SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V10I7P110

Volume 10 Issue 7, 97-107, July 2023

Original Article

Dynamic Weight Allocation for Improved Age-Invariant Face Recognition and Age Estimation System: DMT-MFFCNN Approach

M. Rajababu[1], K. Srinivas[2], H. Ravisankar[3]

[1]*Department of CSE, JNTUH, Hyderabad, Telangana, India.*
[2]*Department of CSE, V.R. Siddhartha Engineering College, Andhra Pradesh, India.*
[3]*CTRI, Andhra Pradesh, India.*

[1]*Corresponding Author : raajababu.makineedi@gmail.com*

Received: 05 May 2023 Revised: 30 June 2023 Accepted: 19 July 2023 Published: 31 July 2023

Abstract - *The investigation of Age Invariant Face Recognition (AIFR) and Age Estimation has received more attention in real-time applications. However, it is an exciting and challenging task due to the rapidly growing rate of image uploads on the internet in today's data-driven world. Multitask Learning (MTL) has shown great potential across various applications, outperforming single-task learning, mainly when applied in deep multitask networks. One of the most critical issues in MTL is determining the weight of each task. We propose a Dynamic Multitask and Multi-Feature Fusion CNN (DMT-MFFCNN) framework that can jointly perform face recognition and age estimation tasks by adopting the weights of the tasks dynamically based on the difficulty in training them. This approach avoids the exhausting and time-consuming process of manually tuning the weights. In particular, the proposed method can successfully enhance the multitask learning model's performance and can be quickly applied without hyperparameter tuning. We evaluated our model on the FGNET, CACD, and UTK Face datasets and a Live dataset AS-23 containing face images of different ages. In the proposed system, the dataset's images are first converted to grey level and Local Binary Pattern (LBP)images. Next, a Merged dataset is created by combining the features by fusion of original, grey level and LBP datasets. The experimentation is carried out with VGG-16, DenseNet-201 and ResNet-50 individual CNN models and the VDeRe-23 feature fusion CNN model. The final step is cross-validating the Merged dataset against a new unseen dataset. Among all the comparisons on Merged Dataset, the VGG-16 model achieved the best face recognition accuracy of 94.06%, and for age estimation with the same model, obtained an MAE value of 1.7 years. Additionally, with the feature fusion model (VDeRe-23), face recognition and age estimation tasks achieved an accuracy of 99.47% and an MAE of 1.67 years, respectively. The results of the extensive experiments have demonstrated that the proposed DMT-MFFCNN yields superior performance than state-of-the-art methods for both face recognition and age estimation tasks using single-task learning.*

Keywords - *Age estimation, Age Invariant Face Recognition, Convolutional Neural Networks, Cross-dataset training, Merged datasets, Multitask Learning.*

1. Introduction

Age classification and facial recognition have become more crucial in various industries, including security, entertainment, and healthcare[21, 22]. However, precisely identifying faces and determining age can be difficult due to variables in lighting, pose, and age-related changes in facial features. Researchers have created several techniques to increase the accuracy of these tasks. Popular deep learning approaches have recently shown promising AIFR and age estimation accuracy. However, most suggested solutions use single-task learning, which handles each job independently. We believe face recognition is not a standalone problem and can be tangled with other related tasks. This motivates us to investigate MTL for AIFR and age estimation. Recently, MTL has become a popular method for enabling a model to learn numerous tasks simultaneously by sharing the features and parameters common to all tasks.

From speech recognition and natural language processing [1, 2] to computer vision, numerous machine learning applications [15, 34] have successfully used MTL. It combines learning by several tasks to represent features challenging to learn by one task but simple to learn by another. An innovative CNN model created by Ranjan et al. [14] is a thoroughly supervised multitask CNN-based method that completes various face analysis tasks at once.

In Multitask Learning (MTL), determining the appropriate weights to assign to different tasks is crucial. The weights dictate the relative importance of each task during the learning process, and they can significantly impact the overall performance of the MTL model. Most strategies for learning multiple tasks use experimental techniques to find the best task weights.

The Hyperface [14] manually adjusts the weights of tasks by their significance in the overall loss, and [20] finds the ideal weight for pedestrian detection assignment using greedy search. In contrast to static weights methods, the approaches [3, 24] change the task weights during the network training. In [24], the authors have set a weight value of 1 for the primary task and dynamic weight values for the auxiliary tasks.[17] modifies the SoftMax loss function for each job by including an uncertainty coefficient. A robust and adaptable method for dynamic weight assignment in a multitask CNN uses a SoftMax layer, particularly when the tasks are associated with classification issues. Multitask learning is complex with these approaches because they need to add new hyperparameters to update the task weights. This necessitates careful calibration of hyperparameters such as learning rate, regularization strength and task weighting.

In order to solve these problems, we propose DMT-MFFCNN, which uses two task-specific subnetworks and a shared feature extraction backbone to accomplish the face recognition and age estimate tasks simultaneously. To address the challenge of balancing the two tasks' trade-offs, we introduce a novel joint cross-entropy loss function that dynamically assigns weights to each task during training based on their relative performance. This enables the model to learn to prioritize the more complicated task and adapt to the changing task difficulty and correlation.

This technique uses a combination of VGG16, DenseNet-201, and ResNet-50 Convolutional Neural Networks (CNNs), simultaneously performing multiple tasks, such as recognizing faces and predicting age. The CNNs are designed to dynamically adjust their architecture based on the input image, allowing them to extract the most relevant features for each task.

