Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.17.(1) 2025,pp.Comp 145–156

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Gender Classification from Human Face Images Using Deep Learning Based on MobileNetV2 Architecture

Nisreen Ryadh Hamza

Computer Science Department, College of Computer Science and Information Technology, University of Al –Qadisiyah Al-Diwaniah, Iraq

E-mail: nesreen.readh@qu.edu.iq:

A R T I C L E I N F O

Article history:
Received: 19 /1/2025
Rrevised form: 3 /2/2025
Accepted : 16 /2/2025
Available online: 30 /3/2025

Keywords:

Deep learning

Transfer learning

CNNs

MobileNetV2

Classification

A B S T R A C T

A person's face provides information about many identifying characteristics, including age, gender, and race. Among these characteristics, gender prediction has drawn much attention due to its many applications and use cases. Using a face to determine an individual's gender and assign it to the appropriate category of "male" or "female" is called gender recognition. Typically, there are multiple steps in the process, such as face detection and feature extraction to record the distinctive features of the face. Efficient feature extraction for gender classification is possible through a number of approaches, including deep learning-based convolutional neural networks (CNNs), which have shown excellent performance in learning hierarchical representations directly from raw pixel data. Previous attempts at gender recognition have focused on several static physical features, such as fingernails, body shape, hand shape, eyebrow, face, etc. In this study, we propose a gender classification using deep learning from human face images. This study uses a deep learning model based on the MobileNetV2 architecture for gender classification. The model was pre-trained and fine-tuned on the dataset using transfer learning techniques. During training, the model was optimized using Adam optimizer with focus loss function and learning rate scheduler. The dataset name is the largest gender-specific face recognition dataset from kaggle with sample images (man and woman). Experimental results show perfect performance and F1score value of 96%.This means that the model achieves excellent performance in the balance between precision and recall, and it is close to perfect performance on many classification tasks.

MSC..

https://doi.org/ 10.29304/jqcsm.2025.17.11970

1.Introduction

Soft biometric characteristics are helpful for identifying, validating, and defining individuals [1]. These features were categorized using facial and body geometry [2]. The two best examples of soft biometrics are age and gender estimation ; Intelligent Systems IS can be utiling gender data in the realms of biometric-based access control, smart

∗Corresponding author

Email addresses:

Communicated by 'sub etitor'

environments, and healthcare. Based on gender data, An improved user experience can be achieved by configuring Intelligent Technologies" IT" in a smart area [3]. Fig.1 displays an overview of the face-based estimating system.

