Electrical Engineering Technical Journal, ISSN: 3007-9772, Vol. 1, No. 1, 2024, Pages 13-19
DOI: https://doi.org/10.51173/eetj.v1i1.5

Electrical Engineering Technical Journal
Journal homepage: https://eetj.mtu.edu.iq

RESEARCH ARTICLE

Computer Vision for Automated Facial Characteristics Detection

Yasser Emad Salman [1*], Ammar Yahya Daeef [2], Asanka G. Perera [3], and Ali Al-Naji [1,4]

[1] Electrical Engineering Technical College, Middle Technical University, Baghdad 10022, Iraq
[2] Technical Institute for Administration, Middle Technical University, Baghdad 10074, Iraq
[3] School of Engineering, University of Southern Queensland, Springfield 4300 QLD, Australia
[4] School of Engineering, University of South Australia, Mawson Lakes, SA 5095, Australia

[*] Corresponding Author Email: bdc0050@mtu.edu.iq

Article Info.

Article history:

Received
02 February 2024

Accepted
20 April 2024

Publishing
28 June 2024

Abstract

Diagnosing diseases in their early stages allows people to see a specialist doctor before the disease reaches advanced stages and avoid any future complications. The use of a real-time imaging system to reliably give various information about the patient's condition, including gender classification, pterygium, and Bell's paralysis, contributes to reducing the duration of diagnosis and human errors. This study focuses on the use of three artificial intelligence algorithms based on deep learning, namely Visual Geometry Group (VGG16), Vision Transformer (ViT), and Xception, and evaluates their performance in detecting gender, pterygium, and Bell's paralysis. VIT has the highest overall performance results from the rest of the algorithms.

 Publisher: Middle Technical University

Keywords: Facial characteristics; Gender recognition; Bell's paralysis; Pterygium; Deep Learning; Xception; ViT; VGG16.

1. Introduction

Automatically categorizing facial characteristics and illness (BP, Pterygium, Gender) is one of the last medical approaches based on AI technology [1]. The advancement of an automatic facial character detection system is essential to strengthening patient care and diagnosing diseases valuably by analyzing a person's facial characteristics and recognizing these features [2]. A large number of infants were diagnosed with eye disorders numbered 4,223 infants, resulting in a rate of 20,242 / 100,000 births per year, or 1 in 4.9 live births [3]. BP is estimated at a rate of 15-23 per 100,000 people [4]. This project faces difficulties in reaching high accuracy due to the enormous variation in personal characteristics, facial data analysis, and changes in lighting [5, 6]. rising detection speed can enhance facial characteristics and disease detection systems [7]. The accuracy of the system can be improved by increasing the numbers and resolution of images and separating them into training and test categories [8].

In the last few years, many researchers have used facial characteristic detection systems to determine certain characteristics. For the GR imaging system, Kumar et al. (2019) used SVM and SIFT for automatic GR and tested them with three datasets. The first dataset FEI contained images of Brazilian individuals' faces and was divided into two groups: smiling and non-smiling. The second database was live images, and the third was the SCIEN database, containing 400 images divided into two halves for males and females. The accuracy of the results was 98%, 94%, and 91%, respectively. Even with this accuracy, the techniques could have performed better due to the time it takes to crop images and the difficulty of finding facial landmarks and extracting features [9]. Another study by Greco et al. (2020) used Viola-Jones and MobileNet v2 for a digital signage system for real-time GR. The system achieved 94.99% GR accuracy on a standard dataset and 92.70% in a real fair scenario [10]. In the following year, a study by Tursunov et al. (2021) used CNN and MAM to develop a GR and AR system. Datasets used were Common Voice and Korean speech recognition. Accuracies achieved were 96% and 97%, respectively [11]. A recent study by Feng (2023) used DCGAN and CNN to improve GR accuracy on a small CelebA dataset. The highest accuracy rate achieved in this study was 82.40%. However, the study used limited data and needed to provide information about the devices used [12]. Another study by Kumar et al. (2023) improved the GR system using DTL classification, KNN, NNs, SVM, Random Forest, and Boosting algorithms using the GR dataset. The boosting algorithm gave the highest accuracy of 98.4% [13].

