[1] Karima BOUKARI

# Deep Learning-Based Iris Recognition System Using Unprocessed Images

Journal of Electrical Systems

***Abstract: -*** The iris image is a powerful and distinctive feature in biometrics that serves as a reliable instrument for human identification. Extracting significant features is crucial for developing iris-based recognition systems. While preprocessing for iris region detection is typically the first step in such systems, it can often fail in real acquisition conditions, leading to decreased performance. This study proposes a new iris-based recognition system that directly exploits original (noisy) images for feature extraction, avoiding the bottleneck of preprocessing. Additionally, a multimodal architecture is proposed for both right and left iris images to strengthen the decision step. Pretrained CNN models extract features classified by SoftMax and support vector machines (SVM). The performance of the proposed system is tested on four public datasets collected under different conditions. The results show that classification accuracy for the original iris dataset without pre-processing is higher than for the normalized database due to the rich CNN features that provide more cognitive information, resulting in greater discrimination power. The CNN multi-modal recognition system achieves the best accuracy compared to most state-of-the-art models, demonstrating the strength of the proposed fusion recognition system.

***Keywords:*** Deep learning and Feature extraction, Iris recognition using original images, Multimodal architecture.

## I. INTRODUCTION

The advent of the technology revolution has brought about significant changes to how we live, work, and learn and has enabled the expansion of Artificial Intelligence (AI) into a wide range of fields, including medicine, education, security, military, and commerce, through the utilization of massive data. Machine learning (ML), a branch of Al technology, emerged in the 1980s and is primarily aimed at analyzing large datasets with high precision and accuracy. Two primary approaches are used for this purpose: supervised learning, which involves the assistance of a data analyst or subject matter expert, and unsupervised learning, which does not. While Supervised learning requires more time and effort from experts, it allows them to learn extensively. In contrast, unsupervised learning is less time-consuming but may result in less valuable insights. Machine learning is currently being implemented in many applications, including person recognition, monitoring, and autonomous driving [1-4].

Deep learning (DL), a recent term that refers to a concept of learning based on Al is à subset of machine learning. It has attracted much attention due to recent advances in voice and image analysis and the feasibility of applications in many other fields that generate massive amounts of data. However, developing and implementing deep learning models pose challenges for multidisciplinary professionals. Convolutional Neural Networks (CNNs) [5-6], Feedforward Deep Neural Networks (D- FFNNs), Long-Term Memory networks (LSTMs), Deep Belief Networks (DBNs), and autoencoders (AEs) [7-11] are some of the approaches included in deep learning. These fundamental architectural components can be combined dynamically to create new network topologies specific to different applications.

Biometrics is becoming increasingly integrated into our daily lives as part of efforts to create a safer world. People must frequently identify themselves for various purposes, such as accessing their workplace or withdrawing money from an ATM, electronic voting, and securing bank payments or online transactions. Biometrics provides an alternative to passwords and other identifiers, removing any doubts about an individual's identity. Biometric systems authenticate a person by establishing the authenticity of a physiological trait and/or specific behavior that person possesses, such as facial recognition/fingerprints, and iris [12-13-14].

Iris identification is one of the best ways to provide individuals with single-factor authentication. The iris is a region of the human eye's ring between the pupil and the sclera that comprises a variety of textural properties. The iris texture provides a high degree of randomness and distinctiveness, which is unlikely to be unique in any of the two iris patterns, whether for identical twins or a person's left and right eyes [15-16]. However, the iris identification process is challenging, particularly in non-ideal environments. Various obstacles such as low resolution, distortion, scaling, and occlusion can cause difficulties in recognizing the iris.

---

[1] *Corresponding author: *Department of Electronics, Faculty of Engineering, Badji Mokhtar University Laboratory for Study and Research in Instrumentation and Communication Annaba (LERICA), Annaba, Algeria karima.boukari@univ-annaba.dz*

Deep learning, particularly Convolutional Neural Networks (CNNs), has been proposed as a solution to improve the feature extraction phase and achieve high accuracy in person identification. CNNs can automatically learn and extract the best features from input images. Transfer learning is also used to learn the features of some classical deep learning models and apply them to other tasks. Previous works in iris recognition applications have achieved high precision rates with a great deal of preprocessing. However, the authors propose to use deep learning to generate deep features from noisy databases like the CASIA-Iris-Lamp V4 image database without any preprocessing. They aim to evaluate the CNN features for two iris datasets and propose a multimodal recognition system with a two-input CNN structure. The main objective of this study is to investigate the performance of iris recognition by convolutional neural networks under non-ideal conditions and to propose a powerful system that can recognize noisy iris images without any preprocessing. The authors demonstrate that CNN vector characterization and classification are better for raw images without any preprocessing compared to those that have undergone all preprocessing steps. They also propose a bi-stream system with right and left iris inputs that uses only the original images in the bimodal recognition system. The two features are fused and fed into the SVM classifier. paper is organized as follows: Related work is described in Section 2, the proposed iris recognition system is detailed in Section 3, the experimental work analysis and results are presented in Section 4, and the conclusion and future work are drawn in Section 5.

## II. Related Works

Recently, computer vision researchers have shown significant interest in deep learning (DL), particularly convolutional neural networks (CNNs). Among all DL models, CNNs have achieved the best performance and accuracy in identifying and classifying objects. In a review article, Malgheet et al. [17] various techniques employed in the iris recognition system, including classical and deep learning approaches. They discussed the advantages and disadvantages of earlier strategies and the limitations of both traditional and deep learning iris recognition techniques.

