University of Baghdad
College of Engineering
JE
JOURNAL OF ENGINEERING

Journal of Engineering

journal homepage: www.joe.uobaghdad.edu.iq

Volume 30 Number 1 January 2024

Face-based Gender Classification Using Deep Learning Model

Buraq Abed Ruda Hassan[1,] *, Faten Abed Ali Dawood[2]

Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq
asds.yy899@gmail.com[1], faten.dawood@sc.uobaghdad.edu.iq[2]

ABSTRACT

Gender classification is a critical task in computer vision. This task holds substantial importance in various domains, including surveillance, marketing, and human-computer interaction. In this work, the face gender classification model proposed consists of three main phases: the first phase involves applying the Viola-Jones algorithm to detect facial images, which includes four steps: 1) Haar-like features, 2) Integral Image, 3) Adaboost Learning, and 4) Cascade Classifier. In the second phase, four pre-processing operations are employed, namely cropping, resizing, converting the image from(RGB) Color Space to (LAB) color space, and enhancing the images using (HE, CLAHE). The final phase involves utilizing Transfer learning, a powerful deep learning technique that can be effectively employed to Face gender classification using the Alex-Net architecture. The performance evaluation of the proposed gender classification model encompassed three datasets: the LFW dataset, which contained 1,200 facial images. The Faces94 dataset contained 400 facial images, and the family dataset had 400. The Transfer Learning with the Alex-Net model achieved an accuracy of 98.77% on the LFW dataset.
Furthermore, the model attained an accuracy rate of 100% on both the Faces94 and family datasets. Thus, the proposed system emphasizes the significance of employing pre-processing techniques and transfer learning with the Alex-Net model. These methods contribute to more accurate results in gender classification. Where, the results achieved by applying image contrast enhancement techniques, such as HE and CLAHE, were compared. CLAHE achieved the best facial classification accuracy compared to HE.

Keywords: Alex-Net, CLAHE, Deep learning, Gender Classification.

*Corresponding author
Peer review under the responsibility of University of Baghdad.
https://doi.org/10.31026/j.eng.2024.01.07

Article received: 15/04/2023
Article accepted: 25/12/2023
Article published: 01/01/2024

تصنيف الجنس بالاعتماد على الوجه باستخدام نموذج التعلم العميق

براق عبد الرضا حسن[1]،*، فاتن عبد علي داود[2]

قسم علوم الحاسبات، كلية العلوم، جامعة بغداد، بغداد، العراق

الخلاصة

تصنيف الجنس هو مهمة حرجة في رؤية الحاسوب. تحمل هذه المهمة أهمية كبيرة في مجموعة متنوعة من المجالات، بما في ذلك المراقبة والتسويق والتفاعل بين الإنسان والحاسوب. في هذه العمل يتكون نموذج تصنيف جنس الوجه المقترح من ثلاث مراحل رئيسية: تتضمن المرحلة الأولى تطبيق خوارزمية فيولا-جونز لاكتشاف صور الوجه، والتي تشمل أربع خطوات: 1) Haar-Like Features ، 2) Integral Image ، 3) Adaboost Learning ، و 4) Cascade Classifier. في المرحلة الثانية، يتم استخدام أربع عمليات ما قبل المعالجة، وهي القص، وتغيير الحجم، وتحويل الصورة من (RGB) إلى (LAB)، وتحسين الصور باستخدام (HE,CLAHE). المرحلة النهائية تتضمن استخدام نقل التعلم، وهو تقنية قوية في التعلم العميق، يمكن استخدامها في تصنيف جنس الوجه باستخدام بنية (Alex-Net). يتضمن تقييم أداء النموذج المقترح لتصنيف الجنس ثلاث مجموعات بيانات: مجموعة بيانات LFW التي تحتوي على 1200 صورة وجه، ومجموعة بيانات Faces94 التي تحتوي على 400 صورة وجه، ومجموعة بيانات العائلة التي تحتوي على 400 صورة وجه. حقق نقل التعلم باستخدام نموذج Alex-Net دقة بلغت 98.77% على مجموعة بيانات LFW. وعلاوة على ذلك، حقق النموذج معدل دقة بلغ 100% على كل من مجموعة بيانات Faces94 ومجموعة بيانات العائلة.وبالتالي ، يؤكد النظام المقترح على أهمية استخدام تقنيات المعالجة المسبقة ونقل التعلم باستخدام نموذج Alex-Net. تساهم هذه الأساليب في نتائج أكثر دقة في تصنيف الجنس. حيث تم مقارنة النتائج التي تم تحقيقها من خلال تطبيق تقنيات تحسين تباين الصور ، مثل HE و CLAHE. حقق CLAHE أفضل دقة في تصنيف الوجه مقارنةً بـ HE.

الكلمات المفتاحية: Alex-Net، CLAHE, التعلم العميق, تصنيف الجنس.

1. INTRODUCTION

The face is the essential focus of attention in society and is usually used as identification. The facial image is considered to be a reliable indicator for gender detection. Every human has a distinctive face that may be used to identify them from other people **(Kanget al., 2017; Joodi et al., 2023)**.

The problem statement of gender classification from face image can be summarized into Pose Variations, Varying Illumination Conditions, and Background Variation. Gender classification plays a crucial role in demographic analysis and behavior tracking. Gender-aware systems also find application in social robotics and Human-Computer Interaction (HCI) **(Lahariya et al., 2021)**. They worked to design a framework based on deep learning with transfer learning and fine-tuning for gender classification from face images, remove the background from the face image, and use it in the proposed model by applying theViola–Jones algorithm, enhance the face image and use it in the proposed model by applying several preprocessing techniques, which are cropping, resizing, Images Convert from RGB to LAB, and Contrast enhancement using Contrast Limited Adaptive Histogram Equalization (CLAHE). Finally, using transfer learning Alex-Net to train the classification model for face images and evaluate the best classification result using Accuracy. In the literature review, several methods for gender classification from facial images have been proposed.

(Dhomne et al., 2018) proposed an efficient convolutional network architecture called a VGG-Net that may be applied in severe situations with a shortage of training data for learning a D-CNN based on this architecture. Particularly for gender identification, this model is

learned. Several face images from the celebrity and LFW datasets were used in the experiments. The test results have shown a gender classification accuracy of 95%.

