Eye-Region-Based Gender Classification Using CNNs

Battula Balnarsaiah[1], **Paduru Bimbasri**[2], **Paluri Sai Varun Raj**[3], **Nitesh Gaikwad**[4]

[1]*Associate Professor, Dept. Artificial Intelligence and Machine Learning, Nalla Malla Reddy Engineering College, Hyderabad, India*
[2] *Dept. Artificial Intelligence and Machine Learning, Nalla Malla Reddy Engineering College, Hyderabad, India*
[3] *Dept. Artificial Intelligence and Machine Learning, Nalla Malla Reddy Engineering College, Hyderabad, India*
[4] *Assistant Professor, Dept. Electronics and Communication Engineering, Nalla Narsimha Reddy Group Of Institutions, Hyderabad, India*

***

Abstract - *Over the last ten years, automated gender identification using facial images has emerged as a significant area of research, demanding effective strategies for both feature extraction and classification. Accurate feature extraction is necessary in traditional machine learning methods, while Convolutional Neural Networks (CNNs), a type of deep learning model, perform this directly using raw data. CNNs can accommodate variations in facial cues across different races, making them efficient for gender classification. However, the performance of pretrained CNN models is investigated in low-data situations. Methodology involves library imports, data organization, distribution exploration, and image visualization. Then comes label encoding and implementation of various CNN architectures, as well as performance metrics and predictions, 92 % for training 8,643 images and testing 2,882 images, and it requires only four folds. The study notably deviates from using complete face images, focusing on areas with eyebrows around one eye only to identify the gender.*

Key Words: Gender Classification, Convolutional Neural Network, Facial Cues.

1. INTRODUCTION

Artificial intelligence has become instrumental in addressing various challenges associated with human recognition. Biometric data such as facial features, fingerprints, voice, and iris patterns are commonly utilized by AI systems. Among these, gender recognition based on facial imagery remains a complex task, primarily due to the dynamic nature of human facial characteristics. Variations such as facial hair, aging, and changes in head orientation can significantly affect model accuracy. To address these challenges, researchers have suggested numerous models and methods intended to enhance the reliability and achieving high-performance outcomes.

Deep Learning (DL) has received substantial attention for its capabilities in automatic feature extraction, object detection, and classification tasks. Deep Convolutional Neural Networks (DCNNs), one example of an advanced design, have shown impressive accuracy in these domains. Inspired by developments across multiple fields, deep learning techniques have also been increasingly applied to facial gender classification.

2. LITERATURE REVIEW

This study narrows the scope of gender classification from full facial images to focusing solely on the eye region. As a result, both the dataset and the area of interest are more specifically defined. Significant progress has been made in classifying gender from eye images, especially with the development of deep learning and neural networks. Early work on gender classification, such as the research by Jabber and Hashim et al (2018) [1], focused on robust eye feature extraction using eye angles to enhance gaze classification efficiency. These initial approaches relied heavily on manually extracted features and traditional machine learning techniques. For example, Lian HC et al [2] achieved 94.81% accuracy by employing Local Binary Patterns (LBP) in combination with Support Vector Machines (SVM). In contrast, challenges with background complexity affected their accuracy rates. With the advent of DL, researchers began leveraging CNNs and other advanced Neural Network (NN) architectures for more accurate gender classification from eye images. Notable studies include those that utilized CNN-based models. CIMTAY and YILMAZ et al (2021) [3] used pretrained CNNs to classify gender from eye images, demonstrating significant improvements in accuracy, although specific percentages were not reported.