In another study [18], a precise iris identification system was developed using pre-processing methods, segmentation with the Circular Hough transform il canny edge detector, and normalization with Daugman's procedure. The system used CNNS to extract information from normalized input iris images, and the Softmax classifier assigned the data to one of 224 classes from the IITD iris dataset and 108 classes from the CASIA V1 iris dataset. However, the accuracy rates of the system depend on the selection of hyperparameters and the tuning of the deep networks and optimizers.

Establishing the classification threshold for large-scale iris recognition jobs remains challenging due to the complexity of iris samples, particularly for real-world applications where the sample space is expanding rapidly. To overcome such threshold determination difficulties, the authors of [19] trained deep CNNS using à large number of iris samples to extract iris characteristics. Moreover, they resolved the issue of inadequate discrimination caused by the conventional Softmax loss function by introducing an improved center loss function known as Tight Center (T-Center) Loss.

In [20], researchers proposed an iris segmentation method that uses an "Interleaved Residual U-Net" (IRUNet) neural network model for semantic segmentation and iris mask synthesis. The algorithm estimates the inner and outer boundaries of the iris picture by recovering the inner border Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

## III. Deep Learning-Based Iris Classification without Prior Segmentation

Deep learning techniques are used to classify iris images directly, without any prior pretreatment. In contrast to prior works that classify segmented and normalized iris images, this approach focuses on the texture and morphology characterization of the iris eye. To extract features from the original iris image database, several transfer learning CNN architectures, including EfficientNet, Xception, ResNet50, VGG19, and MobileNetv2 are used. This approach avoids the segmentation step, improving the results for low-quality images. The iris region is extracted using the localization method, then normalized and enhanced using histogram equalization. For the normalized iris input image database, CNN features are extracted using the pre-trained convolutional neural network models mentioned above. Both SVM multiclass and Softmax classifiers are used for the classification step. To improve the system's robustness, the authors proposed a multi-input CNN architecture for the left and right iris eyes, as Daugman demonstrated that they differ from each other [16]. The system's performance is evaluated on several databases, including CASIA-Iris-Lamp V4, CASIA-Iris-Interval V3, ITDD, and MMU2.

### *A. Data pre-processing*

The pre-processing phase of the iris biometric recognition system is composed of several modules, including [26]:

#### *A.1 Elimination of white iris dots*

In most standard iris databases (CASIA iris database,2012; UBIRIS iris database, 2012; IITD iris database, 2012 and MMU iris database,2012), [27-30] the input eye image may contain white dots. If the white dots are not adequately eliminated they can compromise the iris localization process Inspired by [31], the authors proposed an effective system to suppress white dots, the result is shown in Fig.1.b.

#### *A.2 Median Filtering (Noise Removal)*

Generally, the acquired image contains some ‘irrelevant’ noise in addition to ‘useful’ information. The median filtering method is used to suppress the salt and pepper noise (Fig.1.c).

#### *A.3 Iris location (segmentation)*

The first stage of iris localization involves employing edge detectors to detect edges, followed by boundary detection techniques. The work steps of iris localization include detecting iris boundaries: pupillary boundary (or inner boundary) and outer boundary (or Limbic boundary). The inner boundary is so much darker and easy to detect. However, due to the low contrast outside the iris, the outer boundary is difficult to discern. By maximizing variations in the perimeter-normalized sum of gray level values along the circle, the outside boundary can be located. For this purpose, the Canny filter is applied to sharpen the edges. The resultant filtered image is then passed through à circular Hough transform to find the circular boundaries of the iris (Fig. 1.d).

#### *A.4 Pseudo-polar transformation*

Iris normalization is a step in which the iris is unwrapped to a rectangular ribbon. To map the circular iris region, a homogeneous Rubber sheet model is used (Fig.1.f).

#### *A.5 CLAHE adjustment*

The quality and contrast of an image can be refined by image enhancement, done to the above-normalized image. This will result in a well-distributed texture image. Contrast-limited adaptive histogram equalization (CLAHE) is a contrast adjustment method that produces an image with equally distributed intensity levels. In this paper, CLAHE is applied to the normalized images (Fig.1.g).

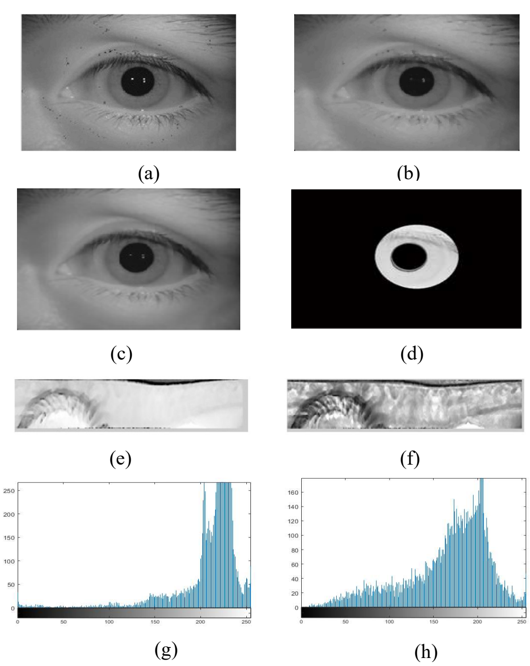



Figure 1. Iris image (a) original, (b) after removal of white spots, (c) after median filtering (3*3), (d) detection, (e) normalized, (f)after CLAHE enhancement, (g) histogram normalized and (h) histogram after CLAHE enhancement

*B. Proposed architecture*

The proposed unified iris recognition system is depicted in Fig.2. As à supervised biometric system, it requires a learning stage before the testing step. Our study specifically emphasizes the feature extraction step, which is illustrated in Fig.3. The recognition system involves two parallel CNN structures that take two input images of the left and right iris, respectively. The feature extraction blocks are used separately for each eye, and the resulting deep features are fused and used as input for the classifier mentioned below. In both the training and testing stages, extraction of CNN features is obtained using transfer learning for the two groups of images from used databases (Fig.4):

- PrepIrD (Preprocessed Iris image Database)
- Org-IrD (Original Iris image Database)

After the training step, testing CNN features are then classified using Softmax and multi-class SVM [32].