(VenkateswarLal et al., 2019) suggested building a method for face recognition by feature descriptors and preprocessing face image techniques, which works well in natural settings. The mined color descriptors, texture, and preprocessed face images were copied and categorized using the support vector machine technique. The experimental consequences of the proposed method are presented using various data samples, such as LFW and FERET datasets. Due to using additional pre-processing and derived feature descriptors, the results indicate that the recommended strategy has admirable classification accuracy and a blend of pre-processing approaches. The average classification accuracy for the manipulated two data samples is 99% and 94%, respectively. **(Galla et al., 2020)** converted the visualized data to audio data to submit better sensation for blind persons by modeled feature extraction and classification. The proposed Feature extraction used the Multiscale-Invariant Feature Transform (MSIFT), while the support vector machine technique and the LASSO classifier fare used to perform feature optimization and type. The obtained classification results are illustrated as EN with an accuracy of 93.5%. LR with an accuracy of 93.2% and RR with a precision of 89.6%. **(Zaman, 2020)** propose for gender classification a custom CNN architecture based on face images into female and male. This CNN contains seven convolutional layers two fully connected layers, and between the convolutional layers are used batch normalization layers. Compared to Google-Net and Alex-Net, the number of parameters in bespoke CNN is three times and 30 times less, respectively. The proposed custom CNN also produced results on par with cutting-edge techniques, producing the top result with 96% accuracy on the CelebA dataset. Custom CNN trained on the CelebA dataset correctly classified Gender for the IMDB and WIKI datasets with 97% and 96% accuracy, respectively. The proposed method obtained an accuracy of 95% in LFW datasets.

(Azimi and Meghdadi, 2021) proposed a real-time deep neural network model that performs gender classification more quickly and with less computation by decreasing the model parameters and computation process. The proposed model is a combination of multifold filters and is relatively light. With an accuracy of 91.4%, 95%, 95%, and 95%, respectively, the performance of this proposed method was assessed in Adience, LFW, FEI, and CAS datasets. **(Alamri et al., 2022)** proposed using machine learning techniques in the LFW dataset for face detection and gender identification. The SVM classifier, the LBPH technique, and the SIFT feature extraction method were trained and evaluated utilizing the LFW dataset. The SIFT with an SVM classifier with an accuracy of 65% is reported, while face recognition with LBPH with a precision of 88% is addressed, exceeding earlier research. Moreover, the LBPH provided a gender detection accuracy of 91%.

This work aims to successfully develop an accurate and reliable gender classification system that can have numerous practical implications. It can contribute to Demographic analysis and crowd monitoring in surveillance applications. Moreover, accurate gender Classification can enable personalized interactions and tailored experiences in the human-computer interaction field. This work serves to remove the background from the face image. Enabling its utilization in the proposed model by applying the Viola-Jones Algorithm is also a target for this work. It also aims to address challenges related to lighting conditions and expression variations using preprocessing techniques.

2. MATERIALS AND METHODS

2.1 Proposed Model for Gender Classification

The proposed model aims to classify Gender from face images and consists of three main phases. In the first phase, the Viola-Jones algorithm detects the face image. In the second phase, four steps are used as pre-processing techniques, including cropping, resizing, Image Convert from RGB to LAB, and Images Enhancement. The final phase is the proposed model of gender classification deep learning using Transfer Learning Alex-Net. The detailed overall framework for Classification according to Gender is shown in **Fig. 1**

