Library Progress International *Print version ISSN 0970 1052*
Vol.44 No.3, Jul-Dec 2024: P.5573-5580 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Multimodal Image Fusion for Feature Extraction based on Biometric Person Identification

Namdeo D. Kapale[1], Agarkar Balasaheb S.[2]

[1]Department of Electronics and Telecommunication Engineering, Sanjivani College of Engineering, Kopargaon, Savitribai Phule Pune University, Pune, Maharashtra, India, Email: kapalenamdeoetc@sanjivani.org.in
[1]Department of Electronics and Telecommunication Engineering, Sanjivani College of Engineering, Kopargaon, Savitribai Phule Pune University, Pune, Maharashtra, India, Email: bsagarkar977@gmail.com
*ndkapale@gmail.com

How to cite this article: Namdeo D. Kapale, Agarkar Balasaheb S. (2024) Multimodal Image Fusion for Feature Extraction based on Biometric Person Identificatione. *Library Progress International*, 44(3), 5573-5580.

Abstract
The method of identifying a person using behavioral attributes, such as fingerprints, eye structures, or facial characteristics is known as biometric person identification. The procedure utilizes complex algorithms to analyze and compare biometric data against recorded templates. The purpose of this research is to develop a novel multimodal image fusion model for feature extraction based on biometric person identification. To build an intelligent biometric person identification framework, we provide a new Versatile Convolutional Neural Network (VCNN) technique for obtaining vital characteristics using multimodal data. The features gathered from modalities are combined and provided into the (Multi-Kernel Support Vector Machine) MKSVM classifier for classification by combining ensemble learning for classification with deep knowledge for extracting characteristics; this method offers precise biometric person identification. Initially, we gathered a Dataset that contains two biometrics, such as the Faces94 database for face images and fingerprints gathered fromFVC2002, to train our suggested model. We employed histogram equalization to pre-process the gathered raw data, it improves the quality of the obtained data. The finding assessment phase evaluates the suggested model performance using various metrics. We compared the recommended technique to various existing methods to assess its effectiveness. The experimental findings demonstrate that the recommended model outperforms conventional approaches.

Keywords: Multimodal Image Fusion, Feature Extraction, Biometric Person Identification, Versatile Convolutional Neural Network (VCNN), Multi-Kernel Support Vector Machine (MKSVM), Face and Fingerprint.

1. Introduction:

At present, biometric data collection is an essential part of authorization and safety systems[1].A person who can be correctly traced using biometric systems that depend on unique physical or attitudinal indicators such as an individual's face, genetic makeup, or eye routines. Because they were simple to use and unobtrusive, biometrics based on images were frequently used. Fortunately, conventional single-modal biometric technologies have drawbacks such as being susceptible to outside interference, having a narrow precision lineup and readily tricked [2]. A common solution to those problems was the integration of several biometric segments, which raised the biometric identification precision. One technique to combine data from several biometric inputs into one format was called verbal vision merging. Due to its potential to increase the precision and robustness of biometric identity, that kind of technology has drawn lot of interest. Since merging enhanced data from several dimensions, multilingual fusion enhanced dependability, resilience to deceit attacks, and biased abilities in a range of operational scenarios [3].One of the numerous uses for versatile merging was to increase the lifespan and accuracy of biometric human being authentication by combining a variety of biometric features.
To extract biometric features for identity verification, it examined multisensory image fusion. The focus was using

the extra features of different biometric approaches to derive distinctive characteristics that allow accurate and reliable human being recognition[4].A complete representation of a person's biometric characteristics that go beyond individual paradigms may be achieved by combining a variety of modalities, including face explosions, monitoring, eyeball frameworks and cognitive dynamics. Multimodal image fusion for biometric person identification was based on the extraction of robust and discriminative features from unified representations of disparate dimensions. Conventional knowledge recovery strategies may operate beside the bounds of certain techniques, overlooking crucial information that was accessible in other settings. Versatile synthesized leveraged supporting data gathered through many approaches to provide more comprehensive and personalized attribute extraction [5].

Increasing awareness of bidirectional fusion is driven by accuracy, especially when separate approaches are limited a low standard. Even while recognition systems for faces perform best in unrestricted environments, they may be affected by obstructions or fluctuations in light[6]. On the contrary, fingerprint-based machine learning was very precise yet susceptible to faking. While identification and visage modalities were merged, the resultant approximation can mitigate the disadvantage of each modality, increasing identification ruggedness and effectiveness. Multimodal communication convergence can be utilized to address the issue between class differences commonality that exists in the single-modal biometric data approach. Multimodal fusion allows a detailed characterization of an individual by combining biometric data such as hand prints, lens designs and facial traits. That improved separation capability and reduced the possibility of incorrect mismatches.