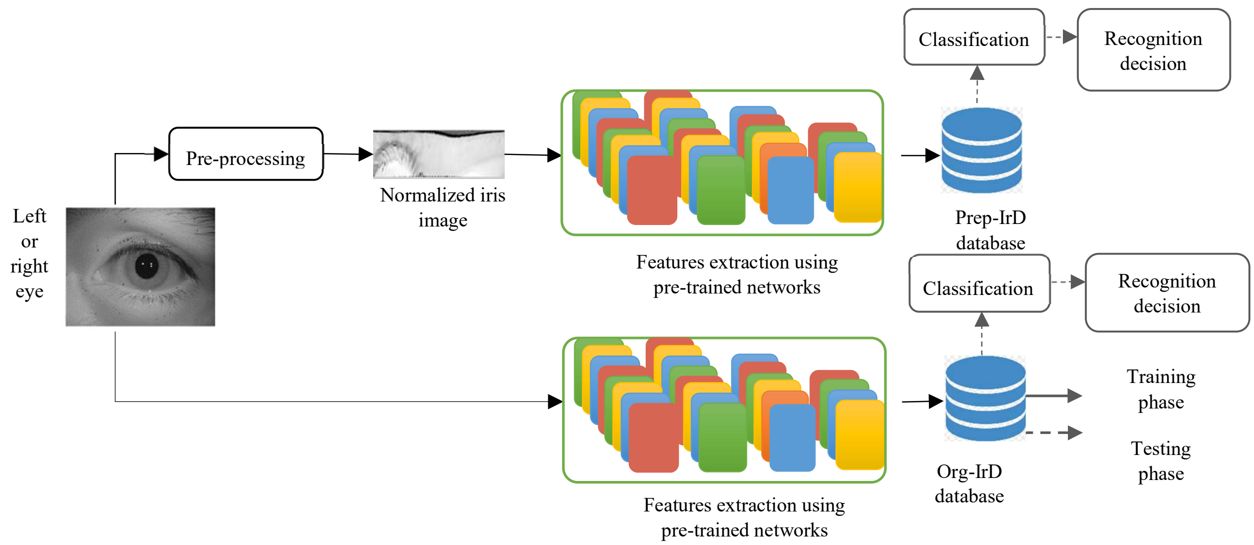



Figure 2. The unified iris recognition system

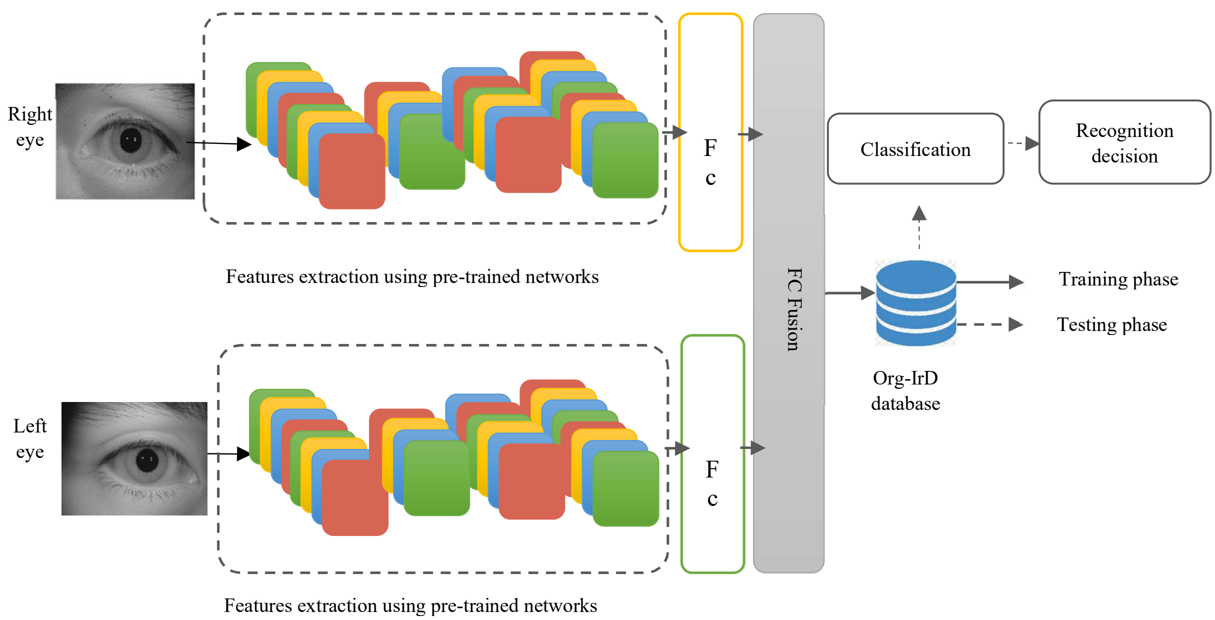
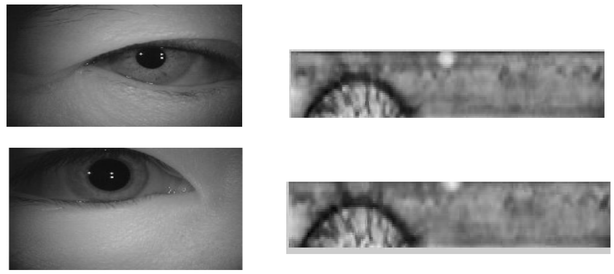