For dynamic weight assignment to the tasks in MTL, we proposed to add a SoftMax layer in the network at the end of the shared hidden layers. The units of the SoftMax layer correspond to task weights, and no additional hyperparameters are used to update the task weights.

A more enhanced loss function for dynamic updation of weights, instead of updating [23] and the networks' filtering weights, allows the networks to focus on training the difficult task by automatically allocating a greater weight. A brief description of this paper's significant contributions is provided as follows.

- Building a Lightweight model for AIFR and Age Estimation using VGG16, Densenet-201 and Resenet-50 Fusion Model.
- Formulating AIFR and Age Estimation as an MTL problem enables the CNNs to use a joint cross-entropy loss function for a dynamic weighting scheme per task significance.
- Conducting a comprehensive evaluation of the proposed approach on the merged dataset to demonstrate its effectiveness in AIFR and Age Estimation.
- Demonstrating significant improvements in performance over existing single-task learning methods for AIFR and Age Estimation.

The remainder of this paper is organized as follows. A summary of the relevant literature is presented in Section II, the method of MTL with dynamic weights is presented in Section III, and the analysis of the experimental findings is presented in Section IV. In Section V, we conclude and outline our plans for the future.

2. Related work

2.1. Age Invariant Face Recognition

The surge in popularity of deep learning techniques, particularly CNNs[31], has revolutionized computer vision applications, especially in solving classification problems. Addressing the challenge of AIFR, Wen et al. [7] introduced the innovative LF-CNN framework as a highly reliable and effective method for AIFR tasks, achieving remarkable accuracy on three AIFR datasets: 97.51% on MORPH Album2, 98.5% on CACD-VS, and an impressive 99.5% on LFW dataset.

Xu et al. [12] applied Coupled Auto-Encoder Network (CAN) to solve AIFR and retrieval problems. It comprises two shallow neural networks that connect encoders. Wang et al. [13] created the innovative network known as Orthogonal Embedding CNN (OE-CNN) to increase the effectiveness of AIFR. A novel end-to-end CNN-based model called AA-CNN was designed by Huang and Hu [16] using training datasets labelled by identity or age.

Yan, Chenggang, et al. [18] introduced an AIFR approach, MFD, that learns more robust and discriminative features while reducing intra-class variations. An AIFR system with a new Artificial approach and a parallel deep CNN was designed and evaluated by Dharavath et al. [19]. A two-part joint multitask CNN (JMCNN) framework was suggested by Yu and Jing [27].

2.2. Age Estimation

Liu et al. [6] introduced the Ordinal Deep Feature Learning (ODFL) method to learn feature descriptors for face representation from raw pixels directly. However, they encountered a sub-optimal issue as ODFL independently learned feature extraction and age estimation. To address this limitation, the researchers proposed an ordinal deep learning (ODL) framework that leverages complementary information from both procedures to strengthen their model.

In a separate study, Fang et al. [28] developed a technique for age and gender prediction by marking the people region and subsequently estimating age and gender. Notably, their approach achieved an impressive Mean Absolute Error (MAE) of 1.84 years by focusing solely on face extraction through a saliency-detecting network.

Nam et al. [29] devised a deep CNN-based model to address age prediction in low-resolution facial images. This model is capable of reconstructing low-resolution faces into high-resolution ones[9, 11].

Zhang et al. [10] introduced the Cross-Dataset Training CNN (CDCNN) as a solution for age estimation using a conventional CNN architecture and cross-dataset validation. Their approach involved utilizing a pre-trained VGG-16 model from ImageNet and treating age estimation as a classification problem. They achieved a promising Mean Absolute Error (MAE) of 3.11 years when evaluated on the AFAD dataset using cross-dataset training.

Even if the performance of all the DNN models mentioned previously was excellent. They are single-task learning models with highly complicated architectures and many hidden layers. In addition to facilitating the learning of several tasks inside a single network, MTL also enhances the primary task through auxiliary tasks [15]. MTL techniques can be categorized as static and dynamic depending on how the task weights are updated.

The static methods [14, 5] manually set the task weights before the networks are trained, and they remain fixed throughout the entire network training process, in contrast to the dynamic methods, which update the task weights as training is processed [3, 25]. Weights used in the static methods can be set in two different ways. The first method assigns each task an identical weight, and the second assigns various weights based on relative relevance. The authors of Fast R-CNN [5] use a loss in MTL to jointly train the classification algorithm.

An MTL system for face analysis, identification and gender recognition was proposed by Hyperface [14] and used deep CNNs. The classification and ranking tasks were combined in a multitask network for the person reidentification challenge [3]. The two tasks were jointly optimized while giving equal weight to each job. According to Tian et al. [20], the main task weight is fixed at 1, and a greedy search is used to determine the weights of the side tasks between 0 and 1. In [4], a new loss function is used, and a hyperparameter is added to balance the training of various tasks. An uncertainty coefficient is added by [17] to unite the various loss functions. A multitask network is suggested by Zhang et al. [25] to identify facial characteristics and detect face landmarks.

The face analysis tasks are set as dynamic secondary tasks, whereas the tasks for detecting face landmarks are the primary tasks with weight 1. An advanced face pose-invariant identification model with automatic weight learning was created by Yin et al. [23].

The secondary tasks share the dynamic tasks produced by the SoftMax layer, and the primary task is set to 1. In order to accurately estimate the gender and age group of facial images, Abdolrashidi et al. [24] suggested a deep learning architecture based on an ensemble of a ResNet-based and attention-based model.