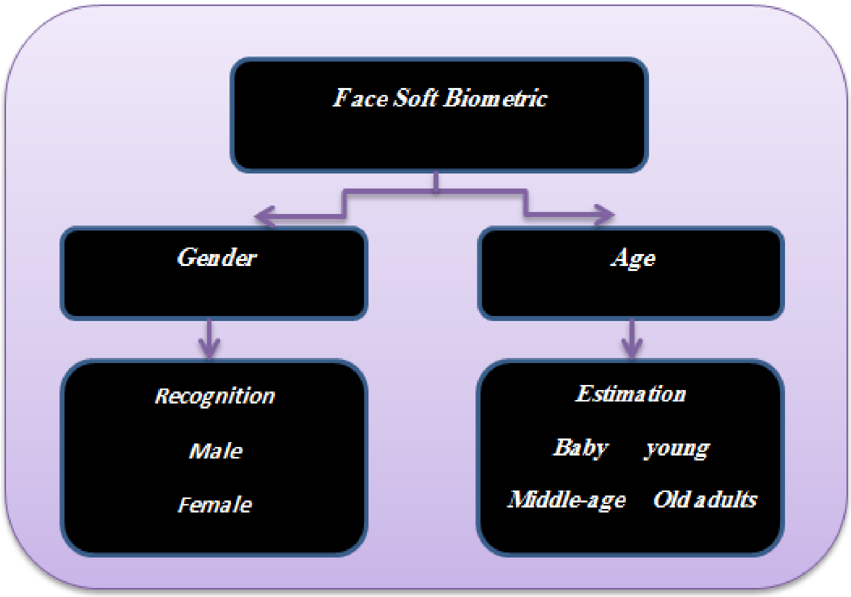

Fig. 1.face-based estimation method

Gender prediction of the face and age estimation applications include, but are not limited to, non-invasive forensic profile evaluation of criminals and victims, surveillance of specific age and gender groups, law enforcement, human-computer interaction, access control, interactive systems, and more. It can be applied to surveillance and access control in a number of ways, such as restricting access to forbidden areas for individuals of a particular sex or age; allowing access to websites based on age or gender; controlling access to web applications; and controlling access to dangerous areas like theme parks, etc. [4]. Vending machines in Japan, for instance, are used to suggest drinks (alcohol, smoke packets) based on the customers' facial adult assessment[5]. It can be applied to any crowd access violation or demographic analysis in commercial CCTV applications. For instance, Certain genders are prohibited in buses, metros, rail compartments, bathrooms, and hostels; travelers or visitors may be automatically redirected and subject to surveillance for any legal infractions. The quantity of users who have profiles (adult, young, juvenile, male, and female) in particular areas, such as public spaces, shopping centers, banks, etc., can also be determined using these forecasts for business planning and sales and marketing strategies. Electronic boards that cater to age and gender groupings that are constantly changing can use it for targeted advertising [6]. Age and gender estimation have recently been incorporated to smartphones as entertainment features. These are useful for automatically rearranging albums to manage functions like rearranging, retrieving, and deleting the taken images based on the gender and age selection. Additionally, gender recognition can be utilized to lower the database's search index in biometric systems. Additionally, it improves the accuracy of identifying a person using facial features like age and gender [7]. Additionally, it can be utilized to create a new face image of a missing child or an elderly family member using facial age synthesis. Facial photos, fingerprints, hand skin, handwriting, voice data, and gait analysis (running, jogging, etc.) can all be used to predict a person's gender. However, anthropology studies of the face or bones can be used to predict age. Because of its easy visibility (not obscured by clothing), collectability, acceptance, and universality, the face is the most appropriate feature. Both using a deep learning-based strategy and The hand-crafted features engineering are regarded as cutting-edge methods for age and gender recognition.According to feature space and feature extraction methodology, facial gender recognition can be divided into two categories: Global and Local Features. While geometric-based approaches extract features from conspicuous facial regions like the nose, mouth, and eyebrow, appearance-based approaches view the entire face as feature space[8]. One area of AI that has made a big impact in the computer world is DeepLearning "DL" [9]. The use of several layers that offer robust abstractions and generalizations sets "DL" apart from "ML" [10].Extracting characteristics is necessary to lessen the computational burden on artificial neural networks [11]. Artificial neural networks must be stripped of the characteristics. To do this, a vast network that can give precise instructions is required. The available processing and memory capacity allow for the successful training of large networks. There are significant benefits to using the same deep learning network for both extraction of features and classifications[12]. In Machine Learning "ML", Transfer Learning "TL" is the process of reusing and optimizing a model that has been trained on one task for a related but distinct task. Rather of starting from zero, you start with a model that has already been trained and use the knowledge it acquired from the initial task to adapt it to a new challenge. This enables the model to function well with less datasets, enhances performance, and cuts down on training time[13].

This study offers a successful method for gender recognition based on MobileNetV 2, a CNN that is well-known for its excellent performance and efficiency. This study demonstrates how transfer learning can achieve exceptional results in gender recognition when combined with appropriate data preprocessing and augmentation approaches.The study is structured as follows: Section 2 gives an overview of the pertinent literature; Section 3 explains the suggested technique; Section 4 presents the findings and discusses them; and Section 5 concludes the paper and suggests possible avenues for further research.