Figure 3. The proposed bi-stream iris recognition system



(a) (b)

Figure 4. Samples images from CASIA-Iris-Lamp V4
(a) Org-IrD and (b) Prep-IrD

### *C. Softmax classification*

The pre-trained model construction in deepest learning applications involves transfer learning. This involves starting with an existing network like EfficientNet or Xception, which has been previously trained on à specific set of classes. To classify iris images from the datasets mentioned above, some adjustments must be made to the deep CNN-based transfer learning networks. The proposed method is outlined in the next pseudocode:

Inputs: Iris images
1. Split dataset images into training and test databases
2. Import pretrained model
3. Input hyperparameters
4. Modifications made to the model
- the three last layers are removed
- a new FC layer, a new Softmax layer, and a new output Layer are added
5. The modified model is fine-tuned
-The training database is the input
-The labels on the training set are the targets
6. Test step on the test database
-Prediction class using Softmax
Output Accuracy performance

### *D. SVM Classification*

In the proposed study, a support vector machine (SVM) is utilized as a supervised learning algorithm for classification or binary regression tasks. The constructs an optimal hyperplane as a decision surface to maximize the separation margin between the two classes of data. Only a small proportion of observations, called support vectors, are

necessary to support the decision surface's optimal position. In this study, a multiclass SVM classifier with a linear kernel and without any optimization is employed for its speed. The proposed method is outlined in the next pseudocode:

Inputs: Iris images
1. Split dataset images into training and test databases
2. Import pretrained model
3. SVM is trained on a training database
- The FC layer output is the FC features
- The FC features are the input
- The labels on the training set are the targets
4. SVM is tested on the test database
Output Accuracy performance.

## IV. Experimental Evaluation and Results

### A. Databases

In the experiments carried out, two iris subsets from the databases were used, as explained in Tables 1 and 2. The Original Iris Image Database (Org-IrD) contains original left (L) and right (R) iris images (without any preprocessing). The Preprocessed Prep-IrD iris image database contains localized and normalized iris images of the same person chosen in Org-IrD; all iris images followed the same preprocessing steps detailed in Section 3.1. To ensure balanced data in the databases used, the same number of samples for the left and right iris images were taken. In the following sections, the focus will be on two cases in particular:

Experiment 1: Investigating the impact of pre-treatment on the system’s performance

Experiment 2: Examining the impact of feature fusion on the system's performance.

Table 1. The specifications of the adopted iris image datasets in Experiment 1

| | IITD | MMU2 | | CASIA-Iris-Lamp V4 | |
|---|---|---|---|---|---|
| Number of classes | 45 | 45 | 99 | 38 | 45 |
| Number of images | 450 | 450 | 990 | 760 | 900 |
| Samples per subject | 5 R and 5 L | 5 R and 5 L | | 10 R and 10 L | |
| Org-IrD image size | 320*240 | 320*238 | | 640*480 | |
| Prep-IrD image size | 48*432 | 32*240 | | 128*404 | |
| Image format | jpeg | Bmp | | Jpeg | |

TAble2. The specifications of the adopted iris image datasets in Experiment 2

| | IITD | MMU2 | CASIA-Iris-Lamp V4 | CASIA-Iris-Interval V3 |
|---|---|---|---|---|
| Number of classes | 208 | 99 | 200 | 120 |
| Number of images | 2080 | 990 | 4000 | 1680 |
| Samples per subject | 5 R and 5 L | 5 R and 5 L | 10 R and 10 L | 7 R and 7 L |
| Org-IrD image size | 320*240 | 320*238 | 640*480 | 320*280 |
| Image format | jpeg | bmp | jpeg | Jpeg |

### B. Deep learning architecture models

#### B.1 EfficientNet

EfficientNetBO is à deep convolutional neural architecture that was proposed in 2019 by Mingxing Tan and Quoc Le from Google AL. The EfficientNet architecture is efficiency in terms of model size and computational resources designed to achieve state-of-the-art accuracy on image classification tasks while maintaining a high degree of required for training and inference. The architecture consists of a novel compound scaling manner to find the optimal and resolution in a principled balance

between model size and method that uniformly scales all dimensions of depth, width, accuracy. Specifically, the architecture uses a combination of mobile inverted bottleneck blocks and squeeze-and-excitation blocks to improve feature representation and reduce the number of parameters. Efficient-Net achieved state-of-the-art performance in several benchmark image classification datasets including ImageNet, while being more efficient in terms of model size and computation compared to state-of-the-art models [33].

#### B.2 Xception

Xception is a convolutional neural network (CNN) architecture that aims to improve the performance of deep learning models by decoupling the learning of channel-wise and spatial-wise features. It employs an extreme version of Inception modules, which consist of à series of parallel convolutional layers with different filter sizes followed by à concatenation operation. Unlike traditional Inception modules, Xception replaces the conventional convolutions with depthwise separable convolutions, which consist of a depthwise convolution followed by à pointwise convolution. This modification significantly reduces the number of parameters and computations required while maintaining accuracy. Furthermore, the Xception architecture uses residual connections to facilitate gradient flow during training and prevent the vanishing gradient problem. These features enable Xception to achieve state-of-the-art performance on image classification tasks with a smaller number of parameters and faster training times compared to other CNN architectures [34].