To ensure eye safety, a number of researchers have created an imaging system to detect pterygium, including Zaki et al. (2018) used HSV-Sigmoid to differentiate the pterygium tissue and the corneal region for automatic detection of eye pterygium. UBIRIS and MILES databases were used for normal eye testing, and Australian Pterygium and Brazil Pterygium databases were used for pterygium eye. 95.6%, 91.2%, 88.7%, and 88.3% values were achieved for area under the curve, accuracy, sensitivity, and specificity, respectively [14]. A study by Zulkifley et al. used pterygium-Net, which is a full CNN for automatic pterygium detection; Data augmentation and Transfer learning techniques are also employed to improve the accuracy of pterygium detection. This study used a different public dataset to pre-train the CNN weights and biases. Values of 98.3%, 95%, and 81.1% were achieved for specificity, sensitivity, and accuracy, respectively. Additionally, a 0.053 failure rate was achieved [15]. A study by Ahmad et al. (2020) used HMC to identify external eye diseases. The dataset used was images of external eye diseases divided into four classes, which divided into several subclasses equal to the number of diseases in that class. The overall accuracy achieved was 75.71% [16].

For nerve system health, Song et al. (2018) used a single CNN for accurate and fast classification of BP. They used 1049 clinical training images divided into seven categories and reported 97.5% accuracy [17]. In 2019, Wang et al. automatically evaluated BP using DL-based CNNs. Due to the lack of available datasets, manual data collection consisted of 500 patients with five facial expressions each performed. Accuracies of 82.90% and 95.60% were achieved using the end-to-end training method and cascaded training method, respectively [18]. Liu et al. (2021) developed an automatic BP detection method based on facial landmarks and ML models. The dataset consisted of 33 facial videos for both training and testing due to the difficulty of collecting facial paralysis samples, achieving an accuracy of 88.9% [19]. Parra-Dominguez et al. (2022) used MLP, SVM, MNLR, and KNN to detect BP and classified it into three stages: strong, mild, and normal. The databases used were YFP and CK+. The detection accuracy was 95.61% and the leveling accuracy was 95.58%. The problems of this study were the difficulties in accurately measuring facial symmetry concerning head tilt angle, the need to claim legal ownership of the image database, and the need for known points for tilt correction in facial symmetry measurements [20]. A recent study by Amsalam et al. (2023) used a learning-based CNN to create a useful, high-accuracy, and safe diagnostic system for BP. 570 facial images were used in this study, containing 200 images of facial paralysis. The accuracy reached was 98% [21]. The system accuracy and number of data used are listed in Table 1.

Table 1. Comparison of previous works.

Work	The used method	Technique	Dataset	Accuracy
[21]	GR	SIFT SVM	FEI live image SCIEN	98% 94% 91%
[9]	GR	Viola-Jones MobileNet v2	Standard dataset Real fair scenario	94.99% 92.7%
[17]	GR	CNN MAM	Common Voice Korean speech recognition	96% 97%
[13]	GR	DCGAN CNN	CelebA	82.40%
[14]	GR	DT, Logistic classification, KNN, Neural Networks SVM, Random Forest, Boosting	Gender classification dataset	98.4%
[19]	Pterygium	HSV-Sigmoid	UBIRIS MILES Australia Pterygium Brazil Pterygium	88.7%
[20]	Pterygium	Fully CNN	Different dataset	81.1%
[12]	Pterygium	HMC	External eye disease images	%75.71
[18]	BP	CNN	1049 clinical images	97.5%
[16]	BP	CNNs	500 patient images	82.9% 95.6%
[15]	BP	Facial landmarks ML models	33 facial videos	88.9%
[10]	BP	MLP, SVM, MNLR and KNN	YFP CK+	95.61% 95.58%
[11]	BP	CNN	570 images	98%

This study addresses the gap in previous studies by comparing the performance of three DL algorithms, namely VGG16, ViT, and Xception in detecting gender, pterygium, and Bell's paralysis using the same number of epochs and datasets. It highlights the limitations and challenges faced in previous studies and aims to increase diagnostic accuracy.