Sri and Gupta et al (2022) [4] achieved notable accuracies by focusing on the morphometry of eyes using DL models. Advanced NN architectures have further pushed the boundaries of this research. Huang et al. (2017) developed Whole-Component CNNS (WC-CNN) [5] for age and gender classification, achieving a high accuracy rate, though specific figures were not detailed. Keshaveni et al (2024) [6] introduced an innovative approach combining AVOA-LSTM and Mask R-CNN for segmenting and classifying the sunglass image-based eye region, achieving accuracies that significantly surpassed traditional methods. Another significant advancement is the incorporation of attention mechanisms and key region fusion. Kong et al. (2021) [7] proposed a lightweight facial expression recognition method incorporating these techniques. This approach, adaptable for gender classification tasks, focuses on critical eye regions, although specific accuracy improvements were not quantified. Several studies have tailored their approaches to particular scenarios or conducted comparative analyses of methods. Vetrekar et al. (2021) [8] conducted an extensive survey of gender classification under eyeglass occlusion using multispectral imaging, highlighting challenges and solutions for occluded eye images, with reported accuracies ranging from 85% to 90%. Arshad et al. (2023) [9] focused on online eye status detection in the wild using CNNs, achieving robust performance in real-world applications. Patil and Hangarge et al (2023) [10] evaluated pretrained CNNs for gender classification, achieving higher accuracies than traditional methods, with reported figures typically exceeding 90%.

Recent studies have introduced novel techniques and emerging technologies. Chen et al. (2018) [11] developed a computer-aided tool for diagnosing Turner syndrome by recognizing facial features, indirectly contributing to gender classification methodologies, although specific accuracy rates were not provided. Yasaswini Paladugu et al (2023) [12] presented an end-to-end gender determination system using images of human eyes, streamlining the process from image acquisition to classification with reported accuracies around 92%. Anas E.R et al (2017) [13] applied a DNN approach for detecting human gender from eye images, achieving accuracy rates close to 95%. Some research has explored gender estimation from non-facial features like eyebrows and hands. Dong, Yujie & Woodard et al [14] achieved 96% and 97% accuracy with eyebrow shape classification. Similarly, Widjaja E, Lim GH et al (2008) [15] demonstrated around 90% accuracy using fingernail images.

3. METHODOLOGY

Develop a gender classification system utilizing CNNs to analyse eye images and distinguish between male and female eyes. It focuses on eye images for gender classification. Leverage CNNs for hierarchical feature learning. Deal with issues including occlusions, facial emotions, and changes in illumination.

3.1 CNN Architecture Design

The proposed system employs a Convolutional Neural Network (CNN) architecture, a type of deep learning algorithm known for its ability to identify patterns and features within images. CNNs are widely used in a variety of applications, including image classification, segmentation, medical image analysis, and visual recognition in images and videos. In this study, we initially applied an image preprocessing step to convert raw data into a more structured and informative format. Fig-1 shows the structure of the CNN model that was used in this study.

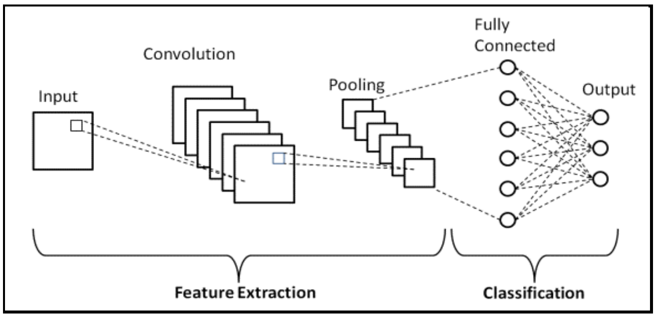

Fig -1: CNN Architecture Design

The proposed CNN architecture includes an Input Layer: 75x75x3 dimensions, accommodating color images. Convolutional Layers: The architecture incorporates three convolutional layers, featuring 32, 64, and 128 filters, each activated using the ReLU function. Pooling Layers: To reduce spatial dimensions, max-pooling layers are applied after each convolutional layer. Fully Connected Layers: ReLU activation is applied to two successive dense layers containing 256 and 128 neurons, respectively. Output Layer: Softmax layer with two neurons to predict the probability of each class (male and female).