2. Related work

Researchers have presented many studies related to gender discrimination utilizing Deep Learning"DL" and Machine Learning"ML" with a variety of databases, but the database used in our research was not found in more than one study mentioned among the studies and the results were compared with it.

To generate texture descriptors, two methods are utilized: "LBP" and "HOG". By using "HOG" attributes with Adaboost and Random Forest classifiers, Yildirim et al. achieved respective ratings of precision of 85.6% and 92.3%[14]. With "LBP"characteristics, the "SVM "linear performs better. While [15] obtained 79.3 % accuracy with "SVM" and "LBP" and "FPLBP". Adience data is gathered from every ages in an unrestricted setting. Alexandre [16] obtained 99.07 % precision on "FERET" faces database captured in a restriction context using the "LBP-SVM" Linear kernel. These findings demonstrate that LBP is only appropriate for images with constraints. Images of children's faces with few characteristics that indicate gender are also included in the Adience dataset. It is also more difficult to identify because of the unregulated environment. CNNs, are employed in the Face and Gender Recognition System proposed by Yuxiang Zhou, Hongjun Ni, Fuji Ren, and Xin Kang. The modules for gender and facial recognition are the two parts of the system. Pre-trained CNN is used by both the gender and face recognition modules to extract features from the image. To be more precise, we train CNN in the module for facial recognition using the public datasets "LFW, "YTF)", and "VGGFace2" [17].By employing CNNs, Kunal, Muskan, Anupma, Rachna, and Preeti saw an improvement in performance on these tasks. We gathered data from "Google" and "IMFDB." They used the "Keras" library to develop a CNN network . After that, three complicated layers were added. Two more CNN layers with an activation function of ReLU were added. A flatten layer was used to make an output flat [18]. The most advanced techniques for determining a person's age and gender from their face were compared by Tania, Paolo, Luigi, and Valerio. Some of these methods suggested new network topologies or the inclusion of additional elements to models that were already established[19]. A new CNN approach based on the efficientnet architecture was presented by Guramritpal Singh, Keshav, and Palvinder. The proposed method outperforms other models when tested and validated on the unfiltered and in-the-wild dataset, and it shows better results and generalization ability than the previous approaches on the adience and other unconstrained face images dataset for age and gender classification [20]. An analysis of models for predictions using one attribute and multi-attribute (attributes : gender & age) was presented by Sandeep Kumar Gupta and Neeta Nain. They provide An outline of both traditional and DL techniques for face age estimation and gender classificationcreated thus far, along with an examination of their advantages, disadvantages, and suggestions for further study. Additionally, it includes the databases that are used to compare findings and their attributes in both limited and unconstrained environments[21]. Neha Sharma, Reecha Sharma, and Neeru Jindal presented a study that used an enhanced CNN to determine both age and gender. Improved computational performance by the suggested model, which may also be applied to other datasets. The core component of a CNN is the convolutional layer. CNNs are made up of neurons that can learn weights and biases, just like other neural networks. The "UTKFace", "IMDB-WIKI", "FG-NET", and "CACD" datasets serve to evaluate the proposed model. Extensive simulations demonstrate the effectiveness of the suggested technique in estimating age with greater accuracy than current algorithms[22]. Aqil Muhammad, Dian Pratiwi ,Agus Salim used Convolutional Neural Networks (CNN) for face classification with a dataset of 27,167 images (17,678 males and 9,489 females). To address data imbalance, the images were adjusted for gender equality. The dataset was divided into two groups to measure differences based on the position and background of the images [23]. İsmail AkgülI presented in the field of AI, two cutting-edge "DCNN" models have been created to estimate both gender and age on a dataset of imbalanced faces. First, a novel model was created dubbed the fast description network "FINet". Second, It reduced the number of parameters in the model structures. These models, along with the "FINet" model, were joined together to form a model called "INFINet". The models were contrasted with other cutting-edge "DCNN "models in AI. TWith the "INFINet" model, the best accuracy results were achieved. [24].

