Inteligencia Artificial 27(74), 37-47
doi: 10.4114/intartif.vol27iss74pp37-47

INTELIGENCIA ARTIFICIAL

http://journal.iberamia.org/

Age-Invariant Cross-Age Face Verification using Transfer Learning

Newlin Shebiah Russel, Arivazhagan Selvaraj, Dhanya Devi S, Dhivyarupini M

Centre for Image Processing and Pattern Recognition,
Department of Electronics and Communication Engineering,
Mepco Schlenk Engineering College, Sivakasi, India

Abstract : The integration of face verification technology has become indispensable in numerous safety and security software systems. Despite its promising results, the field of face verification encounters significant challenges due to age-related disparities. Human facial characteristics undergo substantial transformations over time, leading to diverse variations including changes in facial texture, morphology, facial hair, and eyeglass adoption. This study presents a pioneering methodology for cross-age face verification, utilizing advanced deep learning techniques to extract resilient and distinctive facial features that are less susceptible to age-related fluctuations. The feature extraction process combines handcrafted features like Local Binary Pattern/Histogram of Oriented Gradients with deep features from MobileNetV2 and VGG-16 networks. As the texture of the facial skin defines the age related characteristic the well-known texture feature extractors like LBP and HoG is preferred. These features are concatenated to achieve fusion, and subsequent layers fine-tune them. Experimental validation utilizing the Cross-Age Celebrity Dataset demonstrates remarkable efficacy, achieving an accuracy of 98.32%.

Keywords: Cross-age face verification, Local Binary Pattern, Histogram of Oriented Gradients, Transfer Learning, MobileNetV2, VGG-16

1 Introduction

Face recognition represents a biometric identification technology emerged as a pivotal tool in identity authentication and finds extensive utility across various domains, including law enforcement, identity verification processes, and security measures. The primary differentiation between face verification and facial recognition is in the technology's selection of either one-to-one or one-to-many matching. Facial verification is employed with the objective of ascertaining the authenticity of an individual's claimed identity. The process of face verification involves the identification and quantification of facial characteristics inside an image. Over time, several causes such as the emergence of wrinkles, weight gain, and the proliferation of facial hair, the utilisation of spectacles, and other elements can induce notable alterations in the texture and contour of human faces. Hence, face verification poses a challenging problem. The utilisation of face verification is prevalent in the process of authenticating individuals' identities.

Cross-age face verification has gained significant research attention as it plays a crucial role in scenarios such as locating missing children after a considerable passage of time or identifying individuals with a substantial gap between images, often relevant in apprehending long-absconded criminals. Cross-age face verification is primarily concerned with the task of identifying a person across different time points based on images captured at distinct ages. Cross-age face verification utilizing deep features involves the application of deep learning models to extract distinctive facial characteristics from images of individuals spanning various age brackets. These extracted features are then harnessed to ascertain the legitimacy of two images, discerning whether they depict the same individual or not.

The general framework of Cross-Age Face Verification (CAFR) encompasses a series of steps: data collection to gather facial images spanning diverse age groups, preprocessing to standardize and align these images, feature extraction utilizing deep learning models like CNNs to capture distinct facial features, optional age estimation to

ISSN: 1137-3601 (print), 1988-3064 (on-line)

assess age disparities, verification by comparing the feature vectors of two faces using similarity metrics, threshold setting to determine whether the faces belong to the same individual, evaluation to gauge system performance, optimization for accuracy and efficiency, and eventual deployment in applications like security systems and age progression analysis, making CAFR an essential tool for accurate recognition across varying age ranges. Deep learning architecture finds application in person re-identification using Siamese Networks [1], video classification [2], recognizing motion-blurred CCTs based on deep and transfer learning [3], classification of breast cancer [4], plant leaf classification [5 , 6] , transfer learning for multiple-step-ahead forecasts in monthly time series [7]. Figure 1 shows the general framework of Cross Age Face Verification architecture.

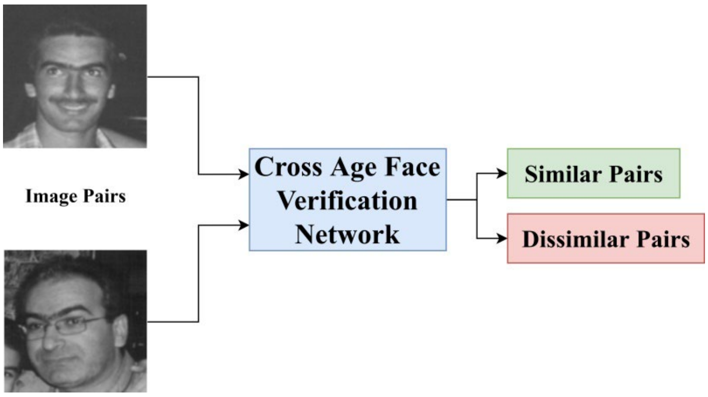

Figure 1 : Framework for Cross Age Face Verification

Our key contributions can be summarized as follows:

- This paper introduces a pioneering methodology for cross-age face verification that combines handcrafted and deep learning features to extract resilient and age-invariant facial characteristics.
- Extensive experimentation using benchmark datasets validates the effectiveness of the proposed approach, achieving notable accuracy rates.