Multimodal image fusion was a compelling method for biometric identification of individuals using extracted feature information. Fusion-based approaches can extract resilient and prejudicial characteristics that optimize identification accuracy, toughness and security by merging data from numerous biometric detectors[7].To build an intelligent biometric person identification framework, we provide a new Versatile Convolutional Neural Network (VCNN) technique for obtaining vital characteristics using multimodal data. The features gathered from both modalities are combined and provided into the (Multi-Kernel Support Vector Machine) MKSVM for classification and this method provides accurate biometric person identification by integrating ensemble learning for classification and deep knowledge for extracting attributes.

2.Related work

The study[8]suggested an approach for human identification using Composite deep learning, Dynamical Sine Evolve, and Lagrange's interpolation-based image processing. The method attained 96.4% accuracy and halved the overall dimension of the data set. Their method provided biometric-based systems with a trustworthy, efficient and secure solution.

The article [9]suggested a method for identifying individuals based on health-state classifications of electrocardiogram (ECG) rhythms, with an emphasis on arrhythmia classifications. With a lightweight CNN, it achieved remarkable accuracy in recognizing (99.28%) both arrhythmic and normal heart rates, testing on nine types of arrhythmias.

The author [10] researched a biometric system with multiple modes that used Gabor filters with unimodal features to use fingerprinting and retinal data. The system's higher recognition ratio and higher rate of false rejection than earlier methods of prospective use for multimodal biometrics in security, privacy and forensics.

The article [11] explored the use of human body odor as a biometric trait for person identification, highlighting its unique chemical composition and potential for non-intrusive, reliable and cost-effective authentication.

The study [12] explored the use of ear biometrics in deep learning algorithms for real-time person identification. The YOLOV5 model performed better than other models about individual identification precision and size (16MB) using a standard dataset.

The author [13] proposed a deep learning model to identify individuals using one's voice, utilizing unique individual characteristics. The model achieved 99.81% accuracy and a loss function of 0.009 over the VoxCeleb1 dataset for 40 subjects.

The paper [14]suggested a profound convolutional neural network design with a total of six layers for ear recognition. The network attained detection scores of 98.42% and 95.88%, correspondingly, using the IITD-II and AMI ear datasets. During COVID-19 outbreaks, the system can be used for passive person identification.

The paper [15] researched that the internet's increased connectivity necessitated robust cybersecurity measures, including biometric systems like fingerprints, speech, and facial recognition. As threats became more endemic, traditional security systems were becoming more vulnerable. Biometric systems reduce time and manpower, increasing efficiency and increasing the need for biometric security.

3. Methodology

In this section, the methodology involves developing a multimodal image fusion model using a Versatile Convolutional Neural Network (VCNN) for biometric person identification. Features from facial and fingerprint data are combined and fed into the Multi-Kernel Support Vector Machine (MKSVM) classifier. Dataset acquisition includes Faces94 and FVC2002, pre-processed using histogram equalization. Figure 1 shows the overview of methodology.

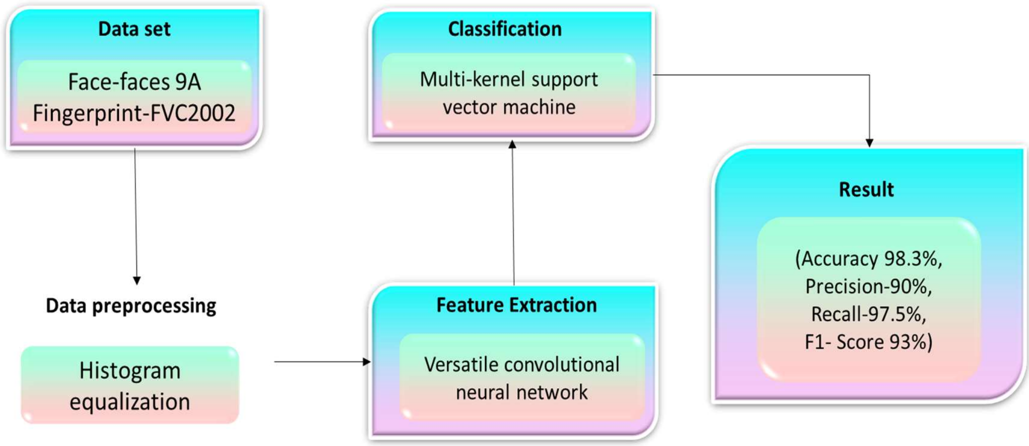

Figure 1:Overview of methodology

3.1 Dataset

The suggested method makes use of two biometrics: the fingerprint and face. Multimodal biometrics were taken into consideration to improve the accuracy of authenticity.