Compared to previous studies, only one study used the Biggest gender/face recognition dataset [25] . The dataset was divided into two groups to measure differences based on the position and background of the photos. The first group achieved an accuracy of 73.33%, while the second group reached 84.34%.

3. Proposed System

Recently, gender-based facial recognition has drawn a lot of attention and is becoming more and more important in people's social life. The design of every system is crucial as it explains works it and the exact steps that must be taken to get the desired result. The proposed gender recognition system uses face images which are important and well-known biometric traits to categorize individuals as either men or women. The phases of the suggested identification system for gender recognition of individuals based on face photographs are shown in Fig.2, which is a block diagram of the classifier construction procedures for the gender recognition system.

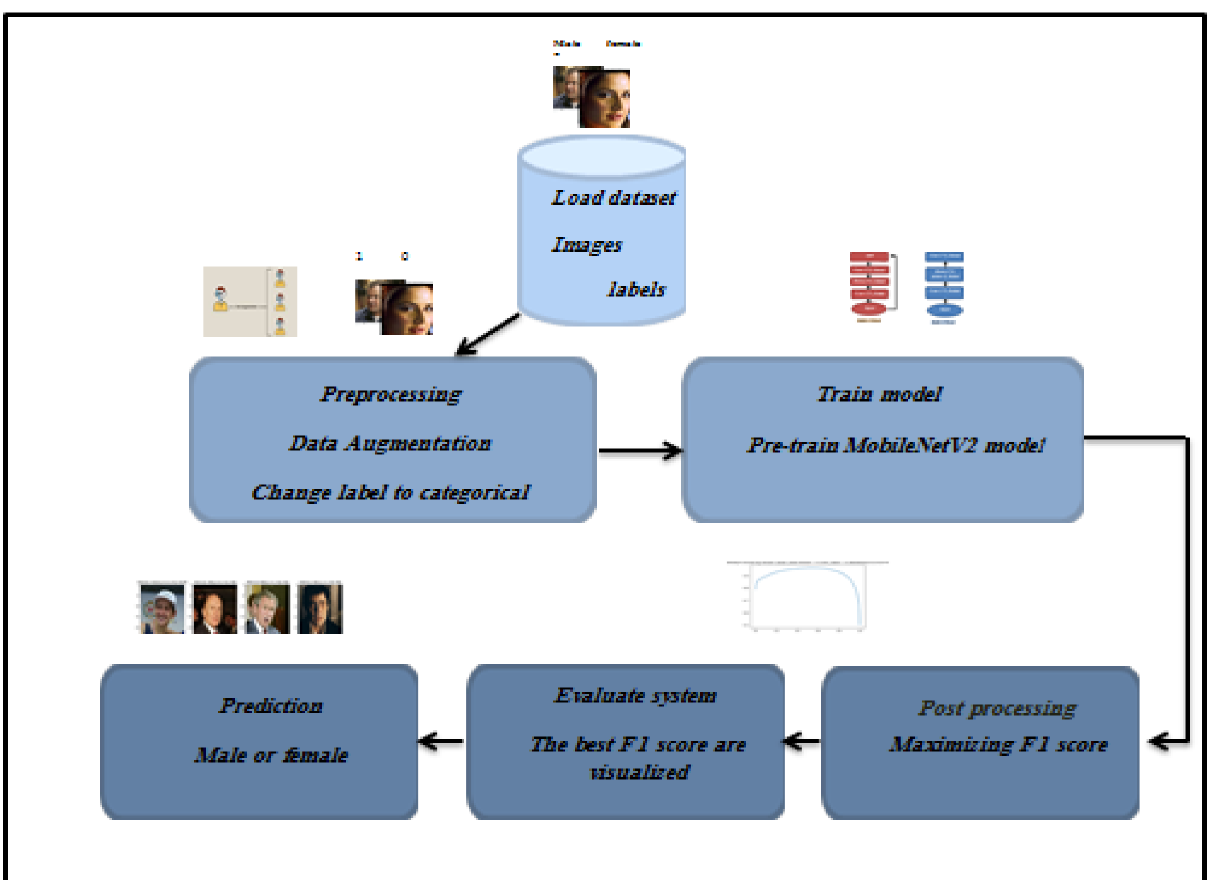

Fig. 2 . Framework of classification system

3.1 DATA SET

The gender image face dataset utilized in this study for gender identification and verification based on facial features will be explained in this part. The dataset, titled "Biggest gender/face recognition dataset from Kaggle," was developed in 2021 and includes cropped face photos [26]. The 27167.jpg files in this dataset, which is used to build gender recognition models, include 9489 images of women and 17678 images of men's faces. Fig.3 displays the photos that were taken in actual settings. The gender distribution in the largest gender face recognition dataset is displayed in Fig. 4. Matching face photos to identify gender (male or female) is the dataset's primary issue.

Figu.3. *Examples of images (men and women) from*

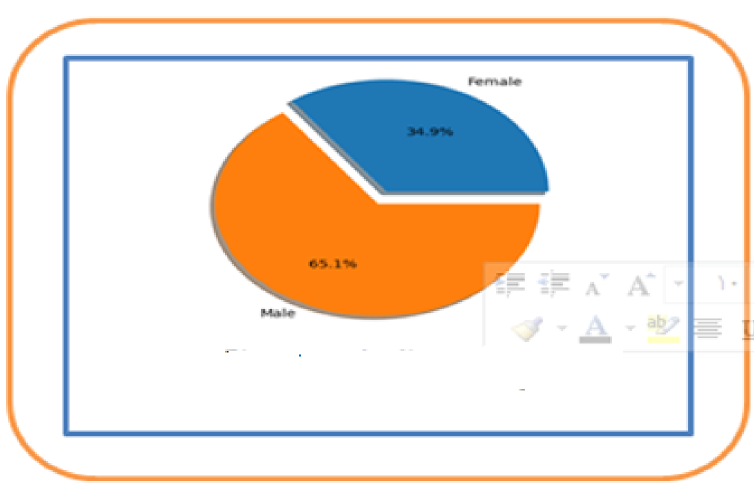

Fig.4. Gender Distribution In Biggest Gender/Face Recognition Dataset

3.2 MODEL ARCHITECTURE

MobileNetV2 is a CNN design intended to perform well on mobile devices. The first fully convolution layer with 32 filters is followed by 19 residual bottleneck layers in the design of MobileNetV2 [27]. As seen in Fig.5 , It contains two distinct types of blocks. The blocks are composed of three layers, there are conv. 1x1 using "ReLU6.". The second layer has a depth-wise "Dwise 3*3". The third layer contains a conv. 1x1 linear. One stride residual block makes up the initial tier. Another residual block with stride 2 is the second layer, which is used for shrinking.

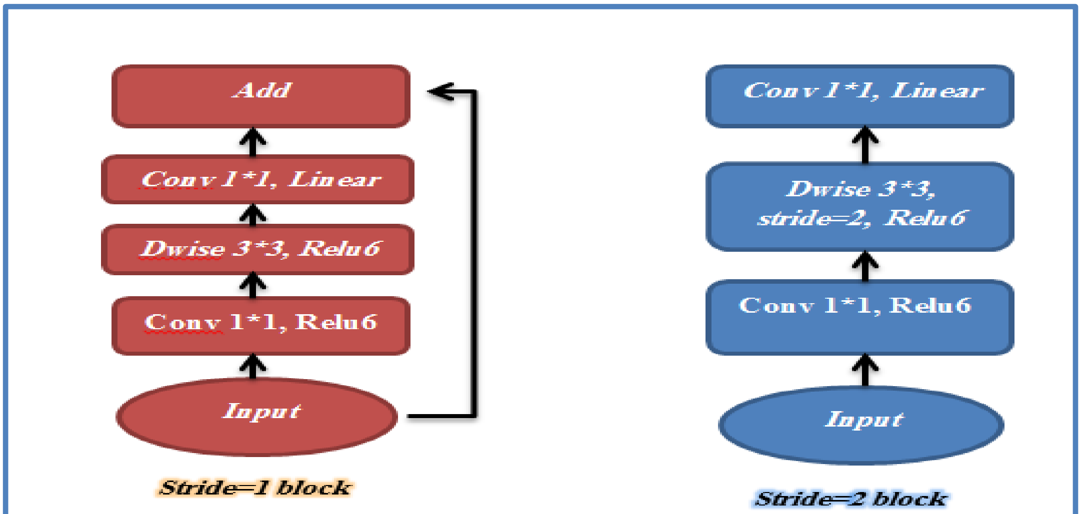

Fig.5. MobileNetV2 blocks architecture

Utilizing the Adam optimizer, which starts with a slightly larger learning rate (le-4) initially and reduces it dynamically based on Reduce LR On Plateau when validation loss pleateaus. Using dataset containing of men and women images . Images are loaded are assigned (1 for man (male), 0 for woman (female)). Training make up 80% whereas testing make up 20%. Figure 6 (a)image label (b) gender distribution. Pre-trained MobileNetV2 model is loaded and modified for binary classification (male or female). Training and validation losses and F1 scores are plotted over epochs.