The proposed methodology offers practical implications for real-world applications such as law enforcement, access control, and identity verification processes, where acurate cross-age face verification is crucial.

2 Literature Survey

This section provides a comprehensive overview of methodologies employed for the purpose of cross-age face identification and verification. The utilization of deep generative model-based networks is widespread in the field of synthesis methods. Zhang et al. [8] proposed a conditional adversarial autoencoder model that was specifically developed to capture the underlying structure of facial features. This model allows for accurate prediction of facial age progression and regression. An further significant development emerged from the proposition of a pyramid design for generative adversarial networks within the context of an age progression model [9]. This innovative approach guarantees that the generated facial images display the intended aging effects while preserving consistent personalized features. In their study, Huang et al. [10] proposed a comprehensive multi-task learning framework called MTLFace, which aimed to address the challenge of Cross-Age Face Recognition (CAFR). The framework successfully achieved the simultaneous goals of generating age-invariant identity-related representation and synthesizing faces.

The deep features obtained from facial images across different age groups often consist of two unique categories of information: qualities connected to age and characteristics particular to facial identity. As a result, cross-age discriminative models have been developed with the purpose of separating identity-dependent components from the retrieved facial traits. The authors Chen et al. [11] introduced a novel coding technique called cross-age reference coding, which involves encoding an image with respect to a cross-age reference point in order to achieve a feature representation that is invariant to age. The cycle age-adversarial model proposed by Du et al. [12] was designed to extract age-invariant features and utilize age labels solely for training objectives. In relation to the Cross-Age Face Recognition (CAFR) problem, Huang et al. [13] introduced the Age-Puzzle FaceNet (APFN) model, which utilizes an adversarial training approach. This model was further improved to boost its compactness and ability to handle age variations (Huang et al., [14]).

Several innovative approaches have been developed to address the intricate task of Cross-Age Face Recognition (CAFR), drawing inspiration from the remarkable achievements of deep learning in the field of Face Recognition (FR). The deep Convolutional Neural Network (CNN) model proposed by Wen et al. [15] demonstrates proficiency in acquiring age-invariant deep features. The model successfully accomplished this task by properly separating identity information from aging information, utilizing a unique loss function specifically designed for preserving identity. The employment of a pre-trained deep face recognition (FR) model within a Siamese architecture employing the contrastive loss function was pioneered by Bianco et al. [16]. The utilization of this approach facilitated the transference of information obtained from the FR job to the complex realm of CAFR.

In their study, Wang et al. (17) introduced a multi-task model that demonstrates the ability to simultaneously address age estimation and face recognition (FR) tasks. In the context of face recognition (FR), a Siamese network was implemented and trained utilizing the contrastive loss function for the purpose of recognition, and the cross-entropy loss function for age estimation. The authors successfully implemented a novel deep-learning architecture [18] by combining an identity network and an aging network, which were designed to utilize the same feature layers. In their study, Wang et al. [19] proposed a novel approach called Orthogonal Component Convolutional Neural Network (OE-CNN) to breakdown deep features into orthogonal components, specifically targeting identification and aging information. The isolated characteristics of identity were subsequently employed for the purpose of identification. The authors Zhao et al. [20] introduced a comprehensive deep architecture that was specifically developed to effortlessly carry out cross-age synthesis and recognition in a holistic manner. The authors Li et al. [21] proposed a method for age-invariant feature learning using convolutional neural networks (CNNs), which includes a novel optimization technique for distance metrics. The strategy proposed by Zhao et al. [22] involves the utilization of Generative Adversarial Networks (GAN) to mitigate the influence of aging components on the acquired face features.

In their study, Du et al. [23] enhanced the effectiveness of pre-trained face recognition (FR) models through the implementation of a transfer learning methodology. This involved extracting distinctive features from a pre-trained generator while simultaneously suppressing age-related information using a discriminator. The technique proposed by Wang et al. [24] involves the utilization of a decorrelated adversarial learning algorithm to decompose identity-dependent components from aging components. The maximum correlation between paired features generated by a deep convolutional neural network (CNN) was determined using this method. The CNN employed a factorization module and a backbone network that were trained to reduce the correlation between identity and aging features. Furthermore, Huang et al. [25] introduced an age-adversarial convolutional neural network (AA-CNN) that encompasses both an age-discriminative network (ADN) and an identity-recognition network (IRN) in a seamless manner. The adopted methodology utilized adversarial loss in order to train the Age-Disentangled Network (ADN), hence enabling the Invariant Representation Network (IRN) to efficiently acquire age-invariant features.