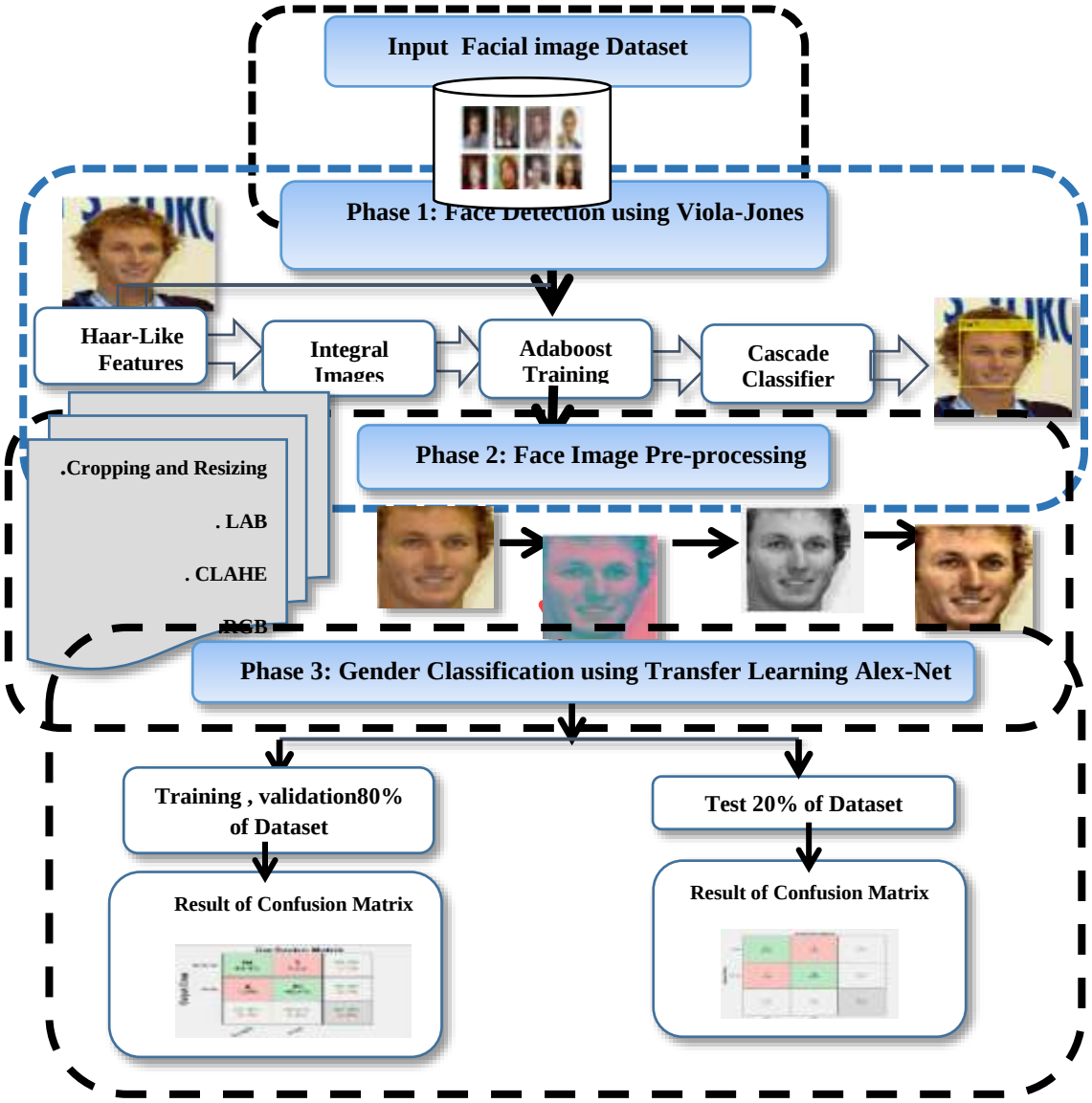


Figure 1. The overall framework for the Proposed Gender Classification

2.2 Data Collection

This step is performed for face image loading using different types of datasets. Brief descriptions of these datasets are given in the following sub-sections.

2.2.1 LFW dataset

LFWdataset is the most often used benchmark for gender classification, even though it was initially developed for unrestricted face recognition such as facial expression, background, lighting, Race, Gender, age, color, saturation, camera quality, clothing, hairstyles, and others. LFW contains 13,233 images of 5749 people (10,256 males and 2977 females) **(Chen et al., 2017; Bajrami et al., 2018)**. The dataset may be accessed and downloaded at the following website: http://vis-www.cs.umass.edu/lfw/. The dataset was created and maintained by researchers at the University of Massachusetts. **Fig. 2** shows Samples of LFW datasets**.**

2.2.2 Faces94 dataset

Each person has 20 images in total. 152 people are depicted, with a 180 by 200 pixel image resolution (portrait format). The collection has 399 images of females and 2,660 images of males **(Thepade et al., 2018; Berbar, 2022).** There are many images in this dataset. The dataset can be observed and obtained at the following web http://cswww.essex.ac.uk/mv/allfaces/faces94.html. Samples of datasets are shown in **Fig. 3**.
Three datasets of facial images are used to test and evaluate the performance of the proposed model for gender classification. The descriptions of these images are illustrated in **Table 1.**

Table 1. Descriptions of the facial images dataset

Type of data set Images.	Number of Images	Female Images	Male Images	Image type & Format	Resolution of each image
LFW	13,233	2977	10,256	RGB JPG	250 × 250
Faces94	3,059	399	2,660	RGB JPG	180 × 200
Real data	400	200	200	RGB JPG	250 × 250

2.2.3 Real dataset

The real dataset is collected from the family individual for unconstrained face classification such as facial expression, background, lighting, race, Gender, age, color, saturation, camera quality, clothing, hairstyles, and others. Real contains 400 images (200 males and 200 females). Face parts in the images were detected through the Viola-Jones face detector. Samples of the real datasets are shown in **Fig. 4.**

Figure 2. Samples of images from the dataset

Figure 3. Samples of images from the faces of 94 datasets

2.3 Face Detection Using Viola -Jones Algorithm

Face detection aims to extract the proposed system's region of interest (ROI), a person's face. This detection is used for further processing by removing features and classifying them. The algorithm that is used for face detection is the Viola-Jones algorithm.

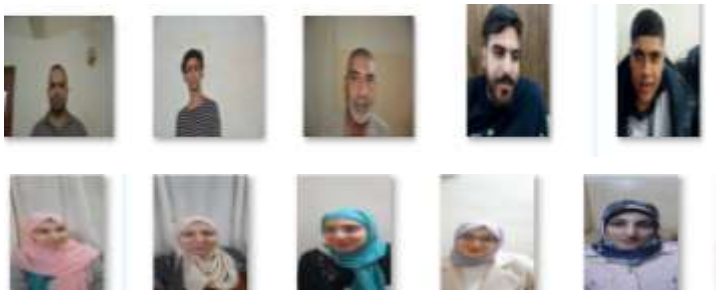


Figure 4. Samples of images from the Real dataset

It is a face-detection technique that produces speedy, accurate, and effective results **(Chaudhari et al., 2018; Suma and Raga, 2018)**. The Viola-jones algorithm consists of four main steps as follows**:**

- Haar-Like Features: The majority of the Features of human faces are identical. We can match their attributes using the Haar-like features that aid in extracting information from the input image **(Heydarzadeh et al., 2010; Hassan and Dawood, 2022)**.
- Integral Image**:** An integral image calculates Haar-like features relatively quickly **(Zhang et al., 2010; Hassan and Dawood, 2022)**.
- Adaboost: The main idea is to train many weak classifiers and then associate between these weak classifiers to acquire a strong classifier **(Peng et al., 2018; Hassan and Dawood, 2022).**
- Cascade Classifier**:** After choosing the greatest qualities in each window, now must choose which of these windows contains faces. Therefore, only the area that completes all steps is categorized as a face. The accuracy of detection increases with the number of steps in the cascade classifier **(Kang et al., 2016; Hassan and Dawood, 2022)**.

2.4 Face Image Pre-processing

The pre-processing step enhances the image quality to achieve the best results. A detailed explanation of the pre-processing steps applied to the input images follows.

2.4.1 Image Cropping

The image's cropping is done to reduce the quantity and keep only the face region **(Ciocca et al., 2007)**. The cropping process is done by taking the values from each side of the face bounding box that has been formed. The result of Image Cropping is illustrated in **Fig. 5.**

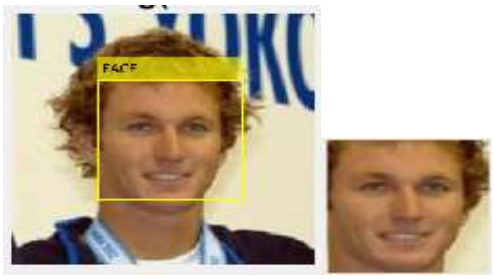


Figure 5. a) Image before Cropping, **b)** Image after Cropping

2.4.2 Image resizing

In this step, the image size is resized before it enters the Alex-Net. The image must be the same size as the other images in the database. After the cropping, the images are of different sizes varying from 130x118 to 132 x120 pixels **(Unal et al., 2016)**. Therefore, all photos in the database were resized to a constant value (227x227) to reduce the calculation cost and complexity of the problem.