*B.3 Resnet-50*

Residual Neural Network is a deep network architecture with 152 layers. This was developed by Microsoft in 2015. It won the ImageNet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015) with an error rate of 3.6%, which is considered better than human-level accuracy. The residual layers in ResNet compute changes in the input. This is then added to the input to produce the output. ResNet-50 is an early adopter of batch normalization. ResNet consists of two blocks CONV and Identity [35].

*B.4 VGG19*

VGGnet is a convolutional neural network proposed by K. Simonyan and A. Zisserman of Oxford University and gained notoriety by winning the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition in 2014. The model achieved an accuracy of 92.7% on Imagenet which is one of the highest scores achieved. It marked an improvement over previous models by proposing smaller convolution kernels (3×3) in the convolution layers than had been done previously. The model was trained over weeks using state-of-the-art graphics cards [36].

*B.5 MobiLeNetV2*

MobiLeNet is a class of lightweight deep convolutional neural networks that are much smaller in size and faster in performance than many other popular models. It uses depth-separable convolutions, which essentially means that it performs a single convolution on each color channel rather than combining all three and flattening it [37]

*C. Network Training Parameters*

The presented work was implemented using MATLAB R2021a on a CPU station with 4GB of memory capacity, and an Intel® Core™ i3-6006U CPU @ 2.00 GHz processor, with Windows 8.1 Professional with a 64-bit operating system.

The database is randomly devised in two parts: 80% for the training process and 20% for testing. A 5-fold cross-validation on this training set is used. For a better generalization of results, the 5 accuracy scores mean of the test folds are calculated.

To be in accordance with the model sizes presented in Table 1, iris databases Prep-IrD and Org-IrD have been resized. For the learning stage, it is necessary to increase and enrich the databases to improve the recognition performance. Additional augmentation operations have been performed on the input images which are random rotation of images along the vertical axis and random translation up to 30 pixels horizontally and vertically. The best hyperparameters were selected by making, a trade-off between the available resources, the learning time, and the accuracy which represents the rate of well-classified images. The training parameters of the suggested system are presented in Table

TABLE 3. ADOPTED PRETRAINED MODELS CHARACTERISTICS

| | AlexNet | GoogleNet | VGG19 | MobileNet | ResNet50 |
|---|---|---|---|---|---|
| Input size | 2227*227 | 224*224 | 224*224 | 32*32 | 224*224 |
| Size | 233 MB | 51 MB | 548 MB | 14 MB | 98 MB |
| Depth | 8 | 22 | 19 | 28 | 50 |
| Hyperparameters | 'sgdm' Optimizer/ Maximum number of epochs =30/ MiniBatch Size =10/ InitialLearnRate'=1e-4 | | | | |

*D. The effect of pretreatment by the performance of biometric recognition systems*

In the first experiment, the goal is to investigate the accuracy of iris image recognition for left and right images separately without any processing (Org-IrD) and with normalized iris images (Prep-IrD), as presented in Table 1. The classifier's accuracy is compared between raw and normalized images. It is important to note that all segmented and normalized images are properly localized, meaning that the external and internal borders of the iris are correctly identified.

The preprocessing phase follows the steps defined in Section 3.1 and is applied to the CASIA-LampV4, CASIA-Interval V3, and MMU2 databases. The normalized IITD database is also available for use. Two approaches were followed in this experiment:

The first approach involves using pre-trained CNN models (EfficientNet, Xception, ResNet50, and VGG19) for feature extraction and support vector machines (SVM) for classification.

The second approach involves applying transfer learning to (pre-trained CNN models (EfficientNet, Xception, ResNet50, and VGG19) for feature extraction and Softmax classification. The results of the experiments are presented in Table 4, which shows the accuracy obtained with different databases.

TABLE 4. ACCURACY EVALUATION USING SOFTMAX AND SVM CLASSIFIERS IN EXPERIMENT 1

| Databases | | Accuracy % by Softmax | | | | Accuracy % by SVM | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | ResNet50 | VGG19 | AlexNet | GoogleNet | VGG19 | ResNet50 | AlexNet | GoogleNet |
| IITD | Prep-IrD | 95.56 | 57.78 | 95.56 | 75.56 | 95.56 | 91.11 | 97.78 | 99.12 |
| | Orig-IrD | 97.78 | 97.78 | 97.28 | 95.56 | 99.01 | 99.20 | 99.04 | 99.23 |
| CASIA-Iris-Lamp V4 (38 classes) | Prep-IrD | 30.26 | 13.16 | 51.32 | 43.42 | 81.58 | 80.03 | 82.89 | 74.21 |
| | Orig-IrD | 96.68 | 98.68 | 99.01 | 99.03 | 96.05 | 94.74 | 98.16 | 95.79 |
| CASIA-Iris-Lamp V4 (45 classes) | Prep-IrD | 45.21 | 55.84 | 64.44 | 27.78 | 67.22 | 69.44 | 80.56 | 72.22 |
| | Orig-IrD | 99.44 | 97.24 | 98.89 | 97.22 | 95.56 | 96.67 | 97.22 | 96.67 |
| MMU2 (45 classes) | Prep-IrD | 81.11 | 89.78 | 78.89 | 38.89 | 88.89 | 80.00 | 92.22 | 88.89 |
| | Orig-IrD | 98.95 | 98.22 | 96.67 | 99.99 | 97.78 | 94.44 | 96.67 | 97.78 |
| MMU2 (99 classes) | Prep-IrD | 78.79 | 59.45 | 46.46 | 16.16 | 72.22 | 76.26 | 81.31 | 80.30 |
| | Orig-IrD | 96.78 | 96.96 | 96.97 | 62.85 | 92.93 | 89.90 | 93.94 | 94.95 |

The highest accuracy rate, reaching 99.12%, was achieved with the normalized images of the IITD database. In terms of comparing the SVM and Softmax classifiers, there is a small difference between them, except for Xception and VGG19, where the difference reaches 40%. VGG19 has the lowest scores, with a recognition rate of 13.16% for the CASIA-Lamp v4 database (38 classes).