A Lightweight Multitask CNN (LMTCNN) was suggested by Lee et al. [26] for the categorization of age and gender at the same time. They used a complex parameter-sharing model to learn the single CNN for both tasks. They divide the general CNN into depth-wise and point-wise convolution networks to reduce the computing cost. For age and gender recognition tasks, they tested with the Audience dataset and got an accuracy of 70.78% and 85.16%, respectively.

3. Proposed Method

Although the CNN models mentioned above have shown good performance, they are limited to single-task applications and have complex architectures with numerous hidden layers. To overcome this limitation and meet the requirement for two models with the same architecture but different weights for age estimation and AIFR, we propose the DMT-MFFCNN model.

This model is a fusion of VGG-16, DenseNet-201, and ResNet-50 CNN architectures, denoted as VDeRe-23. The architecture of our proposed model is illustrated in Figure 1. To address the issue of simultaneous age estimation and AIFR, we apply a regularization term to two different types of features in the model [43]. This regularization promotes mutual enhancement between the two tasks, allowing them to benefit from each other's representations and improving overall performance.

We present a novel joint cross-entropy loss function that can dynamically assign weights to each task during training based on each task's relative performance to address the issue of balancing the two tasks' trade-offs. This section briefly overviews our suggested approach and explains the underlying algorithms. The following stages are necessary for the proposed system to function:

1) Prepare the image datasets for the experiment by preprocessing them.
2) Convert the image datasets into the grey level and LBP individual datasets.
3) Combine the features from the original, grey-level, and Local Binary Pattern (LBP) datasets to create a Merged dataset. Then we apply VGG-16, DenseNet-201 and ResNet-50 models on this Merged dataset.
4) Next, evaluate the fusion model on the unseen AFRD dataset using cross-dataset validation, where the merged dataset serves as the training set.
5) Evaluate and compare the performance of all the CNN above models to identify the model that demonstrates the best results.

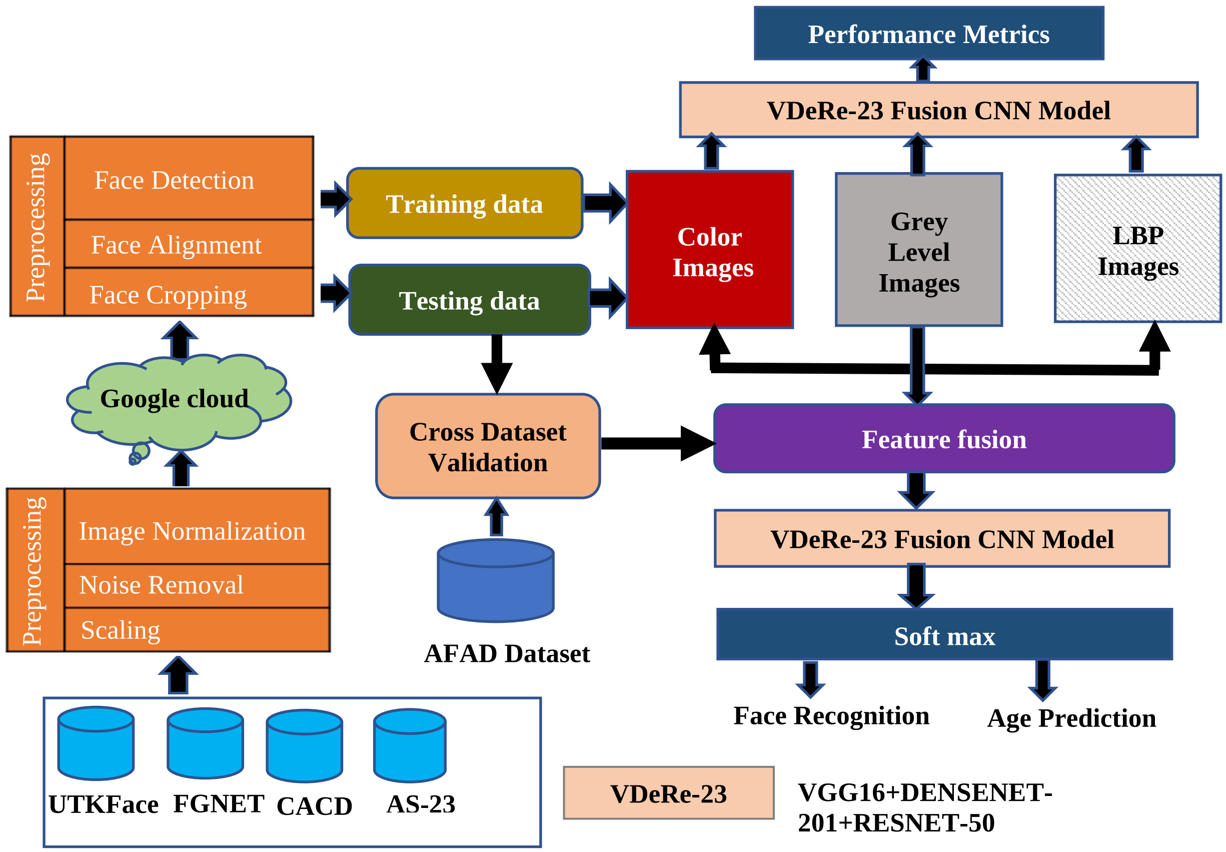

Fig. 1 Block diagram of proposed DMT-MFFCNN model

3.1. Datasets

The proposed method utilizes three publicly available datasets: CACD, UTKFace, FGNET and a live dataset called AS-23. UTKFace dataset comprises an extensive collection of facial images representing diverse age ranges. Each image is annotated with information about age, gender, and ethnicity.