a) Images of faces were gathered from the Faces94 database, which has images of 80 subjects (40 male participants and 40 female participants). Of these, 40 subjects (20 male participants and 20 female participants) had eight images chosen from each. Thus, a total of 210 images were selected.

b) Images of fingerprints were gathered from the FVC2002 (Fingerprint Verification Competition 2002) database. This database includes images of 80 people, 15 images of each person, from 20 people, 8 images of each person, were chosen, for an overall total of 160 images[16].

3.2 Data pre-processing

Data preprocessing for biometric person identification using histogram equalization involves analyzing the distribution of pixel intensities in biometric data (such as fingerprint and facial images) to enhance contrast, normalize intensities, and extract relevant features, aiding in accurate identification and classification.

3.2.1. Histogram equalization

Afacial image$J(A,B)$ that has Npixels and a total of s grey levels for example, a 4-bit image would have 256 grey levels. The goal of histogram equalization is to increase the global contrast of the images by altering the image's pixel density values' allocation J(a,b) into a uniform distribution. It does by dispersing the common pixel intensity values by better distributing the pixel intensity values. Histogram equalization has the following formal definition: Considering the likelihood $p(j) = \frac{ki}{K}$ (that is, the real $J(a,b)$histogram of a pixel that occurs with a grey level of a, where $a \in 0,1,\ldots,k-1$ and ni. The diagram from a given intensity value to a_{new} changedwhichis described and indicates the number of pixels in $J(a,b)$ with the grey level value of s

$$s_{new} = \sum_{a=0}^{k-1} ki/K = \sum_{a=0}^{k-1} p(i) \quad (1)$$

The process of mapping the original range of pixel intensity values (0–256) to the domain of [1,0] is defined by equation (1). Therefore, the values new need to be rescaled to acquire pixel values in the original domain, such as the 4-bit interval.

3.3Versatile convolutional neural network

The Versatile Convolutional Neural Network (VCNN) is a flexible architecture designed for diverse data types and modalities. It integrates convolutional layers with adaptable structures, allowing efficient feature extraction from various inputs. VCNN enhances performance in tasks such as image classification, object detection and

biometric identification through its versatile design.
The Versatile Convolutional Neural Network using Dropout and the Stochastic Gradient Descent (SGD)Optimizer is constructed in conjunction with the CNN framework, the Defective ReLU activation function, and the adaption avoidance approach based on ejection on the Stochastic Gradient Descent (SGD) where it has the subsequent basic operations.

Step 1: Set the filter size pixel to $V1 \times V2$ and pretrain the filter.
Step 2: Load the training image dataset. Receive the information to generate the imagery grid X after processing the training set image to fit the chosen filter dimension.
Step 3: To start parallel operations,Set the weight $(w(l)i, j)$ and biased (bi) initial. Then, call TensorFlow's kernel function def Kernel().

Step 4: Conv2d is used to perform the two-dimensional convolution operation to generate the initial layer combining information array A.

Step 5: The input data for the pool layer is derived from the convolution characteristic array A (1) of the prior level. To obtain the attribute array A for the pool operation.
Step 6: Utilizing Tensor Flow's weights and the freshly skewed interface, create a feature matrix by updating the weighting $wi$and bias parameter biA. Use the Stochastic Gradient Descent optimizer function to get the learning rate for the hierarchy-based refining optimization.
Step 7: Proceed with steps 4, 5, and 6 to construct the second stage of convergence, which will yield the feature matrix A.
Step 8: Combine the feature matrix $A\ (4)$ into a column vector to feed the neurone at the full-joint layer. Multiply by both the bias and the weight matrix. To create an eigenvector $A\,(4)$,utilize the ruptured ReLU activation coefficient.
Step 9: The eigenvector linked to the completely linked layersis used as the dropout layer's feed and then used to calculate the neuron's output probability, where the eigenvector a1 is produced.
Step 10: To get the desired outcomes, use the SoftMax classifier's output and the eigenvector a1 as the input.