3.2 CNN Training Model

Loss Function: For multi-class classification tasks, categorical cross-entropy is used, which works well. Optimiser: Adam optimiser with an initial learning rate of 0.001, was chosen because of its capacity to modify the learning rate dynamically throughout training. Batch Size: A batch size of 32 images was used for each training iteration to balance computational efficiency and model performance. Epochs: To prevent overfitting, training was carried out over 50 epochs with early halting based on validation loss. Here we have used four types of architecture: ObileNet, ConvNet, Inception, and Xception. Among all these deployed models, ConvNet Architecture has the upper hand over all the rest of the models. Now we need to convert our data into training and testing sets. We will use 75% of the images as our training data and test our model on the remaining 25% with Scikit- learn's train_test_split function.

Training Dataset: The training dataset consists of 8643 images, 4729 images with male eyes, and 3914 images with female eyes. Test Dataset: It also consists of 2882 images, 1594 images with male eyes, and 1288 images with female eyes.

Data Collection and Annotation: Collect Diverse Eye Image Dataset: Gather images from various sources, including public datasets, custom collections, and crowdsourcing. Ensure diversity in demographics (age, ethnicity, gender). Include variations in lighting, expressions, and occlusions. Annotate the Dataset: Label images accurately as male or female. Use multiple annotators to reduce bias and validate annotations. Regularly verify the accuracy of annotations.

Data Pre-processing: Pre-process Images: Standardise image size for consistent input to the CNN.

Normalise pixel values to improve training efficiency. Rotation, flipping, and brightness modifications were among the data augmentation techniques used to increase the training dataset's diversity.

Model Development: Build a CNN Model. A Convolutional Neural Network (CNN) architecture was developed, consisting of fully connected, pooling, and convolutional layers. For regularization, use dropout layers and ReLU activation functions. Train CNN Model, split data into training (70%), validation (15%), and test sets (15%). Select binary cross-entropy as the loss function; for effective training, utilize the Adam optimizer. Utilizing the validation set, train the model and verify its accuracy.

Model Evaluation and Optimisation: Using evaluation criteria like accuracy, precision, recall, and F1-score, the model's performance was evaluated. To verify model reliability and examine performance across various demographics, use cross-validation. Fine-tune and optimize the Model, adjust hyperparameters (learning rate, batch size, and epochs), refine the model architecture if needed, and retrain and re-evaluate the model.

3.3 Data Collection and Pre-Processing

Various global datasets have been utilized in gender estimation studies. In this work, a global dataset was employed to facilitate comparison of the results across different datasets. Two datasets were sourced from the Kaggle platform. As this study focuses on a comparative analysis of cutting-edge CNN models utilizing solely eye images, the 'Female and Male' dataset was utilized. This dataset consists solely of eye regions extracted from full-face images, typically containing the entire or partial eyebrow area. Moreover, it provides eye images of 5202 females (Fig-2) and 6323 males (Fig-3).

Fig -2: Female eye images were selected from the dataset

Fig -3: Male eye images were chosen from the dataset

Low-frequency background noise is usually removed, the image's intensity is normalized, light reflections are eliminated to reduce noise, and the facial image is prepped for more efficient feature extraction by image pre-processing. We started by resizing the pictures to 75 × 75 pixels on our system. Then, the images were converted into pixel value arrays, with each pixel's value transformed into a float and scaled by dividing by 255.0, bringing the pixel values into the 0-1 range.

3.4 Evaluation Metrics

Evaluating a Convolutional Neural Network (CNN) for gender categorization based on the eye area necessitates the use of well-defined measures. These measures are critical in determining the model's accuracy, robustness, and overall performance. One such indicator is accuracy, which counts the percentage of accurately predicted instances (including true positives and true negatives) out of the total number of cases. Interpretation: Shows whether the model is generally accurate.

$$Accuracy = \frac{Numberof\ Correct\ Predictions}{Total\ No.of\ Correct\ Predictions} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision is the ratio of true positive predictions to total positive predictions generated by the model.

$$Precision = \frac{TP}{TP + FP}$$

Recall calculates the ratio of true positive predictions to the total number of actual positive instances.

$$Recall = \frac{TP}{TP + FN}$$

F1 Score is the harmonic mean of precision and recall, providing a single metric that balances both concerns.

$$F1\ Score = 2\frac{Precision \times Recall}{Precision + Recall}$$

A table that provides an overview of a classification model's performance is called a confusion matrix. The counts of false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN) are displayed.