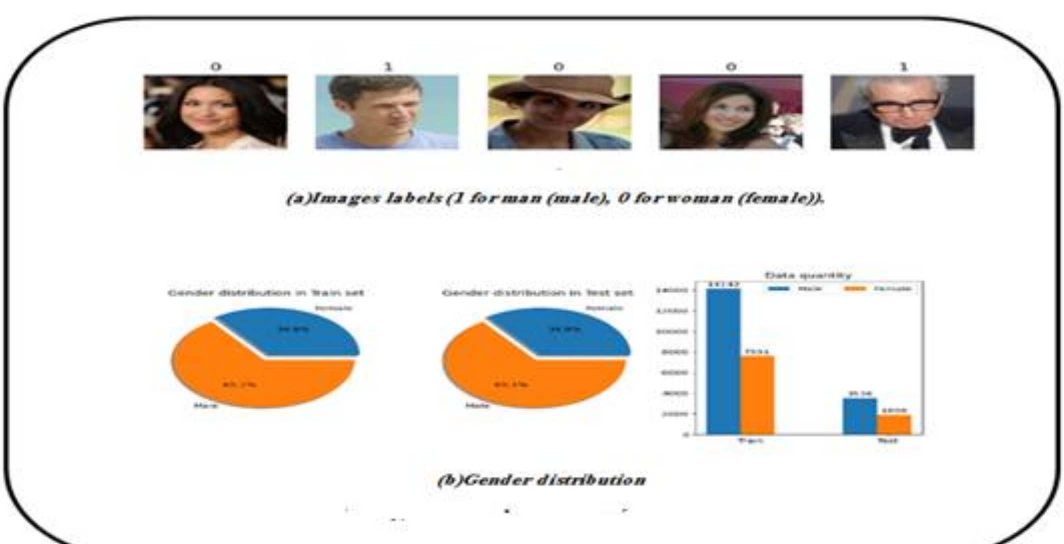

Fig.6. Image Label & Gender Distribution

4. Results and Discussion

The implementation environment description is significant in evaluating the system behaviour and performance and the way it works. Our system runs on Python 3.11 and was trained with an Intel Core i7-7700HQ CPU and 6GB of RAM. Additionally, the Lenovo PC contains a 6 GB Nvidia GEFORCE GTX, which speeds up neural networks' computations on 64-bit Windows 10.

In this research used F1-score as metric to evaluate gender recognition system. F1-score Metric "F1M" is typically employed when the dataset is imbalanced. When the distribution of classes is not balanced, the model's performance is assessed using the F-measure. More seriously, the better the findings, the higher the F-measure. In terms of mathematics:

$$\boldsymbol{F1 - SCORE = 2 * \left(\frac{P \; * \; R}{P \; + \; R}\right)} \qquad (1)$$

Where p represent Precision Metrics "PM" and R represent Recall Metrics"RM" . The ratio of True Positives"TP" to Positive states "TP+FP" is known as Precision Metrics"PM".False Positives"FP" are situations defines as Positive"P" but are actually Negative"N. This can be represented mathematically as:

$$\boldsymbol{P = \frac{Tp}{(Tp \; + \; Fp)}} \qquad (2)$$

Where TP : "True Positives" and FP : "False Positives". Recall Metrics"RM" is the capacity to find each pertinent example in a dataset. the proportion of all accurate observations in the sample space to the number of accurate forecasts. In terms of mathematics:

$$\boldsymbol{R = \frac{Tp}{(Tp \; + \; FN)}} \qquad (3)$$

Where TP: "True Positives" and FN: "False Negatives". In this research training a gender classification model using a dataset containing images of faces labeled with their corresponding gender. The steps of suggested system in detail:

- Loading images (men and women) from the database and assigned (1 male, 0 female). Training make up 80% whereas testing make up 20%
- A pre-trained MobileNetV2 model is loaded and modified for binary classification (male or female). Training and validation losses and F1 scores are plotted over epochs. Table 1 and fig.7 illustrate this.
- The model's predictions are evaluated using different thresholds to maximize the F1-score. Threshold values ranging from 0 to 1 are evaluated, and the threshold that maximizes F1 Score is selected. This threshold is used to classify the predictions into binary ategories (male or female) according to the model's probability scores. The result of model Binary F1Score by thresh value (best thresh = 0.549 | best F1Score = 0.96%). Fig. 8 shows Postprocessing - Maximize F1-score.
- Prediction: Random test samples images are selected, and their predictions (male or female) along with their probabilities and true labels are displayed using Matplotlib. Fig. 9 shows predication result.

Table 1 - Training and validation losses and F1 scores over epoch.