FGNET[32] includes 1,002 images from 82 individuals with varying image counts per person and a wide age range from 0 to 69. CACD[33] comprises 2,000 distinct subjects and contains a vast collection of 163,446 celebrity images. The images cover a wide age range, from 12 years old to 62 years old. AS-23 is a live dataset with images of students aged 2 to 21, totalling 100 subjects, each with 10,000 images. Figure 2 displays sample images from these datasets. A single Dataset (1,98,153 images) was created by combining the images from the multiple datasets. Cross-dataset evaluation is carried out on the combined dataset with the Asian Face Age Dataset (AFAD)[30, 35], and the results are then compared. There are 1,64,432 labelled images in this dataset; however, only 10,000 were used for testing.

3.2. Preprocessing

Preprocessing is crucial when using Convolutional Neural Networks (CNNs) for any task. It involves employing various techniques to enhance image quality and emphasize the features essential for extraction.

3.2.1. Normalization

The normalization process helps improve an image's visual consistency by adjusting its contrast and brightness. Image processing and machine learning can be normalised through two primary methods. The first approach involves scaling the pixel values to a specific range. The second method involves adjusting the mean and variance of the pixel values in an image. The main aim of image normalization is to standardize the image's appearance, making it easier for algorithms and models to process and analyze the images more effectively[36, 37].

3.2.2. Noise Removal

The noise sources in images vary significantly from environmental influences to sensor noise[41], compression artefacts and more. The techniques for removing noise from images try to minimize or eliminate it. Gaussian and median filtering techniques are applied in our methodology. For a given fixed window size, the Median filtering keeps the edges while cutting the noise, and the Gaussian filtering blurs the intended area while reducing the noise with higher frequencies.

3.2.3. Scaling

Scaling is an image processing operation used to change the dimensions of an image while preserving its aspect ratio. The training dataset contained images of varying sizes; thus, before being utilized as input for the model, the images had to be reduced to 224 X 224.

Fig. 2 Sample images from (a) UTKFace (b) FG-NET (c) CACD (d) AS-23

3.3. Google Cloud

Google Cloud Storage is a Google Cloud Platform (GCP) feature. It allows users to store various types of data, including face images, in a scalable and secure manner on Google's infrastructure. We used Google Co-Lab, a cloud-based platform for running Python code and data analysis, to conduct thorough research and analyse the results. Google Co-Lab provided access to powerful computing resources available through Google Cloud, enabling efficient computation and analysis of large datasets.

3.4. Face Analysis

After the datasets are prepared, the next step in the face analysis process involves the following techniques:

3.4.1. Face Detection

Facial detection is a computer technology utilized to Face detection is a computer technology used to identify the presence of human faces within images or videos[42]. The Haar cascade face detection algorithm is a widely adopted method for face detection. It can identify critical facial features like eyes, nose, and mouth, accurately detecting and localising faces within images or videos.

3.4.2. Face Alignment

Face alignment is a method used for repositioning and orienting a face within an image to align it with a predefined reference frame or canonical representation. The primary goal of this process is to facilitate the extraction of facial features for further analysis in computer vision tasks. By locating these landmarks, the position and orientation of the face can be accurately determined, allowing for precise alignment and enabling subsequent facial feature extraction and analysis tasks.

3.4.3. Face Cropping

Following the successful detection and alignment of the face, the subsequent step involves creating a new image that isolates the face by cropping it from the original image. This is achieved using a straightforward Geometric Transformation method called "Bounding Box Cropping".

The bounding box encapsulates the facial region, and then the image is cropped to preserve only the pixels within that box. As a result, the cropped image solely contains the focused facial region, which is ideal for further analysis or processing tasks related to facial features and recognition.

3.5. Training and Test Split

In this experiment, we evaluated three CNN models and a fusion model separately, starting with the Merged Dataset, which was divided into a 70:30 split ratio for training and testing. Subsequently, the Merged Dataset was cross-validated, with the Merged Dataset serving as the training set, and an unseen dataset called AFAD was used as the testing set, containing 10,000 images. Below is Table 1, providing the size of each Dataset and the distribution of the train-test split:

Table 1. Dataset size and train-test split distribution test sets

S.No.	Dataset	#Training	#Testing	#Total
1	FGNET	701	301	1002
2	CACD	1,14,412	49034	1,63,446
3	UTKFace	16594	7111	23705
4	AS-23	3000	7000	10000
5	Merged Dataset	134707	63446	198153

3.6. Model Architecture

According to Figure 3, the deep CNNs form the foundation of the suggested MTL with dynamic weights. AIFR and age estimation are the two branches of the task-specific layers, respectively.

The two branches' very identical structural characteristics make it easier to acquire facial recognition from the previously performed face age estimation task. In particular, branch-1 can retrieve the embedded features for face recognition, while branch-2 determines the probability of age prediction with the help of the fully connected SoftMax layer. Key elements of the DMFT-CNN system are CNNs.

The individual CNNs, in this case, are ResNet-50[39], DenseNet-201[40], and VGG-16[38]. These individual CNNs and the Feature Fusion Model are used in this experimentation to recognise faces from face datasets. During training and fine-tuning, the tasks' weights are dynamically assigned. This is achieved through the use of shared layers and task-specific output layers.