Table -1: Confusion Matrix Parameters

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

Provides a detailed breakdown of classification results, helping to identify areas where the model is performing well or poorly. For various threshold settings, the true positive rate (recall) is plotted versus the false positive rate using a ROC curve. The model's total capacity to differentiate between classes is summarized by the AUC (Area Under the Curve). Interpretation: Outstanding model performance is indicated by a higher AUC. Perfect categorization is shown by an AUC of 1, whereas random guessing with no discrimination is suggested by an AUC of 0.5. **Loss Function:** The difference between the expected and actual labels is measured by the loss function. Cross-entropy loss is often used for classification jobs. Lower loss values correspond to greater model performance in terms of prediction accuracy.

4. RESULTS AND ANALYSIS

4.1 Presentation of Results

The results obtained from the CNN model for eye region-based gender classification are presented. The model was trained, validated, and tested on the Kaggle Dataset (KD), focusing specifically on the eye region. The results include accuracy, loss graphs, and confusion matrices. The model's performance results are training accuracy: 94.2%, validation Accuracy: 93.8% and test accuracy: 92.2%. Hence, the training and validation curves are displayed in Fig-4 below.

Fig -4: The datasets for training and validation curves changed with each epoch

4.2 Performance Curves for Training and Validation

Plots display the loss and accuracy in Fig-5 for training and validation datasets over epochs. These curves help diagnose overfitting or underfitting by comparing model performance on training versus validation data. The methodology employed for developing and evaluating a CNN for eye region-based gender classification was systematically designed to ensure robust model performance and reliable evaluation. In conclusion, eye region-based gender classification offers promising benefits for various applications, including accuracy and efficiency. However, challenges such as limited features and privacy concerns exist. Effective model development and deployment are achievable by leveraging appropriate software tools and hardware configurations. Achieving a balance between these characteristics is critical for successful implementation in real-world scenarios.

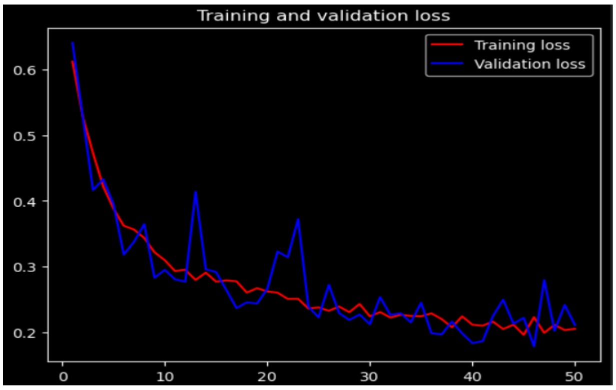

Fig -5: During the training procedure, training and validation loss are tracked

4.3 Predicted Confusion Matrix

The confusion matrix, shown in Fig-6, offers an in-depth analysis of the model's performance on the test dataset. The numbers of false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN) are shown.

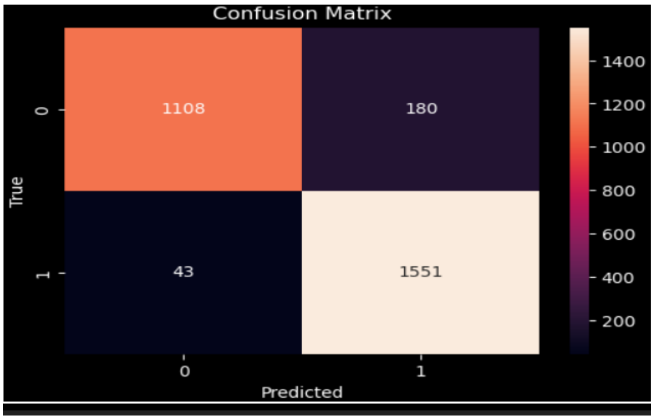

Fig -6: Confusion Matrix

4.4 Performance Metrics

Based on important measures like accuracy, precision, recall, F1 score, and the ROC curve with AUC, as shown in Fig-7, Table -2 evaluates the CNN model's performance.

