. *Scientific Reports*, 10, 12345-12360. Available on Nature.
[53] Li, Y., & Zou, H. (2023). Masked Face Recognition System Based on Attention Mechanism. *Information*, 14(2), 87. Available on MDPI.
[54] Mukhiddinov, M., Djuraev, O., Akhmedov, F., Mukhamadiyev, A., & Cho, J. (2023). Masked Face Emotion Recognition Based on Facial Landmarks and Deep Learning Approaches for Visually Impaired People. *Sensors*, 23(3), 1080. Available on MDPI.
[55] Pudyel, M., & Atay, M. (2023). An Exploratory Study of Masked Face Recognition with Machine Learning Algorithms. *IEEE SoutheastCon 2023*, 877-882. Available on arXiv.
[56] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
[57] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*. Available at: https://arxiv.org/abs/1409.1556
[58] OpenGenus IQ. Understanding the VGG19 Architecture. Available at: https://iq.opengenus.org/vgg19-architecture/
[59] Fundus Image Classification Using VGG-19 Architecture with PCA and SVD. *Symmetry*. Available at: https://www.mdpi.com/2073-8994/11/11/1350
[60] Modi, K., & Ghayvat, H. (2021). CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronics*, 10(20), 2470. Available at: https://doi.org/10.3390/electronics10202470
[61] Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing*, 13(22), 4712. Available at: https://doi.org/10.3390/rs13224712
[62] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *NeurIPS*. Available at: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
[63] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*. Available at: https://arxiv.org/abs/1409.1556
[64] Datagen.tech. Understanding VGG16: Concepts, Architecture, and Performance. Available at: https://datagen.tech/understanding-vgg16-concepts-architecture-and-performance/
[65] ResearchGate. Very Deep Convolutional Networks for Large-Scale Image Recognition. Available at: https://www.researchgate.net/publication/275538560_Very_Deep_Convolutional_Networks_for_Large-Scale_Image_Recognition.
[66] Munduku Munduku, Deo "Neural Horizons: Comparison of Advanced Deep Learning Models for the Revolution in Breast Cancer Diagnosis" https://www.ijnrd.org/papers/IJNRD2404702 , 2024
[67] Varela-Aldás, J. et al. (2021). Facial Recognition System for People with and without Face Mask in Times of the COVID-19 Pandemic. Sustainability, 13(12), 6900. https://doi.org/10.3390/su13126900
[68] Hariri, W. (2021). Efficient Masked Face Recognition Method during the COVID-19 Pandemic. arXiv preprint. https://doi.org/10.48550/arXiv.2105.03026
[69] Yang, W., Wang, S., Hu, J., & Zheng, G. (2019). Security and Accuracy of Fingerprint-Based Biometrics: A Review. Symmetry, 11(2), 141. https://doi.org/10.3390/sym11020141.
[70] Okta (2021). Fingerprint Biometrics: Definition & How Secure It Is. Retrieved from https://www.okta.com.

Deo Munduku, the primary author of this article, is a researcher in artificial intelligence. He obtained his Master's degree in Computer Science from the University of Kinshasa in the Democratic Republic of Congo in 2017-2018. Since 2018, he has been teaching as a teaching assistant in the Department of Mathematics, Statistics, and Computer Science at the Faculty of Science and Technology at the University of Kinshasa. After his Master's degree, he continued his studies through a dual degree program between the University of La Rochelle in France and the National University of Vietnam, with the Institut Francophone International (IFI). His research areas include computer vision/deep learning, explainability of machine learning predictive models (XAI). Currently, he is also focusing on generative AI and neuromorphic computing with spiking neural networks. He has already published several articles in his field of research.

Jean Didier BATUBENGA M. Nz. holds a PhD in Computer Science (Pattern Recognition) from the University of Kinshasa in the Democratic Republic of Congo, obtained in 2019. He is a Professor and Secretary in charge of Teaching in the Department of Mathematics, Statistics, and Computer Science at the Faculty of Science and Technology at the University of Kinshasa.

Blaise Muhala obtained his Bachelor's degree in Computer Science from the University of Kinshasa in the Democratic Republic of Congo in 2016-2017. Since 2018, he has been teaching as a teaching assistant in the Department of Mathematics, Statistics, and Computer Science at the Faculty of Science and Technology at the University of Kinshasa. His field of expertise is information systems.

Boaz Nzazi obtained his Bachelor's degree in Computer Science from the University of Kinshasa in the Democratic Republic of Congo in 2013-2014. Since 2014, and a graduate of the National School of Administration (ENA RDC) since 2020, he has been teaching as a teaching assistant in the Department of Mathematics, Statistics, and Computer Science at the Faculty of Science and Technology at the University of Kinshasa. He also serves as a Civil Administrator of the State. His research area is deep learning.