Li & Lee [26] introduces the Age-Invariant Features Extraction Network (AIFEN), an attention-based feature decomposition model that effectively reduces age interference and learns discriminative representations. AIFEN achieves superior performance on benchmark datasets, outperforming existing methods with relative improvements on CACD-VS, AgeDB, CALFW, and LFW datasets. Cross-age facial image datasets often lack supervised data due to challenges in collection, hindering the effectiveness of age-invariant face recognition methods. To overcome this limitation, a novel semi-supervised approach named Cross-Age Contrastive Learning (CACon) is proposed, leveraging face synthesis models to generate additional samples. CACon introduces a new contrastive learning method and associated loss function, achieving state-of-the-art performance in homogeneous-dataset experiments and surpassing existing methods by a large margin in cross-dataset experiments [27]. Hast [28] investigation examines how the implementation of Embedded Prototype Subspace Classifiers can enhance face recognition accuracy in the presence of age-related variations, solely relying on face feature vectors, amidst the challenge of age progression affecting facial characteristics.

The aforementioned methodologies typically emphasise the feature learning phase, neglecting considerations for the execution and integration of the classification and verification stages. Hence, these processes encompass a dual-stage procedure that may potentially lead to the elimination of certain intermediate data. In this paper, we consider the construction of a discriminative feature learning network. By employing cross-age domain adversarial training, our model is able to generate cross-age face representations that are explicitly separated from age variations.

3 Proposed Methodology

In this section, a novel approach for cross-age face verification is presented, utilizing advanced deep learning methodologies to extract resilient and distinctive features that are less susceptible to age-related fluctuations. The proposed cross-age face verification network is detailed, and the processing mechanism and model structure are illustrated in Figure 2.

As illustrated in Figure 2, the proposed model takes a pair of images as input. It undergoes a series of steps, including feature extraction, feature fusion, feature difference computation, and multiple layers of processing. This process allows the extraction of features from both images. The final classification layer determines the outcome by evaluating the feature of identity difference, which is categorized into two types: intra-class difference (similar pair) and inter-class differences (dissimilar pair). "During the feature extraction process, a combination of handcrafted features, such as Local Binary Pattern/Histogram of Oriented Gradients, and deep feature approaches including MobileNetV2 and VGG-16 network, are utilized to extract the desired features. These handcrafted and deep features are then combined through concatenation, leading to feature fusion. Subsequently, the features are fine-tuned by subsequent layers within the network. Finally, the classification layer evaluates the similarity or dissimilarity of the images based on the extracted features.