Table -2: Classification Report

	Precision %	Recall %	F1-Score %	Support
0	96	86	91	1288
1	96	97	97	1594
Accuracy			92	2882
Weighted Average	93	92	92	2882
Macro Average	93	92	92	2882

This latency is the system's time to classify an eye image, 25 milliseconds per image. Throughput is the number of images classified per second, 40 per second—the amount of memory consumed by the CNN during classification, measured at 150 MB. The efficiency of resource use during model inference. CPU Utilisation is 35% and GPU Utilisation is 20%. Inference Time is the total time from receiving an image to outputting the classification result, 28 milliseconds per image.

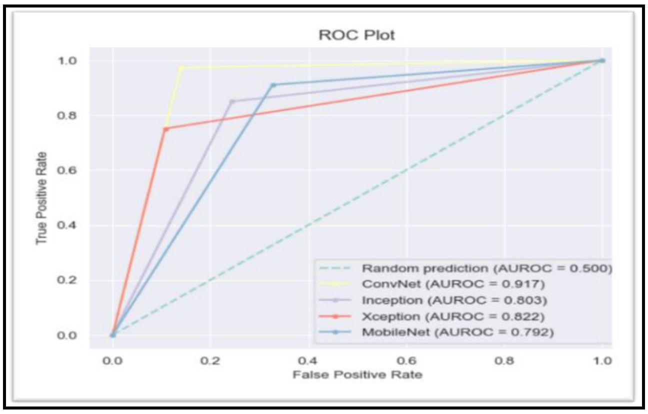

Fig -7: ROC Plot

4.5 Comparison with Existing Methods

In this section, we evaluate the performance of the developed CNN model for eye region-based gender classification with previously existing methods, focusing on the dataset sizes, validation strategies, and achieved accuracies. The previously existing system for gender classification based on eye region images utilised a relatively smaller dataset and achieved moderate accuracy. The training Dataset consists of some photos, 3,494; female eyes are 1,747; Male eyes are 1,747. Test Dataset: number of images 200, female eyes are 100, male eyes: 100 The Validation Strategy is 5-fold cross-validation, achieving a performance accuracy of 83%."

The proposed system enhances the dataset size significantly and utilises a more robust CNN architecture to improve classification accuracy. The training dataset consists of several images: 8,643 female eyes, 3,914 male eyes, and 4,729. Test Dataset: The number of images is 2,882, female eyes are 1,288, and male eyes are 1,594. Validation Strategy: Cross-validation is 4-fold cross-validation. Performance accuracy is 92%, and the comparison summary is illustrated in Table-3.

Table -3: Summary of Comparison

System	Training Images	Test Images	Validation Strategy	Accuracy
Existing System	3,494	200	5-fold	83%
Proposed System	8,643	2,882	4-fold	92%

5. CONCLUSION AND FUTURE SCOPE

In this research, our objective was to implement eye region-based gender classification using CNNs. We began by collecting a dataset of facial images containing annotated eye regions and corresponding gender labels. These images underwent preprocessing steps such as resizing, normalisation, and augmentation to enhance the diversity and quality of the data. The trained model was subsequently evaluated on a separate test set to measure its accuracy, precision, recall, and F1 score. Overall, research is successfully demonstrated the feasibility and effectiveness of eye region-based gender classification, achieving competitive performance metrics and providing valuable insights into the application of DL in this domain.