EPOCH	TRAIN		VAL	
	[LOSS]	[BIARY F1-SCORE]	[LOSS]	[BIARY F1-SCORE]
EPOCH1	0.17	0.7033	0.203	0.4878
EPOCH2	0.044	0,9565	0.017	1.0
EPOCH3	0.039	0.9777	0.023	0.9767
EPOCH4	0.02	0.9743	0.082	0.8999
EPOCH5	0.054	0.9268	0.025	0.9523
EPOCH6	0.009	1.0	0.063	0.9729
EPOCH7	0.017	0.9767	0.003	1.0
EPOCH8	0.011	1.0	0.033	0.9499

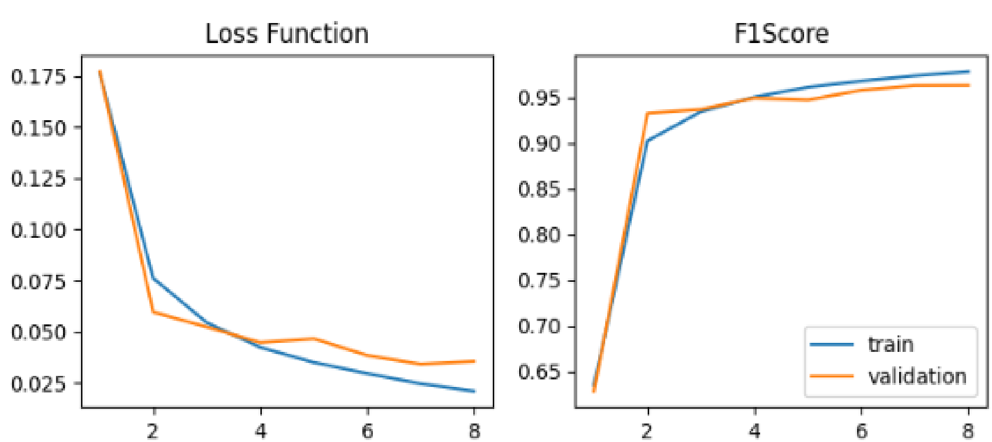

Fig.7. Training Graph

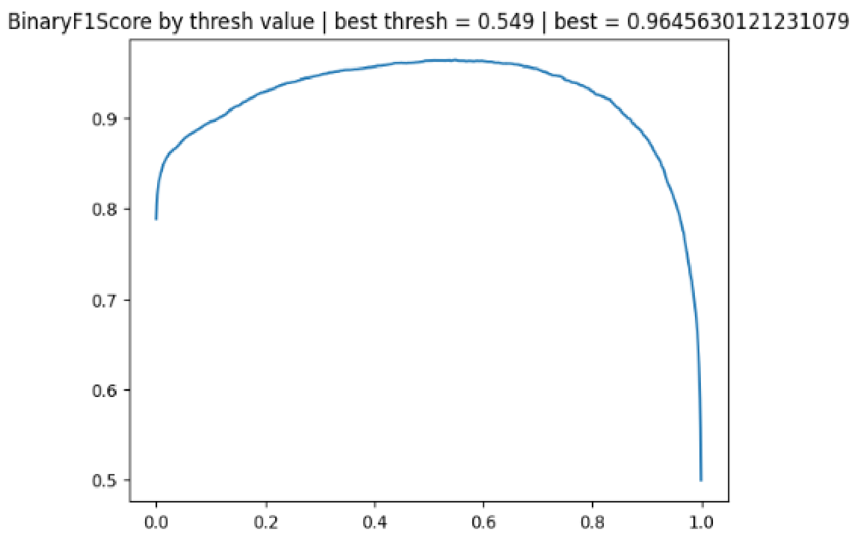

Fig. 8. Post processing - Maximize F1-Score

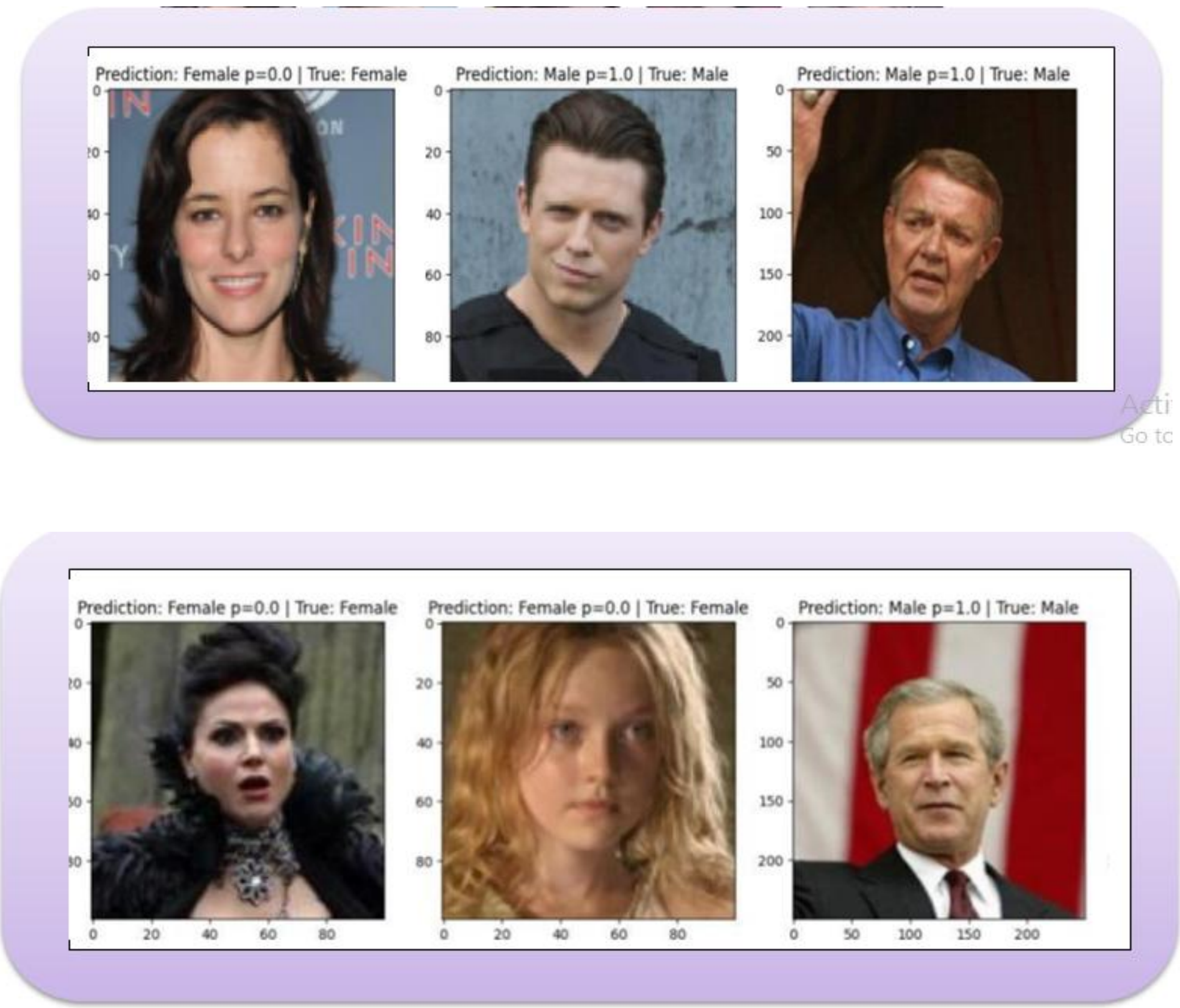

Fig.9 .prediction results

5. Conclusion

The main goal of the study is using neural networks to determine human gender, whether it be Male"M" or Female"F".The issue of the image's environment-caused changing illumination and dull angles was resolved by the pre-processing process, which applied "data augmentation" to increase the efficiency of the system that was suggested. The testing results of choosing the pretrained algorithm with multiple tasks considerably enhanced the system's performance, according to the results of the suggested system that uses artificial intelligence techniques (pre-trained DL models) to estimate whether the gender recognition is Male "M" or Female "F". Utilizing the MobileNetV2 model, testing achieved 96% outcomes for accuracy. This means that the model performs very well in a balance between precision (not having many false positive predictions) and recall (the ability to detect most true positives). In many applications, 96% F1 Score is considered excellent performance, as it means