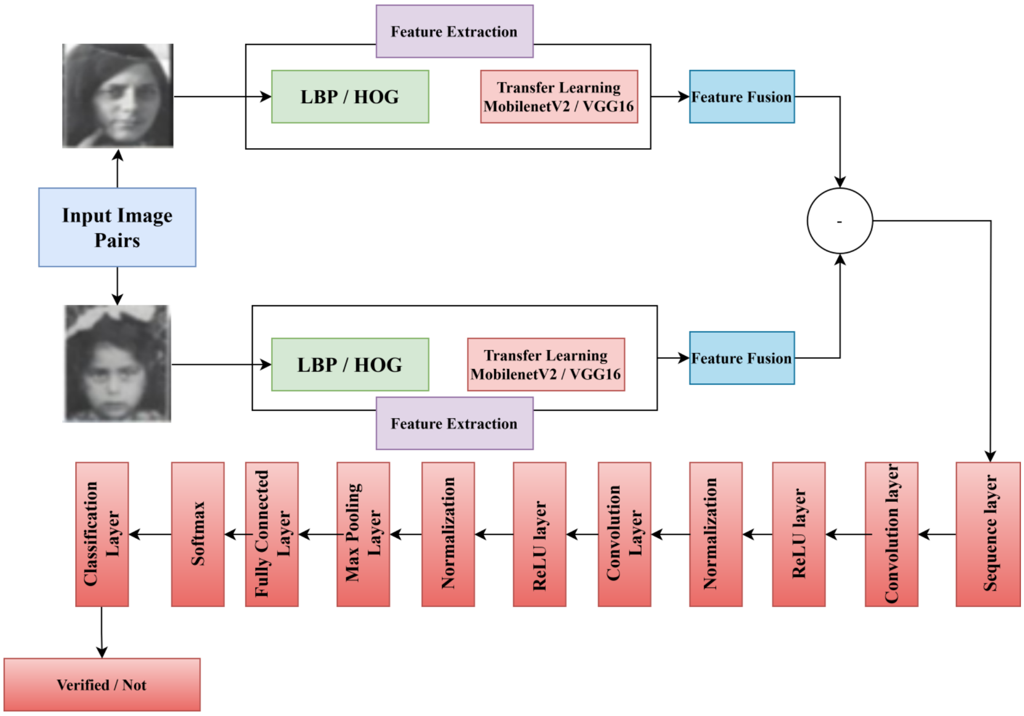

Figure 2. Block diagram for verification using handcrafted features and transfer learning.

3.1 Feature Extraction Using LBP

The Local Binary Pattern (LBP) is a very effective operator utilised for the characterization of local image features. By considering the centre pixel (x_c, y_c) as the reference point, the LBP algorithm generates an ordered binary set by the comparison of the grey value of the centre pixel (x_c, y_c) with its eight neighbouring pixels. The LBP code is represented as a decimalized octet integer.

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^n \tag{1}$$

where *ic* represents the centre pixel (x c, y c) ,and *in* the pixels of its eight neighbors.

$$S(i_n - i_c) = \begin{matrix} 1 & i_n - i_c \geq 0 \\ 0 & i_n - i_c < 0 \end{matrix} \quad (2)$$

3.2 Feature Extraction Using HOG

The Histogram of Oriented Gradients (HOG) is a feature descriptor commonly employed in the fields of image processing and computer vision, particularly for object recognition tasks. This descriptor captures image features based on gradient orientations. The algorithm used for HOG extraction records edge orientations within specified regions of an image.

The computation of HOG features involves several key steps. Initially, the image is divided into cells, and for each cell, gradients in the x and y directions are calculated to determine their magnitudes and orientations. These orientations are then grouped into histogram bins, covering specific ranges. Cells are further organized into blocks, allowing for local gradient normalization to accommodate variations in illumination and contrast. The final HOG descriptor is formed by concatenating the normalized histograms from these blocks. This comprehensive representation effectively captures shape and gradient information and finds wide application in computer vision tasks such as object detection and recognition.

3.3 Feature Extraction Using MobilenetV2

The MobileNetV2 model was utilized for image categorization, emphasizing improved mobility. Rooted in the framework of its predecessor, MobileNetV1, MobileNetV2 incorporates the Depthwise Separable Convolutions (DSC) technique to enhance portability. This approach tackles information loss in non-linear layers by employing Linear Bottlenecks within convolution blocks. Furthermore, MobileNetV2 introduces a unique architectural design called Inverted Residuals, preserving vital information throughout the network layers. Depthwise Separable Convolutions, as utilized in MobileNetV2, are a combination of Depthwise convolution and Pointwise convolution. This integration significantly reduces the total number of parameters and computing cost, approximately to 18% of that required by regular convolutions. In MobileNetV2, the structure starts with a fully convolutional layer comprising 32 filters, followed by 19 bottleneck layers. These bottlenecks involve a depth-separable convolution with residual connections. MobileNetV2 improves model efficacy across various workloads and benchmarks. It achieves this by replacing conventional convolutional layers with depthwise convolutional blocks. These blocks consist of a 3 x 3 depthwise convolutional layer for input filtering and a 1 x 1 pointwise convolutional layer for combining filtered values into new features. While the outcomes are similar to conventional convolution, MobileNetV2 demonstrates significantly higher speed. The model comprises 53 convolutional layers and 1 average pooling layer. Initially, MobileNetV1 used standard 3x3 convolutions, followed by 1x1 depthwise separable convolutional blocks. The concept of inverted residuals involves bottleneck blocks with similarities to residual blocks, including initial input, multiple bottlenecks, and subsequent expansion. Introducing shortcuts between these bottlenecks enhances gradient propagation in multiplier layers due to two main factors: firstly, most information resides within bottlenecks, and secondly, the expansion layer can be seen as a tensor transformed non-linearly. The use of the inverted architecture, or Inverted Residual, reduces memory consumption compared to the standard structure.

3.4 Feature Extraction Using VGG16

The VGG-16 network features a well-defined architecture and exceptional classification performance, making it a straightforward framework to extend for various applications. This architecture incorporates a max-pooling layer with a 2 x 2 filter size and a stride of 2. The convolutional layers employ 3 x 3 filters with a stride of 1, consistently applying the same padding. This arrangement of convolution and max-pooling layers is maintained throughout the design.