Nomenclature & Symbols

AI	Artificial Intelligence	ML	Machine Learning
AR	Age Recognition	BP	Bell's paralysis
CNN	Convolutional Neural Network	LR	Logistic Regression
DCGAN	Deep Convolutional Generative Adversarial Network	RF	Random Forest
DT	Decision tree	VGG	Visual Geometry Group
DTL	Decision tree Logistic	ViT	Vision Transformers
GR	Gender Recognition	HMC	Hierarchical Multi-label Classification
KNN	K-Nearest Neighbors	RF	Random Forest
MAM	multi-attention module	DL	Deep Learning
NN	Neural Networks	HSV	Hue, Saturation, and Value
SIFT	Deep Learning Scale-Invariant Feature Transform		
SVM	Support Vector Machine		

2. Materials and Methods

2.1. Research Ethics and Participants

This study is based on the guidelines of the Helsinki Declaration of Research Ethics (Finland 1964) and was agreed upon by the Research Ethics Committee of the Dhi Qar Health Department, the Iraqi Ministry of Health (research protocol no.: 362/2022). The datasets were collected from Roboflow, Kaggle, hospitals, and the internet to create a reliable, robust, high-accuracy system, as shown in Figure 1.

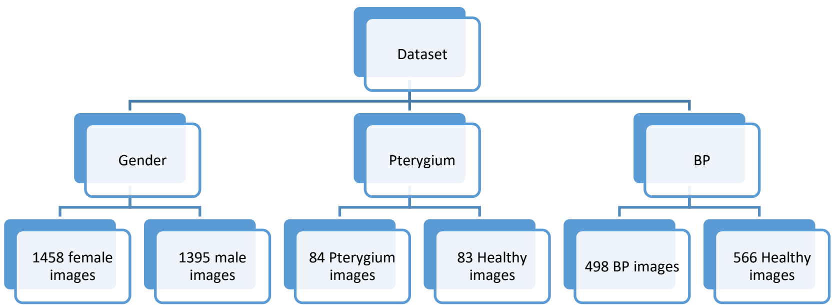


Fig. 1. The used datasets.

2.2. Experimental Setup

The system used a camera to extract information from key facial regions and divided it into three categories to recognize gender if a patient had pterygium, and BP or not based on training datasets (80% training and 20% testing) using the Python version 3.11 including a set of packages (cv2, numpy, torch, DeepFace, ViT for Image Classification and ViT Image Processor). The experimentation setup of the proposed solution is shown in Figure 2. The camera was placed at a maximum distance of 1m from the patient's face. The face was oriented to get the optimum room lighting for accurate feature detection. The experiments used a DELL CORE i5 computer with 16GB RAM and 4GB GPU.

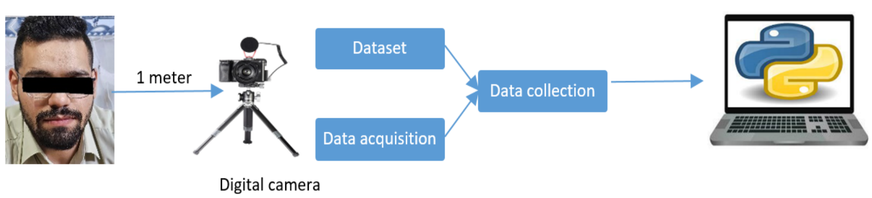


Fig. 2. The experimental setup of the proposed diagnostic system using a digital camera to obtain facial information and a Python program for processing.

2.3. System Design

The proposed system consists of different stages, starting with the data acquisition using a digital camera. Then, the data was divided into three categories - the first category determines the gender, the second category determines whether the patient suffers from pterygium or not, and the last category diagnoses BP. The size of the data captured by the camera was standardized so that the program could easily read it, and then the process of AI and DL began. After completing these stages, the patient was diagnosed according to the trained data, and the results were displayed on the computer screen. Figure 3 shows the main block diagram of the proposed system. Data programming begins after collecting it using the Python program, where a number of packages are defined to detect facial features, determine the size of the image, and then store it. The image is then classified according to storage folder categories and data. Face detection begins by pre-processing images and extracting key points on the face, such as the nose, mouth, eyes, and eyebrows [22]. This technique depends on the compatibility of the right and left halves of the face, where the movement of a healthy person's face is consistent in both halves, while a person with BP suffers from difficulty in closing the eye and limited movement of one side of the lip.