The SoftMax layer connected at the end of the hidden shared layers assigns the dynamic weights for the tasks. The loss function is typically designed to optimize the performance of both tasks simultaneously. This is achieved through a joint loss function, which combines the output of the shared layers with the task-specific output layers for face recognition and age classification. The joint output vector for the neural network can still be constructed as [yf, ya], where yf represents the predicted probabilities for face recognition, and ya represents the predicted age values. However, the loss function will combine the cross-entropy loss for face recognition and the MAE loss for age estimation.

The cross-entropy loss function for face recognition can be defined as:

$$Lf = \sum_{i=1}^{n} ti * \log{(yfi)} \tag{1}$$

Where ti is the one-hot encoded true label for the i-th image, n is the number of classes in the face recognition task, and yfi is the predicted probability of the i-th image belonging to the fth class.

The MAE loss function for age estimation can be defined as:

$$La = \sum_{i-1}^{n} |ya - ti| \tag{2}$$

Where ya is the predicted age value, ti is the actual age value, and n is the total number of images in the dataset.

The joint loss(L) is computed using Lf and La as follows:

$$L = Lf + La \tag{3}$$

By incorporating both cross-entropy loss for face recognition and MAE loss for age estimation in the joint loss function, the network can be trained to balance both tasks effectively and improve the overall performance on both tasks simultaneously.

3.7. Evaluation Metrics

Face recognition accuracy is determined using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) metrics as follows:

Accuracy = (TP+TN)/(TP+FP+FN+TN) (4)

The Mean Absolute Error (MAE) is a popular choice for assessing the accuracy of age estimation models because it provides a straightforward and interpretable measure of how well the model predicts ages. The MAE measures the absolute difference between the predicted age and the actual age of individuals in the dataset. It is calculated using equation (5),

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}i - yi| \quad (5)$$

Where n denotes the total number of data samples, yi denotes the actual age, and ŷi denotes the predicted age of the i-th sample. A lower MAE value indicates better age prediction success by the model.

4. Results

The experimentation for age estimation was conducted in two stages using a merged dataset. In the first stage of the experimental investigation, three individual Convolutional Neural Network (CNN) models and feature fusion models were evaluated on the merged dataset. In the second stage of the experimental investigation, the fusion CNN model was evaluated using a cross-dataset validation approach. The merged dataset was used as the training set, while an unseen AFRD (Age Face in the Wild) dataset was used as the testing set. The performance comparison results of individual CNN models on the merged dataset are presented in Table 2.

Additionally, Table 3 displays the performance results for the merged dataset and cross-dataset training using the VDeRe-23 feature fusion model. Figures 4 and 5 show the comparative metrics analysis of AIFR and age estimation for individual CNN and feature fusion models, respectively.

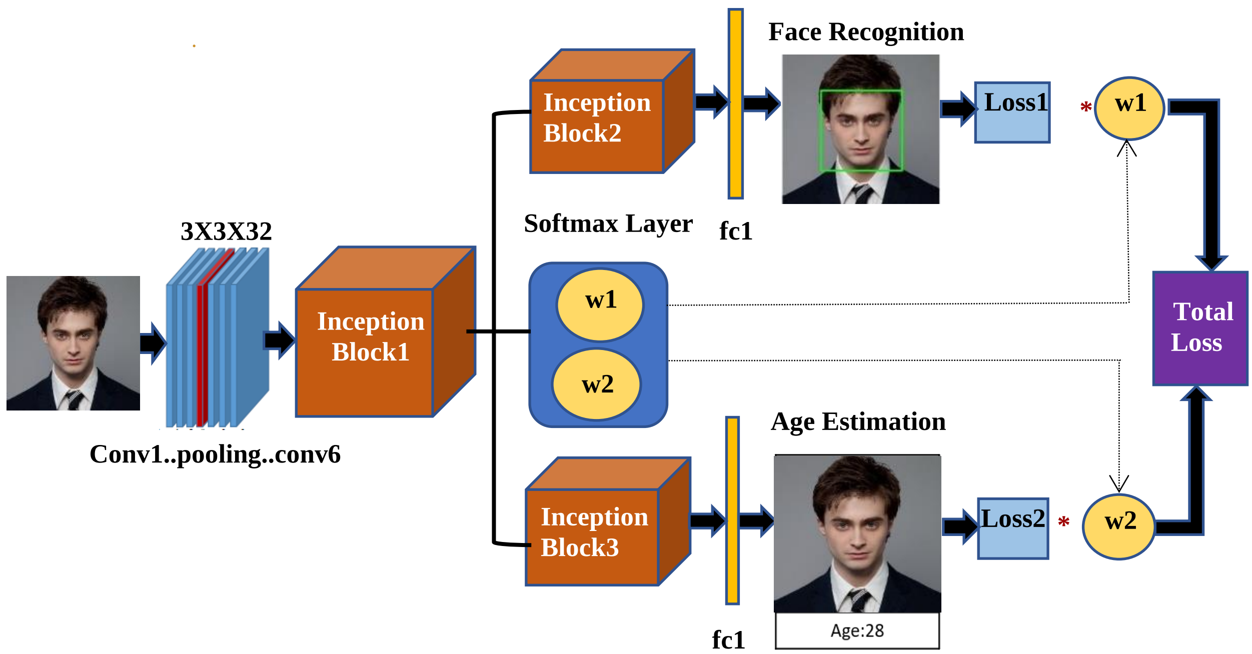

Fig. 3 The proposed MTL framework to perform AIFR with age estimation

Table 2. Performance comparison of AIFR and MAE on merged dataset

Model	Face recognition (Accuracy %)	Age (MAE)
VGG-16	94.06	1.71
Densenet-201	92.4	1.85
ResNet 50	91.43	2.11