VGG-16 consists of 16 levels inside the network architecture, each with associated weights, totaling an impressive parameter count of over 138 million. The initial and subsequent convolutional layers use 64 kernel filters, each with a size of 3 × 3, modifying the input image dimensions to 224 × 224 × 64. The third and fourth convolutional layers feature 128 kernel filters, each with a size of 3 x 3. After two layers and a max-pooling operation with a stride of 2, the output size reduces to 56 x 56 x 128. The fifth, sixth, and seventh layers employ 256 feature maps with 3 x 3 kernel size, followed by a max-pooling layer with a stride of 2.

Layers eight to thirteen consist of two sets of convolutional layers with 512 kernel filters each, using a 3 x 3 kernel size. A max-pooling layer with a stride of 1 follows these layers. The fourteenth and fifteenth layers are fully connected hidden layers with a total of 4096 units. VGG-16 utilizes smaller receptive fields compared to AlexNet, employing 33 units in a single stride. The presence of three Rectified Linear Units (ReLU) enhances the network's discriminative capacity. The network concatenates handmade and deep features, a process known as feature fusion, to achieve a comprehensive representation of the input image. These features are further refined by subsequent layers through feature tuning.

In the context of neural networks, the initial layer, referred to as the sequence input layer, receives sequential data. Data normalization techniques are applied in the input layer to preprocess feature data. One-dimensional convolutional procedures are applied to input data, and the rectified linear unit (ReLU) function is used for activation. Normalization techniques mitigate input fluctuations, while the pooling layer reduces spatial dimensions while preserving essential features. Fully connected layers capture intricate patterns in the input data and create an abstract representation. The softmax layer converts the preceding layer's output into a probability distribution across predefined classes. The classification layer then determines the similarity or dissimilarity of images, indicating whether they are comparable or distinct.

4 Results and Discussion

In this section, a sequence of experiments is carried out on the extensively used cross-age database, the Cross-Age Celebrity Dataset (CACD) [29 - 32]. These experiments aim to assess the efficacy of the proposed methodology.

4.1 Dataset

The Cross-Age Celebrity Dataset comprises a comprehensive assemblage of images depicting celebrities at various stages of their lives. The dataset is frequently employed in the assessment of algorithms pertaining to tasks such as age estimation, face recognition, and face verification. The dataset CACD comprises a total of 163,446 images sourced from the Internet, featuring 2,000 distinct celebrities. The age range of these individuals spans from 14 to 62 years. The visual content was acquired from publically accessible sources, including movie stills, red carpet events, and social media accounts. Samples of images from the dataset are shown in Figure 2.

Figure 2. Sample Images from CACD Database.

4.2 Cross Age Face Recognition Results with Handcrafted and Transfer Learning Features

Images are resized to 224x224 pixels enabling the feature extraction by transfer learning using MobilenetV2 / VGG16. The proposed architecture accepts image pairs and for experimentation two dataset sizes are used: a smaller one with 2,010 pairs each for training and testing, and a larger one with 40,209 pairs each for training and testing.

4.2.1 Local Binary Pattern (LBP)

The LBP feature vector is a 1-by-N vector that encodes local texture information in images. It divides the input image into non-overlapping cells and in this experimentation the dimensionality of feature vector is 1×59. With LBP features, the model achieved a testing accuracy of 62.45% for both smaller and larger set. The reason for low recognition accuracy may be the LBP features might not capture highly discriminative information, especially when dealing with subtle age-related changes. Texture-based methods, such as LBP, face difficulties when distinguishing faces with similar textures but different expressions, poses, or ages. Age-related changes, like wrinkles or fine lines, might not significantly alter the overall texture, making it hard for LBP to differentiate the person between age groups. The 1×59 feature vector might not capture all the relevant facial information necessary for accurate age-invariant recognition. More complex and higher-dimensional feature representations might be required to effectively handle the variability in facial appearances across different ages.

4.2.2 Histogram of Oriented Gradients (HOG)

HOG offers several advantages as a feature extraction method for face verification. It excels at capturing both the shape and appearance details of a face, proving particularly valuable in scenarios with diverse lighting conditions or varying camera angles. Additionally, HOG is computationally efficient, making it a practical choice for many applications. With a feature set comprising 26,244 features, significantly more than LBP, HOG demonstrates robustness in handling complex facial patterns. In experiments using the CACD dataset, the model was trained with 2,010 pairs and tested with another 2,010 pairs, achieving an impressive training accuracy of 94.03% and a testing accuracy of 80.91%. Subsequently, when trained with a larger dataset of 40,209 pairs and tested with an equivalent number of pairs from CACD, the training accuracy surged to an impressive 99.31%, accompanied by a testing accuracy of 95.98%. These results underline the effectiveness of HOG in face verification tasks, especially when dealing with varying lighting conditions and camera angles.