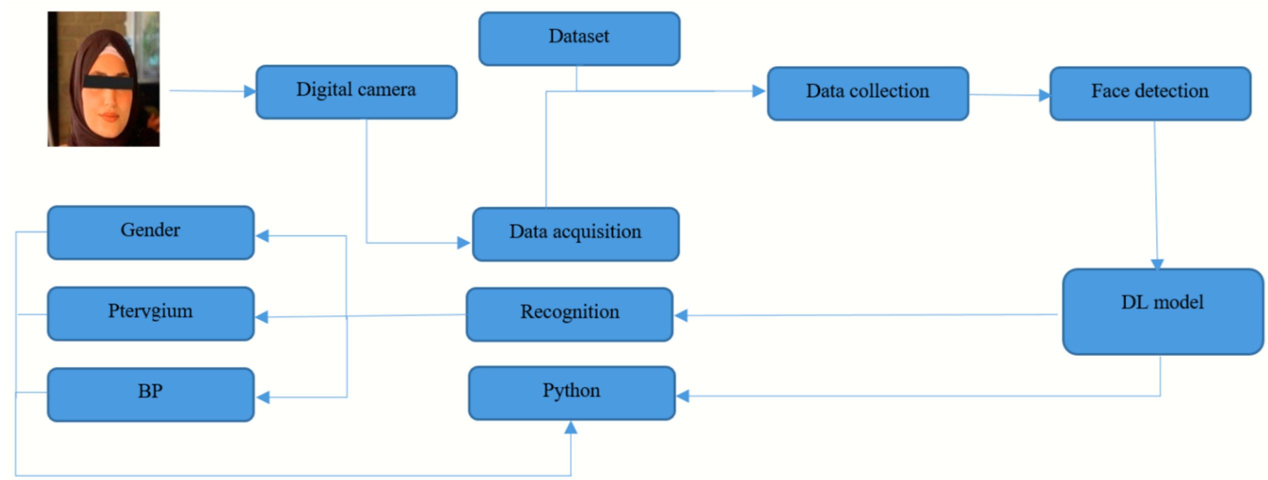


Fig. 3. The block diagram of the proposed system**.**

2.4. Evaluation Metrics

The accuracy, precision, recall, F1 Score, and MCC of the proposed system are calculated as evaluation metrics according to Eq. (1), (2), (3), (4), and (5) respectively [18, 20, 23].

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

$$\text{Precision} = TP/(TP + FP) \tag{2}$$

$$\text{Recall} = TP/(TP + FN) \tag{3}$$

$$\text{F1 Score} = 2 * (\text{Precision} \times \text{Recall}/ \text{Precision} + \text{Recall}) \tag{4}$$

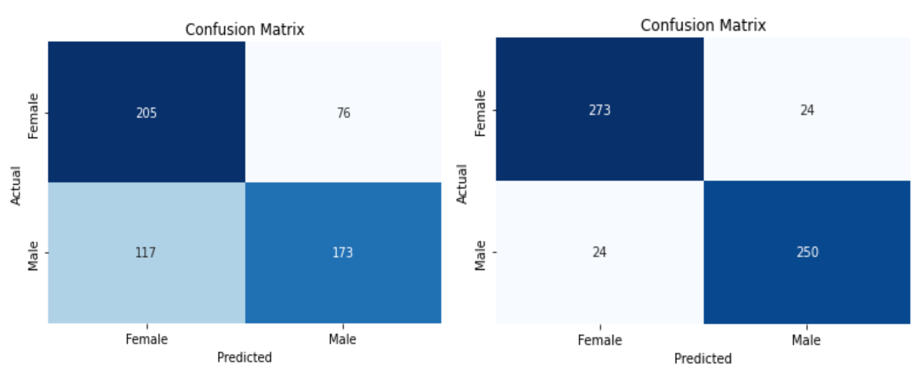
$$MCC = (TP * TN - FP * FN)/ \text{sqrt}((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)) \tag{5}$$

where TP stands for true positives, FP for false positives, TN for true negatives, and FN for false negatives.

3. Experimental Results and Discussion

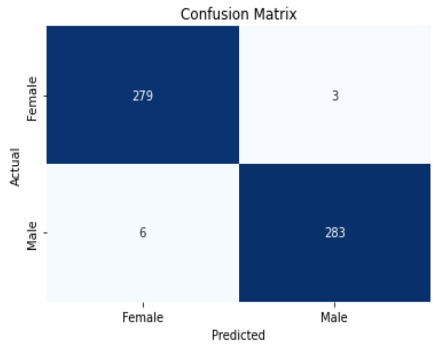
The dataset was collected from several sources: GR images were collected from Kaggle, the pterygium dataset was collected from Roboflow, and the BP images were collected from hospitals and the internet. 2853 images were used as GR images (1458 for females and 1395 for males). The number of eye health images was 167, and they were divided into 84 pterygium images and 83 normal eyes. For BP detection, 1064 total images were separated into 498 BP patients and 566 normal people.