Table 3. Performance comparison with cross dataset validation

Model	Face recognition		Age (MAE)	
	Merged Dataset	Cross Dataset Validation	Merged Dataset	Cross Dataset Validation
VDeRe-23	99.47	90.54	1.67	2.45

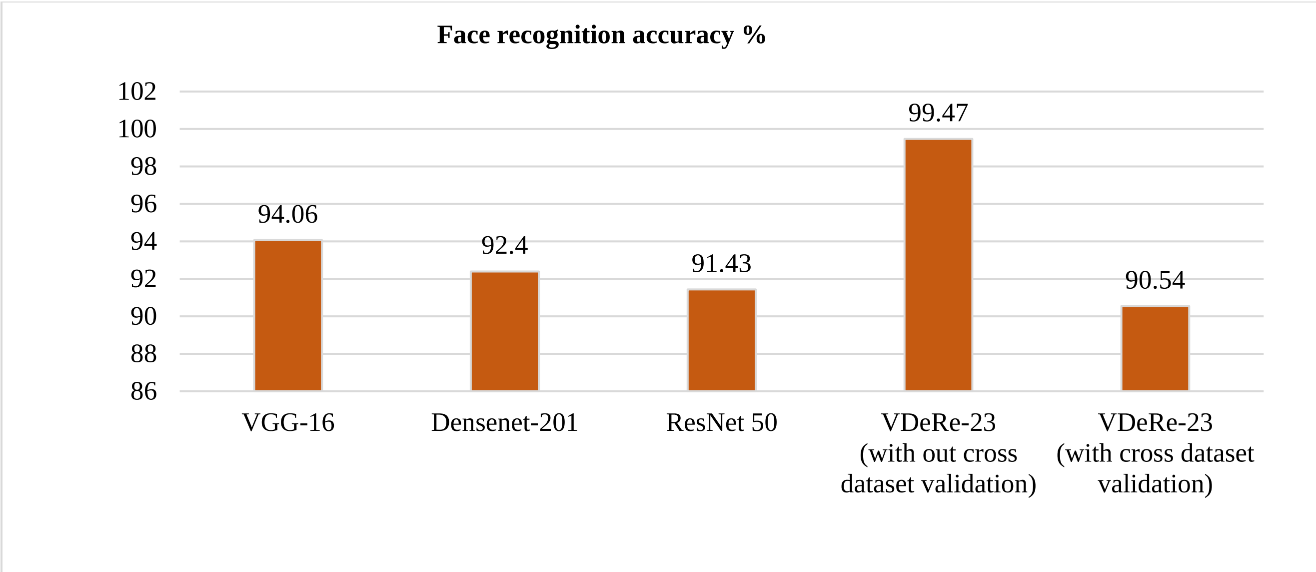

Fig. 4 Performance comparision of AIFR with individual and feature fusion CNN models

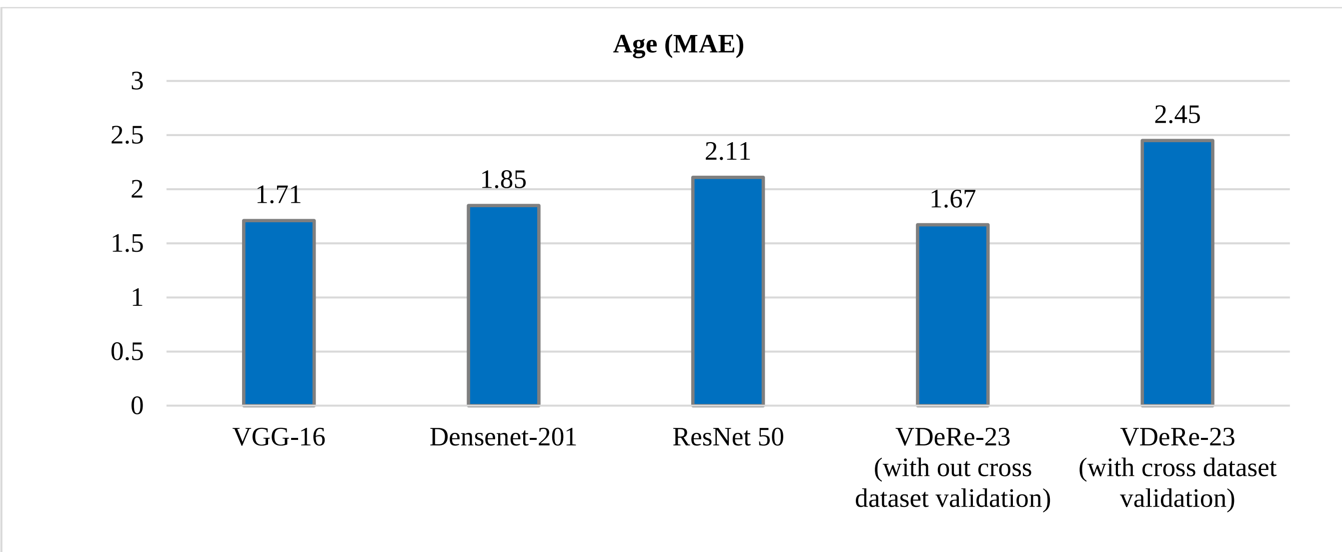

Fig. 5 Performance comparison of MAE with individual and feature fusion CNN models

4.1. Result Analysis

The performance comparison has been conducted between the proposed and state-of-the-art methods on AIFR and age estimation using single-task learning. The results of this comparison are presented in Table 4 for the AIFR task and Table 5 for age estimation using single-task learning. Results showed that the VDeRe-23 model achieved superior accuracy (99.47%) and a low age estimation MAE (1.67) on the merged dataset, outperforming other methods. However, during cross-dataset validation with training on the merged dataset and testing on unseen AFAD data, the model's accuracy dropped to 90.54%, and the age estimation MAE increased to 2.45 years. This highlights the model's challenges in generalizing to new datasets, though it remains promising for age estimation tasks on the merged dataset.