4.2.3 MobileNetV2

In this study, transfer learning was applied using the MobileNetV2 model, focusing on its 'Logits' layer with a feature size of 1000. A batch size of 32 was chosen for the process. Initial validation with 2010 training pairs and a distinct 2010 testing pair, leading to a training accuracy of 97.96% and a testing accuracy of 71.36%. For large image pairs comprising 40,209 training image pairs and an additional 40,208 face image pairs, maintaining the same feature size and batch size, were used for testing, resulting in a training accuracy of 93.56% and a testing accuracy of 85.09%.

4.2.4 VGG16

The implications of utilizing transfer learning with VGG16 model are analysed. The 1000 features were extracted from FC8 layer of VGG16. A batch size for experimentation is chosen to be 32 and the number of epochs is fixed as 20. With the limited image pairs for training and testing the proposed algorithm achieves training accuracy rate of 94.03% and a testing accuracy rate of 73.89%. With large 40,209 training image pairs, characterized by identical feature size and batch size the model achieves a training accuracy rate of 87.31% and a testing accuracy rate of 80.18%.

In analyzing the results several key observations can be made regarding the different methods employed for face verification. Firstly, the HOG method, despite its higher feature count (26,244), demonstrated strong performance with a training accuracy of 94.03% and a robust testing accuracy of 80.91%. On the other hand, the utilization of transfer learning with the VGG16 model, although offering a reduced feature set (1,000), showcased competitive training accuracy (94.03%). However, its testing accuracy slightly dropped to 73.89%, indicating a potential limitation in generalization. MobileNetV2, another transfer learning model with 1,000 features, exhibited an impressive training accuracy of 97.96%, but its testing accuracy fell to 71.36%, suggesting challenges in handling real-world variability. Interestingly, combining MobileNetV2 and HOG features, despite increasing the feature count (27,244), resulted in a lower training accuracy of 93.88% and a testing accuracy of 73.50%. This observation hints at the complexity of integrating different feature extraction methods. In summary, the HOG method stood out with its high testing accuracy, emphasizing its effectiveness in face recognition tasks, while the transfer learning methods, while powerful in training, faced challenges in achieving comparable testing accuracy.

Table 4.1: Accuracy with less number of image pairs

Features	No of training pairs	No of testing pairs	Feature dimension	Training accuracy	Testing accuracy
HOG	40209	40208	26244	99.31	95.98
LBP			59	62.95	62.45
MobilenetV2			1000	93.56	85.09
VGG-16			1000	87.31	80.18
MobileNetV2 and HOG			27244	94.52	90.39

4.2.5 Training with Variable Feature Counts

In the MobileNet V2 model, training at different layers had significant implications on the model's performance and feature dimensions. Initially, training at the Logits layer with 1000 features and a dataset of 2010 pairs led to a high training accuracy of 97.96%. However, the testing accuracy was comparatively lower at 71.36%, indicating a challenge in generalizing the learned patterns to unseen data. Expanding the dataset to 40,209 training and testing pairs maintained the same feature dimension but remarkably improved the testing accuracy to 85.09%, demonstrating the positive impact of a larger dataset on the model's ability to generalize. Furthermore, training on the out_relu layer, which had a higher feature dimension of 62,720, resulted in exceptional performance. The model achieved perfect training accuracy (100%) and an impressive testing accuracy of 98.32%, underscoring its robust learning capabilities and strong generalization to new data.

Table 4.2: Performance Metrics of MobileNet V2 Model at Different Layers for Face Verification

MobileNet V2(layers)	Feature dimension	No.of training pairs	No of testing pairs	Training accuracy	Testing accuracy
Logits layer	1000	2010	2010	97.96	71.36
		40,209	40,208	93.56	85.09
out_relu layer	62,720	2010	2010	100	98.32

These findings highlight the importance of both the layer choice and the dataset size in enhancing the accuracy and reliability of the MobileNet V2 model in face recognition tasks.

Table 4.3: Comparison with the State of the art Methods

Method	**Accuracy (%)**
LFCNN [30] (2016)	98.5
OE-CNN [31] (2018)	99.20
DAL [32] (2019)	99.40
IEFP [33] (2022)	99.57
Attention [26] (2022)	99.63
Proposed Method	98.32

Our proposed method achieved an accuracy of 98.32%, which, while slightly lower than the state-of-the-art methods such as Attention with 99.63% accuracy, demonstrates competitive performance compared to previous techniques such as LFCNN (98.5%), OE-CNN (99.20%), DAL (99.40%), and IEFP (99.57%), highlighting its effectiveness in addressing age-related variations in face recognition tasks.