When comparing the recognition rates between the two subgroups (Org-IrD and Prep-IrD), the difference between the rates is around 2% for most models, except for Xception and VGG19, where the difference reaches 85%. For the CASIA-Lamp V4 and MMU databases, the SVM classifier outperforms Softmax for normalized images. However, With an increase in the number of images to 900, Softmax performs better. It is worth noting that the recognition rates for original images are always better than those for normalized images, regardless of the model and classifier used for all the tested databases. To improve the performance of Softmax, increasing the size of the database using "data augmentation" or the optimal choice of hyperparameters is possible. However, this may result in a large computational cost. Due to hardware limitations, the learning step could not be performed.

The linear SVM used in this study remains fast and requires fewer computational resources, making it superior for relatively small databases. To improve its performance, kernels with parameter optimization can be used.

### *E. Performance analysis of Iris Recognition System using Multi-Modal Fusion of Features*

In the second experiment, the recognition accuracy of the unified recognition system for raw images of the left and/or right iris was compared to that of the bi-stream recognition system. The bi-stream recognition system uses two parallel CNN structures and extracts features from two blocks for the left and right iris eyes, generating two characteristic vectors. These vectors are fused and fed into the SVM classifier. Three types of feature fusions were tested: concatenation fusion (CF), sum fusion (SF), and maximization fusion (MF). The models were evaluated for performance using pre-trained CNN models such as EfficientNet, Xception, ResNet50, Mobilenetv2, and VGGI9 for feature extraction and SVM for classification in both the unified and bi-stream recognition systems. The accuracy results depicted in Figure 5 demonstrate that the fusion system performs significantly better than the uni-stream system across all tested databases and models. Among the uni-stream images, the CASIA-Interval V3 database achieves the highest recognition rates of up to 98.04% using the MobileNet model, while classification rates are similar for other models across all databases in the unistream system. The difference between uni-stream and bistream classification rates is notable, with a 15% gap observed in the IITD and MMU2 databases. Concatenation fusion outperforms the other two fusion types in all models and databases, reaching a rate of 99.51% for the MobileNet model on the IITD database. The MobileNet model is found to produce the best results, confirming the superiority of the biometric bi-stream system due to the robustness and complementarity of the FC characteristics of the left and right irises. To effectively evaluate the proposed method, à comparison is made with the results of Al-Waisy et al. [25] for the CASIA Interval V3 database under the same conditions as described in Table 2. The five models are tested with three types of fusions (concatenation sum, and maximization). Table 5 shows an improvement in the recognition rate for the proposed system, with an accuracy of 99.98% for concatenation fusion in the Mobilenet model. The training time depends on the model size, number of images, and hyperparameters. Although training hyperparameters are different, test time is crucial in the decision-making process, and Table 5 reports the shortest and longest test times based on the model size. The results indicate a better test time in both cases.