4.2. Discussions

We present a detailed performance comparison of our proposed method against state-of-the-art techniques for Age Invariant Face Recognition (AIFR) and age estimation using single-task learning In Table 4 and Table 5. Our method, trained on the Merged dataset, outperforms all other approaches and achieves the highest performance for both tasks. Notably, applying Multitask Learning (MTL) techniques shows superior performance compared to single-task learning methods, underscoring the benefits of simultaneously leveraging multiple tasks to improve AIFR and age estimation outcomes. The success of our proposed method on the Merged dataset validates its effectiveness and potential in pushing the boundaries of AIFR and age estimation research.

Table 4. Comparison of results on AIFR

Dataset	Author	Method	Accuracy %
CACD	Xu et al. [12]	CAN (2017)	77
FGNET			86.5
CACD	Wen et al. [7]	LF-CNN (2016)	93.8
FGNET			88.10
CACD	Huang and Hu [16]	AA-CNN (2020)	95.1
FGNET			89.34
CACD	Yan, Chenggang, et al. [18]	MFD (2022)	97.51
CACD	Dharavath et al.[19]	Parallel Deep CNN (2022)	81
FGNET			81.25
Merged Dataset	Ours	Proposed Method	**99.47**
Merged Dataset	Ours	Proposed Method (with Cross Dataset Validation)	**90.54**

Table 5. Comparison of results on age estimation

Author	Method	Dataset	MAE
Garain et al.[8]	GRA_Net	FG-NET	3.23
Zhang et al.[10]	CDCNN	CACD	3.96
Zhang et al.[25]	AL-RoR-34	FG-NET	2.39
Nam et al.[29]	CNN with GAN	FG-NET	8.3
Liu et al.[6]	ODFAL	FG-NET	3.89
Fang et al.[28]	Multi-Stage Learning	FGNET, CACD	1.81
Ours	Proposed Method	Merged Dataset	**1.67**
Ours	Proposed Method (Cross Dataset Validation)	Merged Dataset	**2.45**

5. Conclusion

We propose a novel method for joint Age Invariant Face recognition and age estimation using Dynamic Multitasking and Multi-Feature Fusion CNN. The proposed method is implemented using the VDeRe-23 fusion model and does not require any hyperparameters, which simplifies the training process and reduces the need for manual tuning. Results from experiments on the merged dataset showed that our model outperformed baseline models with robust accuracy. The joint cross-entropy loss function enables the model to adapt to changing task difficulty and correlation, balancing the two tasks effectively. The proposed method provides a promising approach for joint face recognition and age estimation, with potential security, surveillance, and healthcare applications. Future work can investigate the extension of this method to other related tasks, such as gender classification or facial expression recognition, and explore the possibility of using other CNN architectures for improved performance.

References

[1] Ronan Collobert, and Jason Weston, “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning,” *Proceedings of the 25th International Conference on Machine Learning*, pp. 160-167, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[2] Li Deng, Geoffrey Hinton, and Brian Kingsbury, “New Types of Deep Neural Network Learning for Speech Recognition and Related Applications: An Overview,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8599-8603, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[3] Weihua Chen et al., "A Multi-Task Deep Network for Person Re-Identification," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[4] S. Sandhya Kumari, and K. Sandhya Rani, "Selection of MSER Region based Ultrasound Doppler Scan Image Big Data Classification using a Faster RCNN Network," *International Journal of Computer Engineering in Research Trends*, vol. 9, no. 10, pp. 184-192, 2022. [CrossRef] [Publisher Link]

[5] Ross Girshick, "Fast R-CNN," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448, 2015. [Publisher Link]

[6] Hao Liu et al., "Ordinal Deep Learning for Facial Age Estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 2, pp. 486-501, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[7] Yandong Wen, 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*, pp. 4893–4901, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Avishek Garain et al., "GRA_Net: A Deep Learning Model for Classification of Age and Gender from Facial Images," *IEEE Access*, vol. 9, pp. 85672-85689, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Muhammad Firdaus Mustapha et al., "Age Group Classification using Convolutional Neural Network (CNN)," *Journal of Physics: Conference Series*, vol. 2084, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Beichen Zhang, and Yue Bao, "Cross-Dataset Learning for Age Estimation," *IEEE Access*, vol. 10, pp. 24048-24055, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Duc-Quang Vu et al., "Age and Gender Recognition using Multi-Task CNN," *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 1937-1941, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Chenfei Xu, Qihe Liu, and Mao Ye, "Age Invariant Face Recognition and Retrieval by Coupled Auto-Encoder Networks," *Neurocomputing*, vol. 222, pp. 62–71, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Yitong Wang et al., "Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition," *Proceedings of the European Conference on Computer Vision*, vol. 11219, pp. 764–779, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Rajeev Ranjan, Vishal M Patel, and Rama Chellappa, "Hyperface: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 121-135, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Sebastian Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," *Arxiv Preprint Arxiv:1706.05098*, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[16] Yangjian Huang, and Haifeng Hu, "A Parallel Architecture of Age Adversarial Convolutional Neural Network for Cross-Age Face Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 1, pp. 148-159, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Alex Kendall, Yarin Gal, and Roberto Cipolla, "Multi-Task Learning using Uncertainty to Weigh Losses for Scene Geometry and Semantics," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7482–7491, 2018. [Google Scholar] [Publisher Link]