4.3 Discussion

The findings presented in the previous section shed light on the performance of various feature extraction methods and transfer learning models for cross-age face verification. These results offer valuable insights into the strengths and limitations of each approach, as well as their implications for real-world applications.

The comparison between Local Binary Pattern (LBP) and Histogram of Oriented Gradients (HOG) highlights the importance of feature robustness in handling age-related changes. While LBP exhibited limited accuracy due to its inability to capture subtle variations in facial texture, HOG proved to be more effective in capturing both shape and appearance details, resulting in significantly higher testing accuracies across different dataset sizes. This underscores the importance of selecting feature extraction methods that can adequately capture the intricate details of facial characteristics, particularly in the presence of age-related variations.

The utilization of transfer learning with MobileNetV2 and VGG16 models offered promising results, albeit with varying degrees of success. MobileNetV2, despite achieving high training accuracy, faced challenges in generalization, as evidenced by its lower testing accuracy compared to HOG. Similarly, while VGG16 demonstrated competitive training accuracy, its testing accuracy lagged behind, indicating potential limitations in handling real-world variability.

The comparison of different methods and models underscores the complexity of cross-age face verification tasks and the need for robust and adaptive approaches. While HOG emerged as a strong contender, offering high testing accuracies even with smaller datasets, the findings suggest avenues for further improvement, particularly in enhancing the generalization capabilities of transfer learning models.

5 CONCLUSION

In conclusion, this study addresses a crucial challenge in the field of face verification by tackling age-related disparities, which have long posed significant obstacles to accurate recognition systems. Leveraging cutting-edge deep learning techniques, our methodology pioneers a novel approach to cross-age face verification. By extracting resilient and distinctive facial features that are less susceptible to age-related fluctuations, we have significantly enhanced the reliability and accuracy of face verification technology. Our approach uniquely combines handcrafted features such as Local Binary Pattern (LBP) and Histogram of Oriented Gradients (HoG) with deep features extracted from state-of-the-art networks like MobileNetV2 and VGG-16. By integrating texture-based features like LBP and HoG, which are particularly adept at capturing age-related characteristics, we have devised a robust feature extraction process. These features are intelligently fused, and subsequent layers fine-tune them, resulting in a highly effective and adaptable verification system. The experimental validation of our methodology, conducted using the Cross-Age Celebrity Dataset, yielded exceptionally promising results. With an outstanding accuracy rate of 98.32%, our approach showcases its efficacy in overcoming age-related challenges, making a significant stride toward the development of reliable and precise face verification systems.

References

1. Newlin Shebiah, R., Arivazhagan, S., Amrith, S.G., Adarsh, S., 2023. Person Re-Identification by Siamese Network. Intel. Artif. 26, 25–33. https://doi.org/10.4114/intartif.vol26iss71pp25-33
2. De Oliveira Lima, J.P., Figueiredo, C.M.S., 2021. A temporal fusion approach for video classification with convolutional and LSTM neural networks applied to violence detection. Intel. Artif. 24, 40–50. https://doi.org/10.4114/intartif.vol24iss67pp40-50
3. Shi, Y., Zhu, Y., 2020. Recognition of motion-blurred CCTs based on deep and transfer learning. Intel. Artif. 23, 1–8. https://doi.org/10.4114/intartif.vol23iss66pp1-8
4. Daniel López-Cabrera, J., Alberto López Rodriguez, L., Pérez-Díaz, M., 2020. Classification of breast cancer from digital mammography using deep learning. Intel. Artif. 23, 56–66. https://doi.org/10.4114/intartif.vol23iss65pp56-66