2.4.3 Image Converted from RGB to L*A*B*

The International Commission on Illumination first identified and defined this color space using CAE. One channel in this color space is designated for Luminance (Lightness), while channels a and b are chromaticity layers. Regarding the red-green and blue-yellow axes, the layer color is represented by the a* and b* layers, respectively **(KaviNiranjana et al., 2015).** The cube root of the relative luminance is used to determine the lightness correlation in CIELAB. Additionally, since uniform changes in L*A*B*color space's Components are intended to match consistent differences in perceived color, it is possible to perceive every color as a dot in a three-dimensional space, allowing one to estimate the comparative perceptual variances between any two L*A*B*hues and calculating the Euclidean distance between them **(Rathore et al., 2012)**. The process of converting images from RGB to L*A*B*, using Eq. (1).

$$\begin{bmatrix} L* = 116\, f\left(\frac{Y}{Yn}\right) - 16 \\ a* = 500[f(X/Xn) - f(Y/Yn)] \\ b* = 200[f(Y/Yn) - f(Z/Zn)] \end{bmatrix} \quad (1)$$

X, Y, Z, Xn, Yn, and Zn are the coordinates of CIEXYZ color space. For converting images to the CIEXYZ from the RGB, use Eq. (2).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.608 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

Xn, Yn, and Zn correspond to the white value of the parameter, Eq. (3).

$$f(x) = \begin{bmatrix} x^{\frac{1}{3}} & x > 0.008856 \\ 7.787x + \frac{16}{116} & x <= 0.008856 \end{bmatrix} \quad (3)$$

Converting color reproduction from one basis to another is converting color space. It happens when transforming an image shown in a single color space to a different color space **(Rathore et al., 2012).**

2.4.4 Facial Images Enhancement Techniques

Image enhancement is the process that converts the image at the input into one that is more suited for the needed application. In this proposal, techniques are used, such as HE and CLAHE for image contrast enhancement.

- Histogram equalization(HE)

Histogram equalization is the method that re-distributes all pixel values to be as close to the intended histogram as possible **(Benitez-Garcia et al., 2011)**. Equalization of the histogram allows for increased contrast in areas of low local contrast. It automatically produces a transformation function to provide an image with a consistent histogram at the output. The equalization of the histogram is a technique for adjusting the contrast of an image by utilizing the histogram of the image. Usually, this method is applied to enhance the overall contrast of many images, primarily when near-contrast values represent the image's valid data **(Musa et al., 2018)**. As a result, the intensities can be more evenly dispersed across the histogram. To accomplish this, histogram equalization is used to efficiently spread out the most common intensity values in the data. **(Fadhil and Dawood, 2021)**. The color image and its histogram are shown in **Fig. 6**, followed by the equalized image and its histogram.

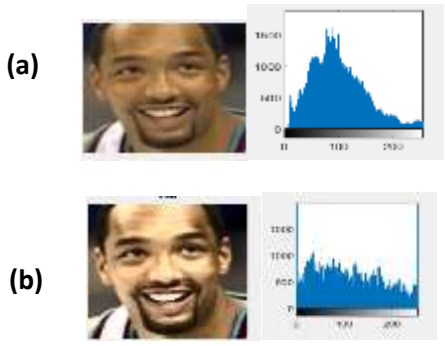


Figure 6. (a): Original Image, Histogram of the original image.
(b) Image enhancement using the HE method, Histogram of HE image

- Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is a well-known technique for enhancing local contrast that has proven effective and practical in several applications **(Dawood and Abood, 2018)**. The fundamental concept of CLAHE is to conduct histogram equalization on non-overlapping portions of the image, utilizing interpolation to fix border discrepancies **(Musa et al., 2018)**. Moreover, CLAHE features two crucial hyper-parameters: the clip limit (CL) and the number of tiles (NT). The first one, called CL, is a numerical number that regulates noise augmentation. Once the histogram for each sub-area has been determined, it is redistributed so that the height does not exceed a chosen "clip limit." To execute the equalization, the cumulative histogram is then constructed **(Wu et al., 2017)**. The second (NT) integer value determines the number of equal-sized non-overlapping sub-areas; depending on its value, the image is divided into several sub-areas of this kind. The proposed method applies CLAHE to LAB color space to improve face images. **Fig.7** shows the color image and histogram, then indicates the CLAHE image, which is the histogram of the CLAHE image.

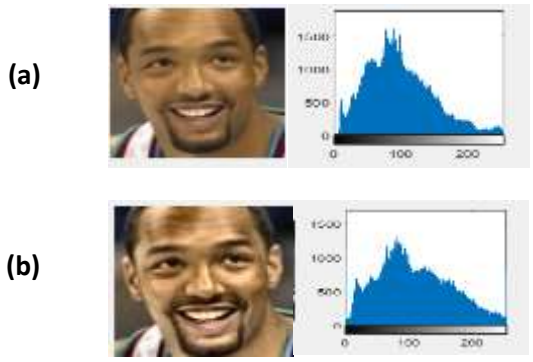


Figure 7. (a): Original Image, Histogram of the original image.
(b): Image enhancement using the CLAHE method, Histogram of CLAHE image.