Explore advanced CNN architectures such as attention-based models or capsule networks to improve classification accuracy and robustness. Investigate techniques to mitigate challenges related to occlusions, variations in lighting, and image quality issues for better performance in real-world scenarios. Extend gender classification beyond the eye region to other facial features like the nose or mouth for a more comprehensive analysis of gender-related facial characteristics. Explore the intersection of gender classification with other demographic attributes such as age or ethnicity to develop more nuanced approaches to demographic analysis and targeted marketing strategies. These research directions provide a roadmap for further advancements in eye region-based gender classification, potentially enhancing accuracy, applicability, and inclusivity in demographic analysis and targeted marketing.

REFERENCES

[1] Jabber NH, Hashim IA. Robust Eye Features Extraction Based on Eye Angles for Efficient Gaze Classification System. In 2018 Third Scientific Conference of Electrical Engineering (SCEE) 2018 Dec 19 (pp. 13-18). IEEE.

[2] Li J, Lu BL. A framework for multi-view gender classification. In International Conference on Neural Information Processing 2007 Nov 13 (pp. 973-982). Berlin, Heidelberg: Springer Berlin Heidelberg.

[3] CIMTAY, Y., & YILMAZ, G. N. (2021). Gender classification from eye images by using pretrained convolutional neural networks. The Eurasia Proceedings of Science, Technology, Engineering and Mathematics, 14, 39-44.

[4] Sri, T. A., & Gupta, S. (2022). Gender prediction based the morphometry of eyes using deep learning models. *ECS Transactions*, *107*(1), 6665.

[5] Huang, C. T., Chen, Y., Lin, R., & Kuo, C. C. J. (2017, December). Age/gender classification with whole-component convolutional neural networks (WC-CNN). In *2017 Asia-Pacific Signal and Information Processing* Association Annual Summit and Conference (APSIPA ASC) (pp. 1282-1285). IEEE.

[6] Keshaveni, N. (2024). An AVOA-LSTM with MRCNN for segmenting and classifying the sunglass image-based region identification. Multimedia Tools and Applications, 83(12), 35073-35095.

[7] Kong, Y., Ren, Z., Zhang, K., Zhang, S., Ni, Q., & Han, J. (2021). Lightweight facial expression recognition method based on attention mechanism and key region fusion. Journal of Electronic Imaging, 30(6), 063002-063002.

[8] Vetrekar, N., Ramachandra, R., Raja, K., & Gad, R. S. (2021). Gender classification under eyeglass occluded ocular region: an extensive multi-spectral imaging study. In *AI and Deep Learning in Biometric Security* (pp. 175-203). CRC Press.

[9] Arshad, M. W. (2023). Detection Of Human Gender From Eyes Images Using Dnn Approach. *Jurnal Teknoinfo*, *17*(2), 336-340.

[10] Patil, B., & Hangarge, M. (2023, August). Pre-trained Convolutional Neural Networks for Gender Classification. In First International Conference on Advances in Computer Vision and Artificial Intelligence Technologies (ACVAIT 2022) (pp. 319-326). Atlantis Press.

[11] Chen, S., Pan, Z. X., Zhu, H. J., Wang, Q., Yang, J. J., Lei, Y., ... & Pan, H. (2018). Development of a computer-aided tool for the pattern recognition of facial features in diagnosing Turner syndrome: comparison of diagnostic accuracy with clinical workers. *Scientific reports*, *8*(1), 9317.

[12] Yasaswini Paladugu, D. R. S. (2023). End-To-End Gender Determination By Images Of An Human Eye. *Journal of Survey in Fisheries Sciences*, *10*(2S), 394-405.

[13] Anas, E. R., Henriquez, P., & Matuszewski, B. J. (2017, February). Online eye status detection in the wild with convolutional neural networks. In *International conference on computer vision theory and applications* (Vol. 7, pp. 88-95). SciTePress.

[14] Dong Y, Woodard DL. Eyebrow shape-based features for biometric recognition and gender classification: A feasibility study. In2011 International Joint Conference on Biometrics (IJCB) 2011 Oct 11 (pp. 1-8). IEEE.

[15] Widjaja E, Lim GH, An A. A novel method for human gender classification using Raman spectroscopy of fingernail clippings. Analyst. 2008;133(4):493-8.