The rest of the section details the performance of three different algorithms and presents which one has the best performance for facial characteristics and disease detection. The data evaluation process used various metrics, including accuracy, precision, recall, F1 Score, and Matthews Correlation Coefficient (MCC). The obtained results of the various DL techniques for GR using a dataset from Kaggle for female and male images are shown in Figure 4 and Table 2.



a) Xception

b) VGG16



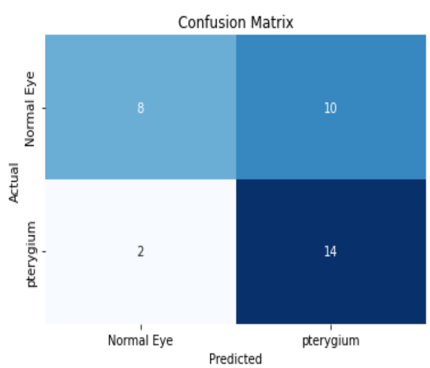
c) ViT

Fig. 4. The confusion matrices for GR are based on a) Xception, b) VGG16, and c) ViT.

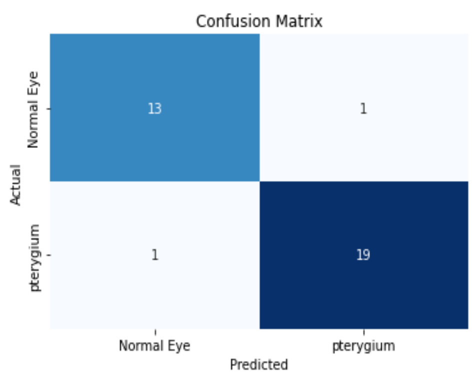
Table 2. Evaluation metrics comparison among different DL algorithms for detecting GR.

Technique	Accuracy	Precision	Recall	F1 Score	MCC
Xception	66.2%	0.7295	0.6366	0.6799	0.3287
VGG16	91.5%	0.9192	0.9192	0.9192	0.8316
ViT	98.42%	0.9895	0.9789	0.9841	0.9685

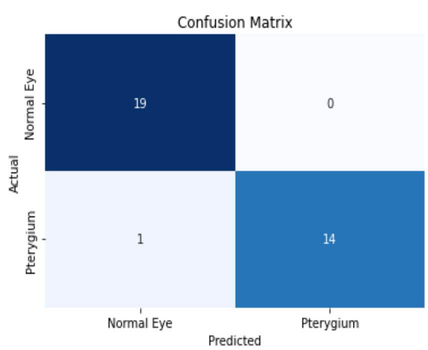
The results for the three DL techniques for pterygium detection using the images from the Roboflow website are shown in Figure 5 and Table 3.



a) Xception



b) VGG16



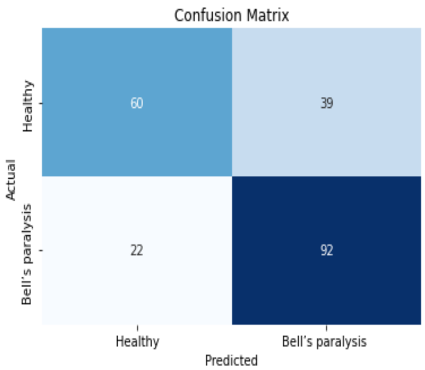
c) ViT

Fig. 5. The confusion matrices for pterygium are based on a) Xception, b) VGG16, and c) ViT.

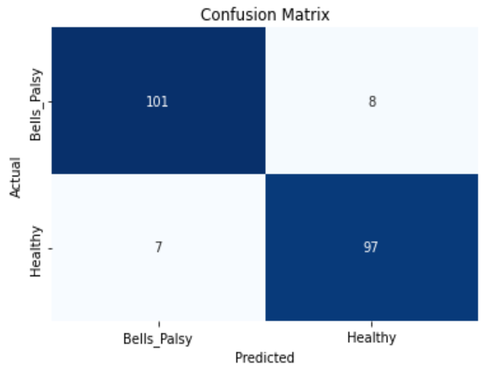
Table 3. Evaluation metrics comparison among different DL algorithms for detecting pterygium.