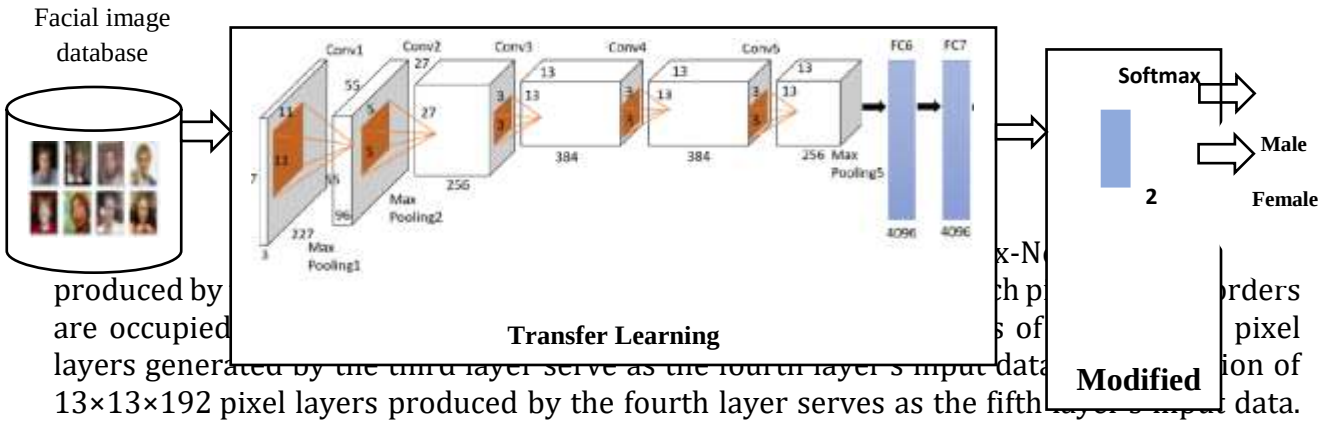
2.5Gender Classification

Classification is the final phase in the proposed model. This phase requires enhancing the performance of gender classification from face images by using Transfer Learning: Alex-Net. In this section we will discuss the basic methods of data needed to train Alex-Net. Alex-Net is given the samples in the form of a fixed-size image. The input image is colored and sized 227×227×3.

2.5.1 Transfer Learning Alex-Net Model

A practical method for resolving a classification issue is transfer learning. In this work, we applied Transfer Learning Alex-Net for gender Classification based on the face image **(Anusriet et al., 2021)**. As shown in **Fig. 8**, The Alex-Net architecture contains three pooling, five convolutional layers, and three fully connected layers **(Rafique et al., 2021)**. We adopted a modified Alex-Net to fit our issue domain. The final layer fully connected layer is swapped out with another layer fully connected layer to classify two classes (female or male) instead of 1000 classes. The parameters of the layers are given in **Table 2.**

- Convolution layer
  The convolution layer (CL) applies to original images using a unique convolution kernel feature detector. The image is scaled to 227×227 ×3 during training. To create new pixels via convolution, 96 numbers of 11×11×3 convolution kernels are applied in the first convolution layer. The 27×27×96 pixel layer produced from the first layer serves as the second layer's input data. Each pixel's edges are filled with two pixels to help with further processing. The computing formula for the second layer's convolution kernel is the same as that for the first layer, with a size of 5×5 and a stride of 1 pixel. A collection of 13× 13×128 pixel layers



produced by ...x-N... ...rders are occupied ... of ... pixel layers generated by the third layer serve as the fourth layer's input data ...ion of 13×13×192 pixel layers produced by the fourth layer serves as the fifth ... data. Each pixel layer's borders are filled with a single pixel to make subsequent processing easier. The convolution layer's KS, stride (S), and padding (P), the dimensions of the output and input images are set at the convolution layer **(Lin et al., 2020)**. The convolution layer's calculation using Eq. (4):

$$\boldsymbol{o} = \frac{(\mathrm{I}-\mathrm{KS}+2\mathrm{P})}{s} + 1 \qquad (4)$$

- Pooling layer
  Max pooling is primarily used to select the maximum value in the array. The max pooling does not impact the judgment when several pixels translate the image. The main benefit of Max pooling is noise reduction. The first, second, and fifth convolution layers are followed by the maximum pooling layers **(Lin et al., 2020)**.
- Fully connected layer
  Alex-Net, There are three completely interconnected levels. The first and second layers each contain 4096 convolution kernels, with a size of 6×6×256. That is due to the exact match between the size of the convolution kernel and the processing feature map. Only one-pixel value from the feature map's size and a one-to-one correlation are multiplied by each coefficient in the convolution kernel. As a result, are 4096 neurons in the pixel layer after convolution, which is 4096×1×1. This study has two classes; thus, two neurons in the third ultimately linked layer are tuned to one type each **(Lin et al., 2020)**. The original Alex-Net requires an input RGB image of dimensionality 227×227, in line with the increased dimensionality of images. In addition, the Alex-Net has an output layer with thousands of neurons corresponding to each type of Image-Net Internal object.
  As a result, the last three layers were modified as part of the debugging phase to address the gender classification problem from the facial image to (male/female). These layers are replaced with a full connection (fc), Softmax, and output layers. By implementing this fine-tuning phase, the network can infer the necessary bias and specificity of the data, which

enhances the ability of the network to make strong discrimination. In addition, to counteract overfitting, Two dropout layers (50% randomly) were added to the grid due to the limited size of the data set. **Fig. 9** shows the implementation details of the entire Alex-Net network layers after modifications.

Table 2. Details of Transfer Learning Alex-Net layers.

Layer	Number of Kernels	Kernel Size	Stride	Padding	Output Size
Input	-	-	-	-	227×227×3
Conv1	96	11×11×3	4	0	55×55×96
Relu1	-	-	-	-	55×55×96
Norm1	-	-	-	-	55×55×96
Max pool1	-	3×3	2	0	27×27×96
Conv2	256	5×5×48	1	2	27×27×256
Relu2	-	-	-	-	27×27×256
Norm2	-	-	-	-	27×27×256
Max pool2	-	3×3	2	0	13×13×256
Conv3	384	3×3×256	1	1	13×13×384
Relu3	-	-	-	-	13×13×384
Conv4	384	3×3×192	1	1	13×13×384
Relu4	-	-	-	-	13×13×384
Conv5	256	3×3×192	1	1	13×13×256
Relu5	-	-	-	-	13×13×256
Max pool5	-	3×3	2	0	6×6×256
Fc6	-	-	-	-	1×1×4096
Relu6	-	-	-	-	1×1×4096
Dropout6 (50%)	-	-	-	-	1×1×4096
Fc7	-	-	-	-	1×1×4096
Relu7	-	-	-	-	1×1×4096
Dropout7 (50%)	-	-	-	-	1×1×4096
fc8	-	-	-	-	1×1×2
Softmax	-	-	-	-	1×1×2
Class output	-	-	-	-	Cross entropyex with ‘male’ and 'female' other class

3. EVALUATION METRICS

A confusion matrix can be used to investigate how well a classification algorithm can do gender recognition from a face image. **Table 3** illustrates the confusion matrix. The confusion matrix's TP and TN values represent the proportion of accurate positive and negative classifications **(Goodfellow et al., 2016)**. At the same time, FP and FN stand for the balance of incorrectly classified negative and positive instances, respectively. Several commonly used metrics can be presented from the confusion matrix for evaluating the classifier performance with varying concentrations of evaluations**(Vujović, 2021; Al Jibory et al., 2022)**.