that the model is able to classify the data effectively. Compared to previous studies, only one study used the Biggest gender/face recognition dataset . The dataset was divided into two groups to measure differences based on the position and background of the photos. The first group achieved an accuracy of 73.33%, while the second group reached 84.34%. These findings show that the recommended method performs the best in terms of classification when compared to previous research.

REFERENCES

[1]. P. TERHÖRST, D. FÄHRMANN, N. DAMER, AND F. KIRCHBUCHNER(2021). ON SOFT-BIOMETRIC INFORMATION STORED IN BIOMETRIC FACE EMBEDDINGS, VOL. 3, NO. 4, PP. 519–534, DOI: 10.1109/TBIOM.2021.3093920.

[2]. A. JAIN AND V. KANHANGAD(2017). GENDER CLASSIFICATION IN SMARTPHONES USING GAIT INFORMATION ANKITA, EXPERT SYST. APPL., DOI: 10.1016/J.ESWA.2017.10.017.

[3]. I. SIDDIQI, C. DJEDDI, AND A. RAZA (2014). AUTOMATIC ANALYSIS OF HANDWRITING FOR GENDER CLASSIFICATION, , DOI: 10.1007/S10044- 014-0371-0.

[4]. OSMAN OF, YAP MH (2018). COMPUTATIONAL INTELLIGENCE IN AUTOMATIC FACE AGE ESTIMATION: A SURVEY, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE 3(3):271–285.

[5]. (2017).JAPANESE SMOKERS TO FACE AGE TEST, HTTP://NEWS.BBC.CO.UK/2/HI/ASIA-PACIFIC/7395910.STM.

[6]. GUO G (2012). HUMAN AGE ESTIMATION AND SEX CLASSIFICATION. IN: VIDEO ANALYTICS FOR BUSINESS INTELLIGENCE, SPRINGER, PP 101–131.

[7]. MAKINEN E, RAISAMO R (2008). AN EXPERIMENTAL COMPARISON OF GENDER CLASSIFICATION METHODS, PATTERN ¨ RECOGN LETT 29(10):1544–1556.