5. Ghosh, S., Singh, A., Kumar, S., 2023. PB3C-CNN: An integrated Parallel Big Bang-Big Crunch and CNN based approach for plant leaf classification. Intel. Artif. 26, 15–29. https://doi.org/10.4114/intartif.vol26iss72pp15-29
6. Gokulnath, B. V., Usha Devi, G., 2020. A survey on plant disease prediction using machine learning and deep learning techniques. Intel. Artif. 23, 136–154. https://doi.org/10.4114/intartif.vol23iss65pp136-154
7. Solís, M., Calvo-Valverde, L.A., 2022. Performance of Deep learning models with transfer learning for multiple-step-ahead forecasts in monthly time series. Intel. Artif. 25, 110–125. https://doi.org/10.4114/intartif.vol25iss70pp110-125
8. Zhang, Z., Song, Y., Qi, H., 2017. Age progression/regression by conditional adversarial autoencoder, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, pp. 5810–5818.
9. Yang, H., Huang, D., Wang, Y., Jain, A.K., 2018. Learning face age progression: A pyramid architecture of GANs, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA, pp. 31–39.
10. Huang, Z., Zhang, J., Shan, H., 2021. When age-invariant face recognition meets face age synthesis: A multi-task learning framework, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Virtual, pp. 7282–7291.
11. Chen, B.-C.; Chen, C.-S.; Hsu, W.H. Cross-age reference coding for age-invariant face recognition and retrieval. In European Conference on Computer Vision; Springer: Berlin/Heidelberg, Germany, 2014; pp. 768–783.
12. Du, L.; Hu, H.; Wu, Y. Cycle age-adversarial model based on identity preserving network and transfer learning for cross-age face recognition. IEEE Trans. Inf. Forensics Secur. 2019, 15, 2241–2252.
13. Huang, Y.; Chen, W.; Hu, H. Age-puzzle facenet for cross-age face recognition. In Asian Conference on Computer Vision; Springer: Cham, Switzerland, 2018; pp. 603–619.
14. Huang, Y.; Hu, H. A parallel architecture of age adversarial convolutional neural network for cross-age face recognition. IEEE Trans. Circuits Syst. Video Technol. 2020, 31, 148–159.
15. Y. Wen, Z. Li, Y. Qiao, Latent Factor Guided Convolutional Neural Networks for Age-Invariant Face Recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4893–4901.
16. S. Bianco, Large Age-gap Face verification by Feature Injection in Deep Networks, Pattern Recogn. Lett. 90 (2017) 36–42.
17. X. Wang, Y. Zhou, D. Kong, J. Currey, D. Li, J. Zhou, Unleash the Black Magic in Age: a Multi-task Deep Neural Network Approach for Cross-age Face Verification, IEEE Int. Conf. Face. Gest. Recognition. (2017).
18. H. Li, H. Hu, C. Yip, Age-Related Factor Guided Joint Task Modelling Convolutional Neural Network for Cross-Age Face Recognition, IEEE Trans. Inf. Forensics Security 13 (9) (2018) 2383–3239.
19. Y. Wang, D. Gong, Z. Zhou, X. Ji, H. Wang, Z. Li, W. Liu, T. Zhang, Orthogonal Deep Features Decomposition for Age-invariant Face Recognition. European conference on computer vision (ECCV), 2018.
20. J. Zhao, S. Yan, J. Feng, Towards Age-invariant Face Recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2020).
21. Y. Li, G. Wang, L. Nie, Q. Wang, W. Tan, Distance metric optimization driven convolutional neural network for age-invariant face recognition, Pattern Recogn. 75 (2018) 51–62.
22. S. Zhao, J. Li, J. Wang, Disentangled representation learning and residual GAN for age-invariant face verification, Pattern Recogn. 100 (2020), 107097.

23. L. Du, H. Hu, Y. Wu, Age factor removal network based on transfer learning and adversarial learning for cross-age face recognition, IEEE Trans. Circuits. Sys. Video. Technology, (DOI: 10.1109/TCSVT.2019.2923262), 2019.
24. H. Wang, D. Gong, Z. Li, W. Liu, Decorrelated adversarial learning for age-invariant face recognition. IEEE Conference on Computer Vision and Pattern Recognition, 2019.
25. Y. Huang, H. Hu, A Parallel Architecture of Age Adversarial Convolutional Neural Network for Cross-Age Face Recognition, IEEE Trans. Circuits. Sys. Video. Technology 31 (1) (2021) 148–159.
26. Li S, Lee HJ. Effective Attention-Based Feature Decomposition for Cross-Age Face Recognition. *Applied Sciences*. 2022; 12(10):4816. https://doi.org/10.3390/app12104816
27. Wang, H., Sanchez, V., & Li, C. (2024). Cross-Age Contrastive Learning for Age-Invariant Face Recognition. arXiv preprint arXiv:2312.11195.
28. Hast, A. (2023). Age-Invariant Face Recognition Using Face Feature Vectors and Embedded Prototype Subspace Classifiers. In *Advanced Concepts for Intelligent Vision Systems: 21st International Conference, ACIVS 2023 Proceedings* (pp. 88–99). Kumamoto, Japan. doi:10.1007/978-3-031-45382-3_8
29. D.Chen,X.Cao,F.Wen,andJ.Sun,“Blessingofdimensionality:High- dimensional feature and its efficient compression for face verification,” in IEEE Conf. Computer Vision and Pattern Recognition, 2013, pp. 3025– 3032.
30. D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang, “Hidden factor analysis for age invariant face recognition,” in IEEE Int. Conf. Computer Vision, 2013.
31. Chen, Bor-Chun, Chu-Song Chen, and Winston H. Hsu. "Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset." IEEE Transactions on Multimedia 17.6 (2015): 804-815.
32. Wen, Yandong, Zhifeng Li, and Yu Qiao. "Latent factor guided convolutional neural networks for age-invariant face recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.