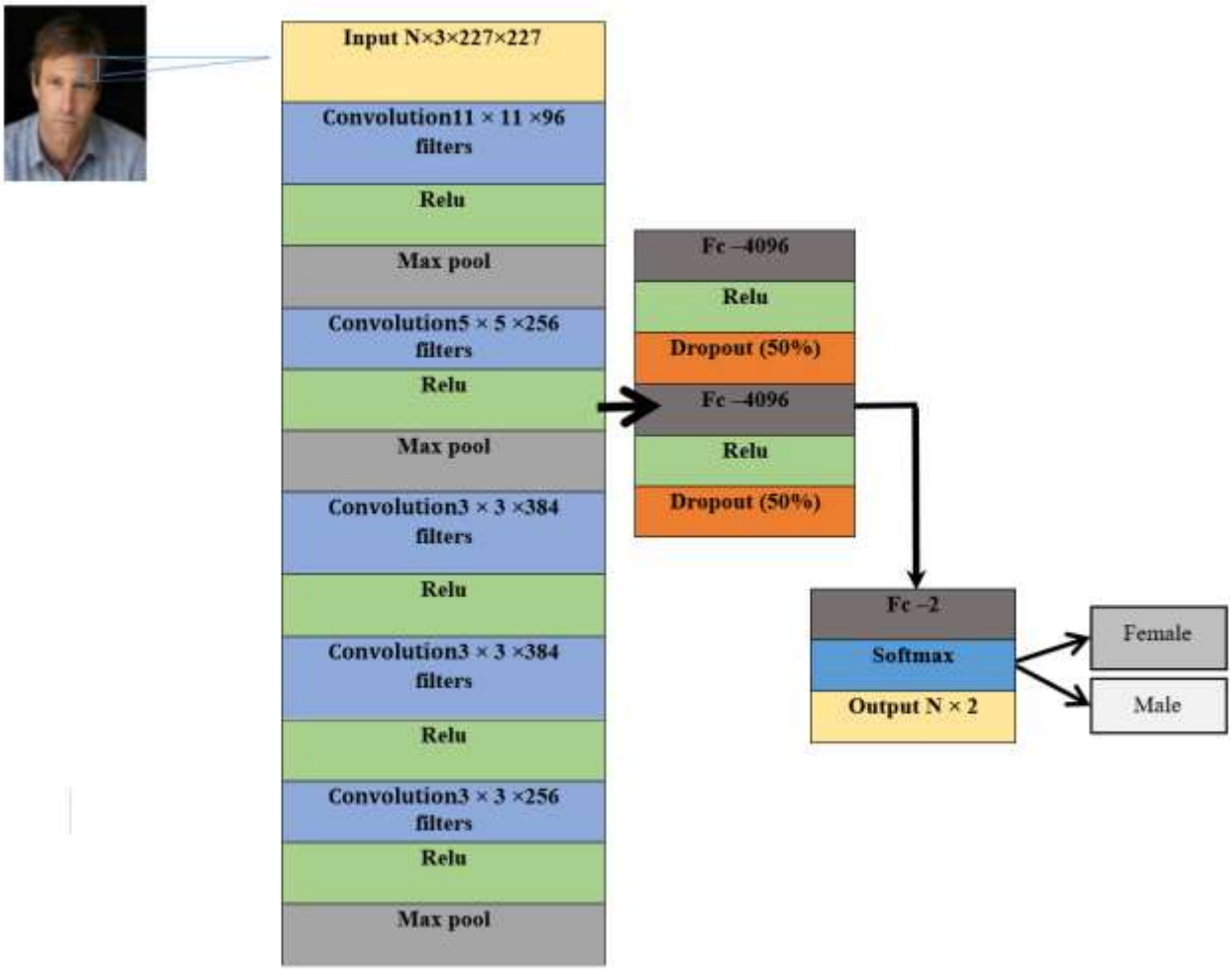


Figure 9. Conceptual of Transfer Learning Alex-Net model.

Table 3. Confusion Matrix

		True condition		
	Total population	condition Positive	Condition Negative	
Predicted condition	Predicted condition positive	True Positive TP	False Positive FP	PP
	Predicted condition negative	False-negative FN	True negative TN	PN
		RP	RN	

• **Accuracy:** This is seen in the ratio of real classifications to all sample tests used during training and testing. It is computed using Eq.(5) **(Vujović, 2021; Alfarhany and Abdullah, 2023):**

$$\text{Accurcy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (5)$$

The volume of accurate predictions that have been tallied is used to determine how well the gender categorization from the facial image technique performs.

• **Recall or sensitivity ratio:** the ratio of true positive cases correctly predicted positive **(Mateenet et al., 2020).** This rate is calculated by dividing the images (Male and female) correctly detected for a total of images (male and female) and false negative detected correctly, using Eq. (6)

$$\text{Recall(Sensitivity)TPR} = \frac{\text{TP}}{\text{TP+FN}} = \frac{\text{TP}}{\text{RP}} \quad (6)$$

• **Precision or confidence ratio**: Precision represents the proportion of predicted positive cases correctly classified as positive. It is calculated by dividing the number of correctly detected images by the sum of false-positive and correctly detected images **(Alkentar et al., 2021)**, using Eq. (7)

$$\text{Precision (Confidence)TPA} = \frac{\text{TP}}{\text{TP+FP}} = \frac{\text{TP}}{\text{PP}} \quad (7)$$

• **F1-score:** it is the average of the weighted values for precision and recall. As a result, this score considers both false positives and false negatives, using Eq.(8) **(Vujović, 2021; Ghadi and Salman, 2022)**:

$$\text{F1score} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall+Precision}} \quad (8)$$

4. RESULTS AND DISCUSSION

The proposed model was implemented using MATLAB (R2020a). The computer system specifications are as follows: Processor intel Corei7-10510U CPU @ 1.80GHz, Memory size: 8 GB, Type of Operating system 64-bit. Before entering the data into the Transfer Learning Alex-Net model, it was divided into main parts: training constitutes 80%, and testing includes 20%.

Another set was taken from the validation set, consisting of 16% of the total training sets. This division is commonly employed to ensure proper evaluation and optimization of the model's performance as given in **Table 4**. The training and validation sets were used during the model's training, while the test set was used for the final model performance. In the proposed model, three datasets were used for training separately using the same metrics in training face image datasets with the Transfer Learning Alex-Net model.