[8]. BALUJA S, ROWLEY HA (2007).BOOSTING SEX IDENTIFICATION PERFORMANCE, INT J COMPUT VIS 71(1):111–119 .

[9]. I. YASAKA AND O. ABE (2018). DEEP LEARNING AND ARTIFICIAL INTELLIGENCE IN RADIOLOGY: CURRENT APPLICATIONS AND FUTURE DIRECTIONS, PLOS MED., VOL. 15, NO. 11, PP. 2–5,, DOI: 10.1371/JOURNAL.PMED.1002707.

[10]. Y. LECUN, Y. BENGIO, AND G. HINTON (2015). DEEP LEARNING, NATURE, VOL. 521, NO. 7553, PP. 436–444, DOI: 10.1038/NATURE14539.

[11]. I. DENG AND D. YU(2013. DEEP LEARNING: METHODS AND APPLICATION.

[12]. ARNOLD, S. REBECCHI, AND S. CHEVALLIER (2011). AN INTRODUCTION TO DEEP LEARNING TO CITE THIS VERSION, EUR. SYMP. ARTIF. NEURAL NETWORKS, NO. JANUARY.

[13]. ZHUANG, FUZHEN, ZHIYUAN QI, KEYU DUAN , DONGHO XI, YONGCHUN ZHU, HUIXIONG, AND QING HE(2020), ACOMPREHENSIVE SURVY ON TRANSFER LEARNING ,PROCESSINGS OF THE IEEE .

[14]. YILDIRIM ME, INCE OF, SALMAN YB, SONG JK, PARK JS, YOON BW (2016) ,GENDER RECOGNITION USING HOG WITH MAXIMIZED INTER-CLASS DIFFERENCE. IN: VISIGRAPP (3: VISAPP), PP 108–111.

[15]. HASSNER T, HAREL S, PAZ E, ENBAR R (2015). EFFECTIVE FACE FRONTALIZATION IN UNCONSTRAINED IMAGES. IN: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP 4295–4304.

[16]. ALEXANDRE LA (2010). GENDER RECOGNITION: A MULTISCALE DECISION FUSION APPROACH. PATTERN RECOGNITION LETTERS ,31(11):1422–1427.

[17]. YUXIANG ZHOU, HONGJUN NI,FUJI REN, XIN KANG (2019). FACE AND GENDER RECOGNITION SYSTEM ASED ON CONVOLUTIONAL NEURAL NETWORKS ,IEEE.

[18]. KUNAL JAIN, MUSKAN CHAWLA, ANUPMA GADHWAL, RACHNA JAIN & PREETI NAGRATH (2020) . AGE AND GENDER PREDICTION USING CONVOLUTIONAL NEURAL NETWORK, CONFERENCE PAPER.

[19]. TANIA DI MASCIO, PAOLO FANTOZZI, LUIGI LAURA & VALERIO RUGHETT (2021). AGE AND GENDER (FACE) RECOGNITION: A BRIEF SURVEY,CONFERENCE PAPER, SEPTEMBER .

[20]. GURAMRITPAL SINGH SAGGU, KESHAV GUPTA & PALVINDER SINGH MANN (2021). EFFICIENT CLASSIFICATION FOR AGE AND GENDER OF UNCONSTRAINED FACE IMAGES, CONFERENCE PAPER , 14 DECEMBER .

[21]. SANDEEP KUMAR , NEETA NAIN1 (2022). REVIEW: SINGLE ATTRIBUTE AND MULTI ATTRIBUTE FACIAL GENDER AND AGE ESTIMATION.

[22]. NEHA SHARMA, REECHA SHARMA & NEERU JINDAL (2022). FACE-BASED AGE AND GENDER ESTIMATION USING IMPROVED CONVOLUTIONAL NEURAL NETWORK APPROACH.

[23]. AQIL MUHAMMAD, DIAN PRATIWI ,AGUS SALIM (2023). PENERAPAN METODE CONVOLUTIONAL NEURAL NETWORKS PADA PENGENALAN GENDER MANUSIA BERDASARKAN FOTO TAMPAK DEPAN, JURNAL KOMTIKA.

[24]. İSMAIL AKGÜL (2024) .DEEP CONVOLUTIONAL NEURAL NETWORKS FOR AGE AND GENDER ESTIMATION USING AN IMBALANCED DATASET OF HUMAN FACE IMAGES.

[25]. AQIL MUHAMMAD , DIAN PRATIWI AND AGUS SALIM (2023), PENERAPAN 25. METODE CONVOLUTIONAL NEURAL NETWORKS PADA PENGENALAN GENDER MANUSIA BERDASARKAN FOTO TAMPAK DEPAN,JOURNAL KOMTIKA.

[26]. HTTPS://WWW.KAGGLE.COM/DATASETS/MACIEJGRONCZYNSKI/BIGGEST-GENDERFACE-RECOGNITION-DATASET/DATA

[27]. SANDLER, ANDREW HOWARD, MENGLONG ZHU, ANDREY ZHMOGINOV,AND LIANG-CHIEH CHEN(2018). MOBILENETV2: INVERTED RESIDUALS AND LINEAR BOTTLENECKS, IN PROCEEDINGS OF THE IEEE CONFERENCE .