3.4 Multi kernel Support vector machine

This method provides accurate biometric person identification by combining deep learning for characteristic extraction with ensemble learning for classification using the MKSVM classifier. This method's main goals were to decrease the gap between a hyper plane attention and bridge the gap between classes. This suggested classifier is justified by the fact that has various combinations among these attributes may need various interpretations of similarity (different portions). Taking into that the section undoubtedly involves a nonlinear change, considerations for the practical kind of adjustment that causes information to become swiftly detached were superfluous. The non-linear operation in the SVM classification is carried out by initiating the kernel functions. This work employs the radial bias function (RBF) and linear kernel functions for the classification procedure.
These categorization methods were the support vectors with c acting to a constant and δ acting to a kernel parameter in the end.

(a) Linear kernel function

Given that the classes can be divided sequentially, minimum of a hyperplane that divided into the classes error-free may be found and is represented by a biassed vector. The simplest and fastest-processing kernel is$V$, which isthe fundamental kernel.

$$Lin(x, y)xVy + z \qquad (2)$$

(b) RBF Kernel

The work makes use of a function called the radial basis function or Gaussian kernel to help support vector machine classification. Because exterior judgements were randomized in relation to the test vector, the identification of a test matrix can benefit from the discovery of unconnected evaluations.

$$RBF(s, t)\exp\ (\delta \,||\, s\ + t\ ||), \qquad \delta < 0 \qquad (3)$$

It's assumed that employing several kernels works to create a better categorization is an effective approach.It is simply obtained as an inclined union of many kernels. The benefits using multi-kernel operations (2) and (3) isthe internal product of primary activity that has to be described ina fresh digital field. The categorization work is found in the conditions below.

$Multikernal = \frac{2}{1}\ (Lin(x, y) - RBF(x, y))$ (4)

The hyperplane is classified using the efficient and optimized findings. A Kernel's judgement depends on the problem at hand since it depends on the model. The aforementioned multi-kernel approach characterizes the facts using the optimal features. To categorize data classifiers in general have development and evaluation steps.

4.Experimental results

The efficiency of the proposed VCNN-MKSVM strategy is investigated by a comparative evaluation of the outcomes. The efficacy of a suggested method is demonstrated by contrasting its accuracy and efficiency with combining the ensemble learning for classification with deep knowledge for extracting characteristics, the features collected from both modalities are combined and supplied into the Multi-Kernel Support Vector Machine (MKSVM) classifier for classification. This method provides accurate biometric person identification. The efficacy of the suggested and current methods (Multimodal Biometric Feature Extraction ((MBFE), CNN, VGG19 with softmax classifier (VGG 19-SC))[17]) was assessed using metrics such as the F1 score, precision, accuracy, and recall.

The percentage of all cases that were accurately detected is known as accuracy. In the context of a biometric person identification, it displays the proportion of individuals who have been correctly recognized (founded on their biometric information) out of all the individuals in the collection. The accuracy is calculated using equation (4).

$Accuracy = \frac{TP+TN}{Totalsamples}$ (4)

- TP-It represents for perfected accuracy predictions or accuracy that exceeds validation.
- TN: The negative predictive value corresponded to the testing level.
- FP: When the expected samples correspond to the values and validation level at the same degree
- FN-When the precise values, which were the validation stage, varied at various points from the items tested that projected.

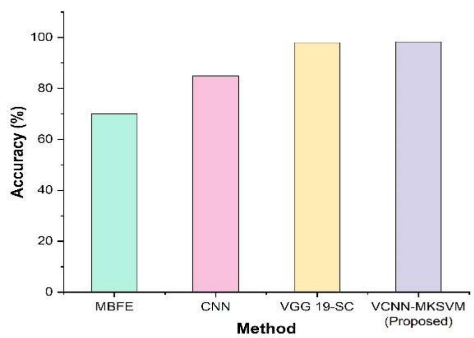

Figure 2: Result of accuracy

Table 1: Worth of accuracy, precision, recall and F1- core

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
MBFE	70	74	50	66
CNN	85	67	65	55
VGG19-SC	98	88	97	92
VCNN-MKSVM (Proposed)	98.3	90	97.5	93

Table 1 and Figure 2 provide the result of accuracy values. Our suggested technique outperformed MBFE, CNN, VGG19-SC in comparison, with a 98.3% improvement over VCNN- MKSVM. The proposed approach VCNN-MKSVM, shown notable gains in Biometric person identification as compared to the current methods.