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

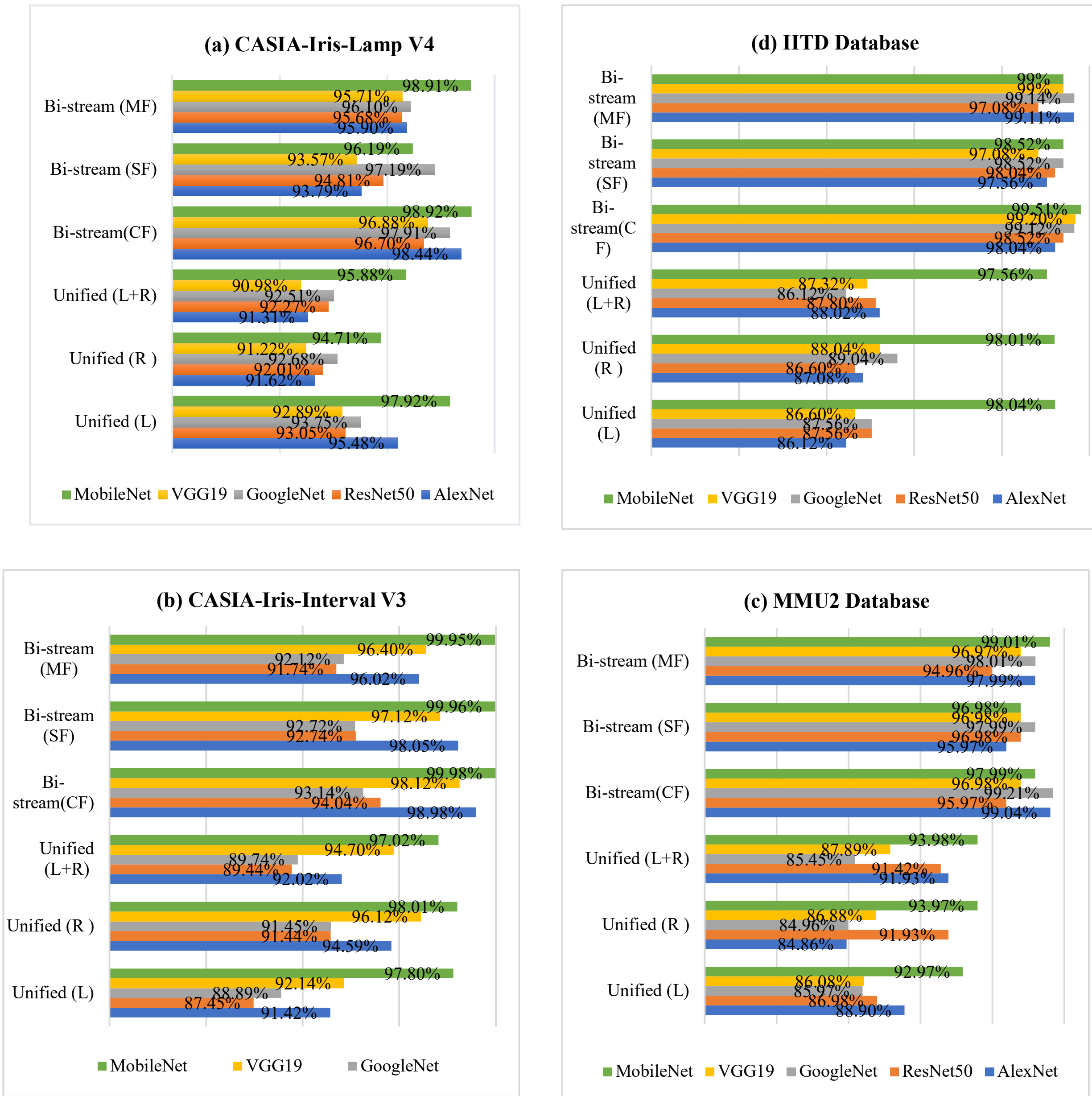



Figure 7. Performance comparison of the unified system and the proposed bitstream system

TABLE 5. PERFORMANCE COMPARISON OF THE PROPOSED RECOGNITION METHOD WITH A STATE-OF-THE-ART SYSTEM

| | **Al- Waisy et al. [25]** | **The proposed bi-stream system** |
|---|---|---|
| Used database | CASIA-Iris-Interval V3 | |
| # of classes | 120 | |
| Method | Multimodal CNN Softmax | Multimodal CNN SVM |

| | | |
|---|---|---|
| Accuracy | 99.82% | 99.98% |
| Testing time | 0.81 s | 0.22 s |
| Configuration system | MATLAB 2015, Windows 8.1, 1ntel Xeon E5-1620 CPUs , 16 GB of RAM. | MATLAB R2021a, Windows 8.1 Professional CPU 4GB RAM, inIntel® Core™ i3-6006U CPU @ 2.00 GHz |

## V. CONCLUSION

Deep learning networks enable direct learning of feature representation schemes from images. This work presents a fast and efficient iris recognition system without relying on a successful iris segmentation step, which may not always be possible under uncontrolled conditions. The system calculates CNN features and compares them for two iris image databases (Org-IrD and Prep-IrD). The accuracy for the Org-IrD subset is better without any processing, indicating the strength of the CNN features in characterizing iris images based on texture and morphology. The recognition rates for original images out-perform those of normalized images, regardless of the used model (EfficientNet, Xception, ResNet50, or VGG19) and classifier (SVM or Softmax), for all tested databases (CASIA- Lamp V4, IITD, and MMU2). In a second experiment, a multi-modal recognition system is proposed using different models (EfficientNetBO, Xception, ResNet50, VGG19, and MobileNetv2). The system employs two parallel CNN structures for the left and right iris eyes, which generate two characteristic vectors that are fused and fed to the SVM classifier. Three fusions of features (concatenation, sum, and maximization) were tested, and the proposed system was evaluated on several public datasets (CASIA-Lamp V4, CASIA-Interval V3, IITD, and MMU2). The system achieved satisfying results, especially with MobileNet. However, these results need to be generalized to larger databases. In future work, the authors plan to increase the number of database classes and test other SVM kernels to compare their classification performance. The system will also be tested on other databases with noisy images encountered in unconstrained recognition situations such as off-angle, distance, etc.

## VI. REFERENCES


[1] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi & M. Dehmer, « An introductory review of deep learning for prediction models with big data", Frontiers in Artificial Intelligence, 3, 4, 2020.

[2] A. Dhillon & G. K. Verma, " Convolutional neural network: a review of models, methodologies, and applications to object detection", Progress in Artificial Intelligence, 9(2), 85-112, 2020.

[3] S. Sony, K. Dunphy, A. Sadhu & M. Capretz, "A systematic review of convolutional neural network-based structural condition assessment techniques", Engineering Structures, 226, 111347, 2021.

[4] I. Stancic, L. Veic, J. Music, T. Gr Grujic, "Classification of Low-Resolution Flying Objects in Videos Using the Machine Learning Approach", Advances in Electrical and Computer Engineering 22(2):45-52, May 2022, DOI:10.4316/AECE.2022.02006

[5] J. Jiao, M. Zhao, J. Lin & K. Liang, "A comprehensive review on convolutional neural network in machine fault diagnosis". Neurocomputing, 417, 36-63, 2020.

[6] S. Hao, F. X. Ge, Y. Li & J. Jiang, "Multisensor bearing fault diagnosis based on one-dimensional convolutional long short-term memory networks". Measurement, 159, 107802, 2020.