[18] Chenggang Yan et al., "Age-Invariant Face Recognition by Multi-Feature Fusion and Decomposition with Self-Attention," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 18, no. 15, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Krishna Dharavath, Amarnath Gaini, and Viod Adla, "A Parallel Deep Learning Approach for Age Invariant Face Recognition System," *International Conference on Computer Communication and Informatics*, pp. 1-6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Yonglong Tian et al., "Pedestrian Detection Aided by Deep Learning Semantic Tasks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5079-5087, 2015. [Google Scholar] [Publisher Link]

[21] D. Betteena Sheryl Fernando et al., "Face Recognition for Home Security," *SSRG International Journal of Computer Science and Engineering*, vol. 6, no. 10, pp. 7-12, 2019. [CrossRef] [Publisher Link]

[22] Vinita Bhandiwad, "A Review Paper on 2D to 3D Face Recognition Technology," *SSRG International Journal of VLSI & Signal Processing*, vol. 2, no. 2, pp. 10-13, 2015. [CrossRef] [Publisher Link]

[23] Xi Yin, and Xiaoming Liu, "Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition," *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 964-975, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[24] Amirali Abdolrashidi et al., "Age and Gender Prediction from Face Images using Attentional Convolutional Network," *Arxiv preprint arXiv:2010.03791*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] Zhanpeng Zhang et al., "Learning Deep Representation for Face Alignment with Auxiliary Attributes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 918-930, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[26] Jia-Hong Lee et al., "Joint Estimation of Age and Gender from Unconstrained Face Images using Lightweight Multi-Task CNN for Mobile Applications," *IEEE Conference on Multimedia Information Processing and Retrieval*, pp. 162-165, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[27] Jinbiao Yu, and Liping Jing, "A Joint Multi-Task CNN for Cross-Age Face Recognition," *25th IEEE International Conference on Image Processing*, pp. 2411-2415, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[28] Jie Fang et al., "Muti-Stage Learning for Gender and Age Prediction," *Neurocomputing*, vol. 334, pp. 114-124, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[29] Se Hyun Nam et al., "Age Estimation by Super-Resolution Reconstruction Based on Adversarial Networks," *IEEE Access*, vol. 8, pp. 17103-17120, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[30] Zhifei Zhang, Yang Song, and Hairong Qi, "Age Progression/Regression by Conditional Adversarial Autoencoder," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5810-5818, 2017. [Google Scholar] [Publisher Link]

[31] Joyassree Sen et al., "Face Recognition using Deep Convolutional Network and One-shot Learning," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 4, pp. 23-29, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] T. Cootes, FG-Net Face and Gesture Recognition Network, 2014.

[33] Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu, "Cross-Age Reference Coding for Age-Invariant Face Recognition And Retrieval," *Computer Vision – ECCV*, vol. 8694, pp. 768–783, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[34] Pooja G Nair, and R. Sneha, "A Review: Facial Recognition using Machine Learning," *International Journal of Recent Engineering Science*, vol. 7, no. 3, pp. 85-89, 2020. [Google Scholar] [Publisher Link]

[35] Zhenxing Niu et al., "Ordinal Regression with Multiple Output CNN for Age Estimation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4920-4928, 2016. [Google Scholar] [Publisher Link]

[36] Anju, and Anu Rani, "Face Recognition using PCA Gobor Filter SVM Techniques," *International Journal of Computer & Organization Trends,* vol. 6, no. 3, pp. 20-23, 2016. [Publisher Link]

[37] Opeyemi Oyelesi, and Akingbade Kayode Francis, "Face Recognition for Access Control using PCA Algorithm," *SSRG International Journal of VLSI & Signal Processing*, vol. 4, no. 2, pp. 22-27, 2017. [CrossRef] [Publisher Link]

[38] Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Arxiv Preprint Arxiv:1409.1556*, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[39] Kaiming He et al., "Deep Residual Learning for Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 770–778, 2015. [Google Scholar] [Publisher Link]

[40] Gao Huang et al., "Densely Connected Convolutional Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 4700–4708, 2018. [Google Scholar] [Publisher Link]

[41] Linkai Li et al., "Integrated Access Control System of Face Recognition and Non-Contact Temperature Measurement Based on Arduino," *International Journal of Computer and Organization Trends*, vol. 12, no. 2, pp. 1-5, 2022. [CrossRef] [Publisher Link]

[42] Jayamala Kumar Patil et al., "A Novel Method of Traffic Rule Compliance Detection using Morphological Operations at Toll Booths," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 1, pp. 148-159, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[43] Nazia Begum, and Syed Noorullah, "Biometric Uncorrelated Face Recognition using Unsupervised Learning," *International Journal of Computer Engineering in Research Trends,* vol. 7, no. 3, pp. 13-18, 2020. [CrossRef] [Publisher Link]