The measurement ratio is the precision proportion of positive to all authentic instances. The sum of the real positive designs and the one-sided models represents the total actual samples. True positives are called TPs, and false positives are called FPs.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

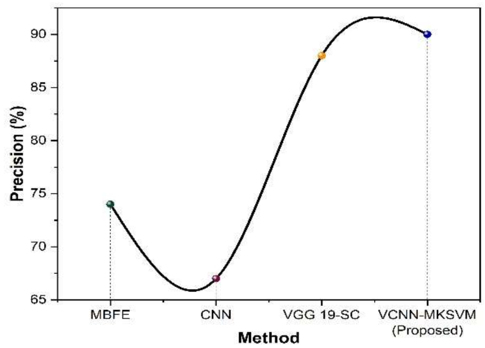

Figure 3: Result of precision

The precision result and value are displayed in Table 1 and Figure 3, respectively. Our suggested technique outperformed MBFE, CNN, and VGG19-SC in comparison, with a 90% improvement over VCNN-MKSVM. The proposed approach VCNN-MKSVM, shown notable gains in Biometric person identification as compared to the current methods.

Evaluation is defined as the recall ratio of the positive genuine trials divided by the sum of good genuine trials and all false-negative true samples. The false negative is represented by FN and True Positive by TP.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

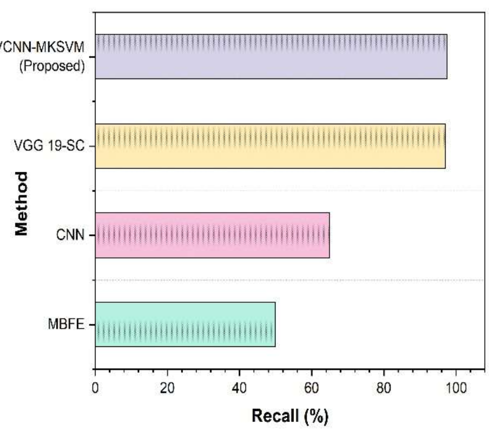

Figure 4: Result of recall

Table 1 and Figure 4 exhibit the recall outcome. Our suggested technique outperformed MBFE, CNN and VGG19-SC in comparison, with 97.5% improvement over VCNN-MKSVM.The proposed approach VCNN-MKSVM, shown notable gains in Biometric person identification as compared to the current methods.

The cumulative total of precision as well as recall represents the F1 score. It compensates for both false positives and false negatives, achieving an equilibrium among recall and precision. By utilizing this equation (7) to compute the F1score, it was possible to minimize false positives and negatives in the event of an uneven class allocation.

$$F1 - Score = 2 * \frac{(precision * recall)}{(precision + recall)} \quad (7)$$

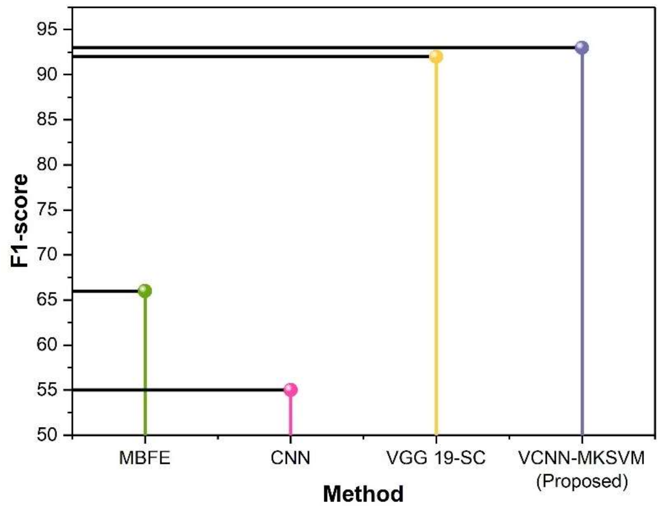

Figure 5: Result of F1-Score

The F1 score value is shown in Table 1 and Figure 5. Our suggested technique outperformed MBFE, CNN, and VGG19-SC in comparison, with 93% improvement over VCNN- MKSVM. The proposed approach VCNN-MKSVM shown notable gains in Biometric person identifications compared to the current methods.