[7] J. Shao, Y. Niu, C. Xue, Q. Wu, X. Zhou, Y. Xie & X. Zhao, "Single-channel SEMG using wavelet deep belief networks for upper limb motion recognition", International Journal of Industrial Ergonomics, 76, 102905, 2020.

[8] J. Yu & G. Liu, "Knowledge extraction and insertion to deep belief network for gearbox fault diagnosis". Knowledge-Based Systems, 197, 105883, 2020.

[9] D. Yu, M. Kolbæk, Z. H. Tan & J. Jensen, "Permutation invariant training of deep models for speaker-independent multi-talker speech separation", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 241-245, March 2017.

[10] T. Zhou, S. Ruan, S. Canu, "A review: deep learning for medical image segmentation using multi-modality fusion", Array, 3–4, p. 100004, 2019

[11] A. Essien & C. Giannetti, « A deep learning model for smart manufacturing using convolutional LSTM neural network autoencoders", IEEE Transactions on Industrial Informatics, 16(9), 6069-6078, 2020.

[12] S. Sangwan and R. Rani, "A Review on -Iris Recognition", (IJCSIT) International Journal of Computer Science and Information Technologies, vol. 6, pp. 3871-3873, 2015.

[13] T. Thomas, A. George & K. I. Devi, "Effective iris recognition system. Procedia Technology", 25, 464-472, 2016.

[14] A. K. Jain, A. Ross & S. Prabhakar, "An introduction to biometric recognition", IEEE Transactions on circuits and systems for video technology, 14(1), 4-20, 2004.

[15] J. G. Daugman, "New methods in iris recognition", IEEE Trans. Syst. Man Cybern. 37(5), 1167–1175, 2007.

[16] J. G. Daugman, "How Iris Recognition Works", IEEE Transactions on circuits and systems for video technology, Vol. 14, No. 1., 2004.
[17] J. R. Malgheet, N. B. Manshor, and L. S. Affendey, " Iris Recognition Development Techniques: A Comprehensive Review", Wiley Hindawi, Article review ID 6641247, 32 pages, 2021.
[18] S. Sujana, D. V. Reddy, "An Effective CNN-based Feature Extraction Approach for Iris Recognition System", Turkish Journal of Computer and Mathematics Education Vol.12 No.6, 4595-4604, 2021.
[19] Y. Chen, C. Wu & Y. Wang "T-center: a novel feature extraction approach towards large-scale iris recognition", IEEE Access, 8, 32365-32375, 2020.
[20] Y. H. Li, W. R. Putri, M. S. Aslam & C. C. Chang, « Robust iris segmentation algorithm in non-cooperative environments using interleaved residual U-Net", Sensors, 21(4), 1434, 2021.
[21] K. Nguyen, C. Fookes, S. Sridharan & A. Ross, "Complex-valued Iris Recognition Network", arXiv preprint arXiv:2011.11198, 2020.
[22] G. Liu, W. Zhou, L. Tian, W. Liu, Y. Liu & H. Xu, « An efficient and accurate iris recognition algorithm based on a novel condensed 2-ch deep convolutional neural network", Sensors, 21(11), 3721, 2021.
[23] K. Nguyen, C. Fookes, A. Ross & S. Sridharan, "Iris recognition with off-the-shelf CNN features: A deep learning perspective", IEEE Access, 6, 18848-18855, 2017.
[24] H. Moyasar Therar, Lect. Dr. E. Ahmed Mohammed and A. Jadaan Ali, "Multibiometric System for Iris Recognition Based Convolutional Neural Network and Transfer Learning", IOP Conf. Series: Materials Science and Engineering 1105 (2021) 012032, 2021.
[25] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, T. Nagem, "A multi-biometric iris recognition system based on a deep learning approach". Pattern Anal Applic Pattern 21:783–802, 2018.
[26] S. Tounsi And K. Boukari "The impact of collarette region-based convolutional neural network for iris recognition International Journal of Electrical and Computer Engineering Systems", IJECES Vol. 13 No. 1, 2022.
[27] Casia iris database, http://www.idealtest.org/findTotalDbByMode.do?mode=Iris, last accessed April 22, 2012.
[28] UBIRIS iris database, http://iris.di.ubi.pt/, last accessed April 22, 2012
[29] IITD iris database, http://www.iitd.ac.in/, last accessed April 22, 2012.
[30] MMU iris databases, http://pesona.mmu.edu.my/~ccteo/, last accessed April 22, 2012.
[31] T. Kyaw, K. T. Swe, P. T. Z. Tun, "Performance analysis of features extraction on iris recognition system" IJCIRAS | ISSN (O) - 2581-5334 | Vol. 2 Issue. 12, May 2020.
[32] X. Qi, T. Wang, J. Liu, "Comparison of support vector machine and softmax classifiers in computer vision". Second Int. Conf. on Mechanical Control and Computer Engineering, 2017.
[33] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," arXiv preprint arXiv:1905.11946, 2019.
[34] H. Zhao, Z. Li, L. Fang, and T. Zhang, "A Balanced Feature Fusion SSD for Object Detection", Neural Process Lett 51, 2789–2806 doi.org/10.1007/s11063-020-10228-5, 2020.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", in Proceedings of the IEEE Conference on computer vision and pattern recognition, pp. 770–778, 2016.
[36] E. Hassan, Y. Khalil, and I. Ahmad, "Learning feature fusion in deep learning based object detector", Journal of Engineering, Hindawi, 2020.
[37] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications". CoRR, abs/1704.04861, 2, 4, 5, 6, 2017.