The training was done for the model for 20 epochs with a batch size of 20, and the learning rate is set to 0.0003. Three cases will be illustrated.

Case 1: Results of training without Preprocessing for three datasets, as shown in **Table 5.**

Table 4. Facial Image Datasets Splits.

Dataset	No. of Training Images	No .of Validation Images	No .of Testing Images	Total
LFW	768	192	240	1200
Faces94	256	64	80	400
Real data	256	64	80	400

Table 5. Results of training without Preprocessing for three datasets.

Dataset	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
LFW	100.00%	92.19%	0.0051	0.2481
Faces94	100.00%	95.00%	0.0515	0.0481
Real data	100.00%	95.31%	0.0088	0.0650

Case 2: Results of training with Preprocessing using Histogram equalization (HE) for three datasets, as shown in **Table 6.**

Table 6. Results of training with Preprocessing using (HE) for three datasets

Dataset	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
LFW	100.00%	91.15%	0.0037	0.2168
Faces94	100.00%	98.44%	0.0006	0.1413
Real data	100.00%	98.44%	0.0015	0.0766

Case 3: Results of training with Preprocessing using CLAHE for three datasets, as shown in **Table 7**.

Table7. Results of training with Preprocessing using (CLAHE) for three datasets

Dataset	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
LFW	100.00%	98.96%	0.0326	0.0259
Faces94	100.00%	100.00%	0.0031	0.0015
Real data	100.00%	100.00%	0.0005	0.0011

5. CONCLUSIONS

The proposed method is used to classify the Gender from the face image, whether (male or female), before inputting the face image in Transfer Learning Alex-Net. We apply the Viola-Jones algorithm for face images to detect faces and exclude the background. After the face detection in images, a pre-processing step is used, such as cropping, resizing, Images Convert from RGB to LAB, and Contrast enhancement using (CLAHE). Deep learning techniques are the most common for gender detection, aiming to learn essential features from raw data automatically. Alex-Net-based Deep Learning model was proposed for gender classification tasks. A comparison has been made using Alex-Net with the dataset before and after preprocessing, where Alex-Net was applied to three datasets. The Alex-Net model delivered the best accuracy after preprocessing in LFW, Faces94, and Real datasets, where 98.77 %, 100%, and 100%, respectively.

NOMENCLATURE

Symbol	Description	Symbol	Description
A*	A(green-red)	CNN	Convolutional Neural Networks
a *	The red-green color component of the LAB color space. It ranges from -128 to 127, where negative values indicate greenness and positive values indicate redness.	D-CNN	Deep Convolution neural networks
b *	The yellow-blue color component of the LAB color space. It also ranges from -128 to 127, where negative values indicate blueness and positive values indicate yellowness.	FC	Fully Connected Layer
B*	B(blue-yellow)	FN	False Negative
I	The size of the input image	FP	False Positive
KS	The dimensions of the convolutional kernel	HCI	Human-Computer Interaction
L*	The lightness component of the LAB color space. It ranges from 0 to 100, where 0 represents black and 100 represents white.	HE	Histogram Equalization
o	The output size after applying the convolution operation	LBPH	Local Binary Pattern Histogram
P	The number of pixels added around the input to control the spatial dimensions and preserve information during convolution.	LFW	Labeled Faces in the Wild
s	The step size or the number of pixels the kernel moves during the convolution operation	Relu	Rectified Linear Unit Function
X, Y, and Z	The amounts of the three primary colors (red, green, and blue) in a given color, while Xn, Yn, and Zn represent the reference white point.	RGB	Red, Green, Blue Color Space
Abbreviations		ROI	Region Of Interest
Adaboost	Adaptive boosting	Tn	True Negative
CL	Clip limit	Tp	True Positive
CLAHE	Contrast limited adaptive histogram equalization		

REFERENCES

Alamri, H., Alshanbari, E., Alotaibi, S., and AlGhamdi, M., 2022. Face recognition and gender detection using SIFT feature extraction, LBPH, and SVM. *Engineering, Technology & Applied Science Research*, *12*(2), pp. 8296-8299. Doi:10.48084/etasr.4735

Alkentar, S.M., Alsahwa, B., Assalem, A., and Karakolla, D., 2021. Practical comparation of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection. *Journal of Engineering*, *27*(8), pp. 19-31. Doi:10.31026/j.eng.2021.08.02

Alfarhany, A.A.R., and Abdullah, N.A., 2023. Iraqi Sentiment and Emotion Analysis Using Deep Learning. *Journal of Engineering*, *29*(09), pp. 150-165. Doi: 10.31026/j.eng.2023.09.11

Al Jibory, F.K., Mohammed, O.A., and Al Tamimi, M.S.H., 2022. Age estimation utilizing deep learning Convolutional Neural Network. *International Journal on Technical and Physical Problems of Engineering*, 14(4), pp. 219–224.

Anusri, U., Dhatchayani, G., Angelinal, Y. P., and Kamalraj, S., 2021. An early prediction of Parkinson's disease using facial emotional recognition. In *Journal of Physics: Conference Series* (Vol. 1937, No. 1, p. 012058). IOP Publishing. Doi:10.1088/1742-6596/1937/1/012058

Bajrami, X., Gashi, B., and Murturi, I., 2018. Face recognition performance using linear discriminant analysis and deep neural networks. *International Journal of Applied Pattern Recognition*, *5*(3), pp. 240-250.Doi: 10.1504/IJAPR.2018.094818

Berbar, M. A., 2022. Faces recognition and facial gender classification using convolutional neural network. *Menoufia Journal of Electronic Engineering Research*, 31(2), pp. 1-10. Doi: 10.1109/AIMV53313.2021.9670898

Chaudhari, M. N., Deshmukh, M., Ramrakhiani, G., and Parvatikar, R., 2018. Face detection using viola jones algorithm and neural networks. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-6. Doi:10.1109/ICCUBEA.2018.8697768

Chen, J., Liu, S., and Chen, Z., 2017. Gender classification in live videos. *IEEE International Conference on Image Processing (ICIP)* pp. 1602-1606. Doi: 10.1109/ICIP.2017.8296552

Ciocca, G., Cusano, C., Gasparini, F., and Schettini, R., 2007. Self-adaptive image cropping for small displays. Transactions on Consumer Electronics, 53(4), pp. 1622-1627. Doi:10.1109/TCE.2007.4429261