5. Conclusion

In conclusion, our research introduces a novel Versatile Convolutional Neural Network (VCNN) for biometric person identification, leveraging multimodal data fusion from facial and fingerprint biometrics. Employing a Multi-Kernel Support Vector Machine (MKSVM) classifier with ensemble learning and deep feature extraction, our model achieves precise identification. We utilized Faces94 and FVC2002 datasets, enhancing data quality through histogram equalization preprocessing. Experimental results showcase superior performance compared to existing methods, achieved accuracy (98.3%), precision(90%), recall (97.5%) and F1- score (93%) affirming the efficacy of our approach in advancing biometric identification technologies, promising in real-world applications. Versatile Convolutional Neural Network (VCNN) may face challenges in handling complex features from diverse biometric data. Multi-Kernel Support Vector Machine (MKSVM) computational complexity may limit scalability in large datasets, impacting real-time processing. Future research could explore enhancements to Versatile Convolutional Neural Network (VCNN) for more robust feature extraction and investigate optimization techniques for Multi-Kernel Support Vector Machine (MKSVM) to improve biometric person identification accuracy and efficiency.

References

1) Singh, G., Bhardwaj, G., Singh, S.V. and Garg, V., 2021. Biometric identification system: security and privacy concern. Artificial intelligence for a sustainable industry 4.0, pp.245-264.
2) Zhang, X., Yao, L., Huang, C., Gu, T., Yang, Z. and Liu, Y., 2020. DeepKey: A multimodal biometric authentication system via deep decoding gaits and brainwaves. ACM Transactions on Intelligent Systems and Technology (TIST), 11(4), pp.1-24.
3) Bala, N., Gupta, R. and Kumar, A., 2022. Multimodal biometric system based on fusion techniques: a review. Information Security Journal: A Global Perspective, 31(3), pp.289-337.
4) Dargan, S. and Kumar, M., 2020. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. Expert Systems with Applications, 143, p.113114.
5) Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B. and Chanussot, J., 2022. Deep learning in multimodal remote sensing data fusion: A comprehensive review. International Journal of Applied Earth Observation and Geoinformation, 112, p.102926.

6) Kuruvayil, S. and Palaniswamy, S., 2022. Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. Journal of King Saud University-Computer and Information Sciences, 34(9), pp.7271-7282.
7) Parvathy, J. and Patil, P.G., 2024. FRMSDNET Classifier for Multimodal Feature Fusion Biometric Authentication. International Journal of Intelligent Systems and Applications in Engineering, 12(6s), pp.169-186.
8) Jadhav, S.B., Deshmukh, N.K. and Pawar, S.B., 2024. Robust Authentication System with Privacy Preservation for Hybrid Deep Learning-Based Person Identification System Using Multi-Modal Palmprint, Ear, and Face Biometric Features. International Journal of Image and Graphics, p.2550049.
9) Al-Jibreen, A., Al-Ahmadi, S., Islam, S. and Artoli, A.M., 2024. Person identification with arrhythmic ECG signals using deep convolution neural network. Scientific Reports, 14(1), p.4431.
10) Amin, P., Murugan, R. and Gupta, M.V., 2024. RELIABLE PERSON IDENTIFICATION USING A NOVEL MULTIBIOMETRIC IMAGE SENSOR FUSION ARCHITECTURE.
11) MANJU, M., HUMAN BODY ODOUR AS A BIOMETRIC INDICATOR FOR PERSON IDENTIFICATION.
12) Hossain, S., Anzum, H. and Akhter, S., 2024. Comparison of YOLO (v3, v5) and MobileNet-SSD (v1, v2) for Person Identification Using Ear-Biometrics. International Journal of Computing and Digital Systems, 15(1), pp.1259-1271.
13) AL-Shakarchy, N.D., Obayes, H.K. and Abdullah, Z.N., 2023. Person identification based on voice biometric using deep neural network. International Journal of Information Technology, 15(2), pp.789-795.
14) Ahila Priyadharshini, R., Arivazhagan, S. and Arun, M., 2021. A deep learning approach for person identification using ear biometrics. *Applied intelligence*, *51*(4), pp.2161-2172.
15) Rai, V., Mehta, K., Jatin, J., Tiwari, D. and Chaurasia, R., 2020, May. Automated biometric personal identification-techniques and applications. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1023-1030). IEEE.
16) Vijayalakshmi, G.V. and Mohana, C., 2006. A Multimodal Biometric Recognition System Based on Decision Level Fusion for User Authentication.

Amin, P., Ganesh, D., Gantra, A. and Singhal, P., 2024. Improved Human Identification by Multi-biometric Image Sensor Integration With a Deep