Dawood, F.A., and Abood, Z.M., 2018. The importance of contrast enhancement in medical images analysis and diagnosis. *International Journal of Engineering Research & Technology (IJERT)*, V7(12), pp. 21–24. Doi:10.17577/ijertv7is120006

Dhomne, A., Kumar, R., and Bhan, V., 2018. Gender recognition through face using deep learning. *Procedia Computer Science*, 132, pp. 2–10. Doi:10.1016/j.procs.2018.05.053

Jia, S., Lansdall-Welfare, T., and Cristianini, N., 2016. Gender classification by deep learning on millions of weakly labelled images. *International Conference on Data Mining Workshops (ICDMW)*, pp. 462-467. Doi:10.1109/ICDMW.2016.0072

Fadhil, S.S., and Dawood, F.A.A., 2021. Automatic pectoral muscles detection and removal in mammogram images. *Iraqi Journal of Science*, 62(2), pp. 676–688. Doi:10.24996/ijs.2021.62.2.31

Galla, D.K.K., Mukamalla, B.R., and Chegireddy, R.P.R., 2020. Support vector machine based feature extraction for gender recognition from objects using lasso classifier. *Journal of Big Data*, 7(1). Doi:10.1186/s40537-020-00371-0

Ghadi, N.M., and Salman, N.H., 2022. Deep learning-based segmentation and classification techniques for brain tumor mri: a review. *Journal of Engineering*, *28*(12), pp. 93–112. Doi:10.31026/j.eng.2022.12.07

Goodfellow, I., Bengio, Y., and Courville, A., 2016. Deep learning. MIT Press. Ghadi, N.M., and Salman, N.H., 2022. Deep learning-based segmentation and classification techniques for brain tumor MRI: A review. *Journal of Engineering*, 28(12), pp. 93-112. Doi:10.31026/j.eng.2022.12.07

Hassan, B.A., and Dawood, F.A.A. 2023. Facial image detection based on the Viola-Jones algorithm for gender recognition. *International Journal of Nonlinear Analysis and Applications*, *14*(1), pp. 1593-1599.

Heydarzadeh, Y., Haghighat, A. T., and Fazeli, N., 2010. Utilizing skin mask and face organs detection for improving the Viola face detection method. In *2010 Fourth UKSim European Symposium on Computer Modeling and Simulation*. pp. 174-178. Doi:10.1109/EMS.2010.38

Joodi, M.A., Saleh, M.H., and Khadhim, D.J., 2023. Proposed face detection classification model based on Amazon Web Services Cloud (AWS). *Journal of Engineering*, 29(4), pp. 176–206. Doi:10.31026/j.eng.2023.04.12

Kang, S., Choi, B., and Jo, D.,2016. Faces detection method based on skin color modeling. *Journal of Systems Architecture*, *64*, pp. 100-109. Doi.org/10.1016/j.sysarc.2015.11.009

Kavi Niranjana, K., Professor, A., and Kalpana Devi, M., 2015. RGB to Lab transformation using image segmentation. *International Journal of Advance Research in*, 3(11), pp. 8–16. Doi:10.33395/sinkron.v3i2.10102

Lin, C.J., Li, Y.C., and Lin, H.Y., 2020. Using convolutional neural networks based on a Taguchi method for face gender recognition. *Electronics (Switzerland)*, 9(8), pp. 1–15. Doi: 10.3390/electronics9081227

Mateen, M., Wen, J., Nasrullah, N., Sun, S., and Hayat, S., 2020. Exudate detection for diabetic retinopathy using pretrained convolutional neural networks. *Complexity*, *2020*, pp.1-11.

Musa, P., Rafi, F. Al and Lamsani, M., 2018. A review: Contrast-limited adaptive histogram equalization (CLAHE) methods to help the application of face recognition. *Proceedings of the 3rd International Conference on Informatics and Computing, ICIC 2018*, (November 2020), pp. 1–6. Doi:10.1109/IAC.2018.8780492.

Peng, Z., Wu, J., and Fan, G.,2018. A rapid face detection method based on skin color model and local binary gradient feature. In *2018 5th International Conference on Systems and Informatics (ICSAI)* pp. 922-927. Doi:10.1109/ICSAI.2018.8599306

Rafique, R., Nawaz, M., Kibriya, H., and Masood, M., 2021. DeepFake detection using error level analysis and deep learning. In *2021 4th International Conference on Computing & Information Sciences (ICCIS)* pp. 1-4. Doi:10.1109/ICCIS54243.2021.9676375

Rathore, V.S., Kumar, M.S., and Verma, A., 2012.Colour based image segmentation using L* a* b* colour space based on genetic algorithm. *International Journal of Emerging Technology and Advanced Engineering*, 2(6), pp. 156–162.

Suma, S. L., and Raga, S., 2018. Real time face recognition of human faces by using LBPH and Viola Jones algorithm. *International Journal of Scientific Research in Computer Science and Engineering*, *6*(5), pp. 6-10. Doi:10.26438/ijsrcse/v6i5.610

Thepade, S. D., and Abin, D., 2018. Face gender recognition using multi-layer perceptron with OTSU segmentation. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)* pp. 1-5. Doi: 10.1109/ICCUBEA.2018.8697480

VenkateswarLal, P., Nitta, G.R., and Prasad, A., 2019. Ensemble of texture and shape descriptors using support vector machine classification for face recognition. *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-8. Doi:10.1007/s12652-019-01192-7

Vujović, Ž., 2021. Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6), pp. 599–606. Doi:10.14569/IJACSA.2021.0120670

Wu, B., Zhu, W., Shi, F., Zhu, S., and Chen, X., 2017. Automatic detection of microaneurysms in retinal fundus images. Computerized Medical Imaging and Graphics, 55, pp. 106–112. Doi:10.1016/j.compmedimag.2016.08.001

Zaman, F.H.K., 2020. Gender classification using custom convolutional neural networks architecture. *International Journal of Electrical and Computer Engineering*, 10(6), pp. 5758–5771. Doi:10.11591/ijece.v10i6.pp5758-5771

Zheng, J., Ramirez, G. A., and Fuentes, O., 2010. Face detection in low-resolution color images. In Image Analysis and Recognition: 7th International Conference, ICIAR. pp. 454-463. Doi:10.1007/978-3-642-13772-3_46