Technique	Accuracy	Precision	Recall	F1 Score	MCC
Xception	64.71%	0.4444	0.8000	0.5714	0.3499
VGG16	94.12%	0.9286	0. 9286	0. 9286	0.8786
ViT	97.06%	1.0000	0.9500	0.9744	0.9416

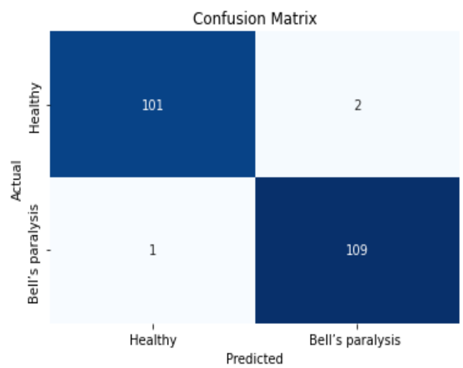
The results of three DL techniques for BP detection were used from hospital and internet images, and the results are shown in Figure 6 and Table 4.



a) Xception



b) VGG16



c) ViT

Fig. 6. The confusion matrices for BP are based on a) Xception, b) VGG16, and c) ViT.

Table 4. Evaluation metrics comparison among different DL algorithms for detecting BP.

Technique	Accuracy	Precision	Recall	F1 Score	MCC
Xception	71.36%	0.6061	0.7317	0.6630	0.4234
VGG16	92.96%	0.9266	0. 9352	0. 9309	0.8591
ViT	98.54%	0.9806	0.9902	0.9854	0.9718

It was shown in this chapter that the VIT algorithm achieved the best results in identifying gender, pterygium, and detecting BP with an accuracy rate of 98.42%, 97.06%, and 98.54% respectively, superior to previous studies.

4. Conclusion

This research introduced a contemporary approach to detecting facial characteristics and diseases using an imaging technique based on DL, diverging from conventional methods reliant on human diagnosis. The study employed three AI techniques based on DL, alongside face detection methods, computer vision, and digital camera utilization, to optimize diagnostic accuracy and facilitate result comparisons. The proposed system illustrated the effectiveness of ViT technology in identifying gender, pterygium, and BP in contrast to alternative techniques. The integration of this system into medical practice reduced physicians' diagnostic and treatment duration. The proposed system also allowed patients to operate it at home, streamlining time and effort. To implement this system in real-time, algorithms needed to be developed, and a large number and variety of high-resolution images needed to be added for training and testing to achieve the best possible performance. Using a diverse dataset and developing approved algorithms can improve the performance of imaging diagnostic systems and their adoption in the medical field.

References

[1] E. Salem, M. Hassaballah, M. M. Mahmoud, and A.-M. M. Ali, "Facial features detection: a comparative study," in *The international conference on artificial intelligence and computer vision*, 2021, pp. 402-412: Springer. doi: 10.1007/978-3-030-76346-6_37

[2] J. Qiang, D. Wu, H. Du, H. Zhu, S. Chen, and H. Pan, "Review on facial-recognition-based applications in disease diagnosis," *Bioengineering*, vol. 9, no. 7, p. 273, 2022. doi: 10.3390/bioengineering9070273

[3] S. A. Mansukhani *et al.*, "Incidence and distribution of ocular disorders in the first year of life," vol. 27, no. 2, pp. 80. e1-80. e5, 2023. doi: 10.1016/j.jaapos.2023.02.002

[4] J. Rajangam *et al.*, "Bell Palsy: Facts and Current Research Perspectives," vol. 23, no. 2, pp. 203-214, 2024. doi: 10.2174/1871527322666230321120618

[5] S. S. Ahamed, J. Jabez, and M. Prithiviraj, "Emotion Detection using Speech and Face in Deep Learning," in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, 2023, pp. 317-324: IEEE. doi: 10.1109/ICSCSS57650.2023.10169784

[6] S. Colaco, Y. J. Yoon, and D. S. Han, "UIRNet: Facial Landmarks Detection Model with Symmetric Encoder-Decoder," in *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2022, pp. 407-410: IEEE. doi: 10.1109/ICAIIC54071.2022.9722657

[7] D. Mamieva, A. B. Abdusalomov, M. Mukhiddinov, and T. K. J. S. Whangbo, "Improved face detection method via learning small faces on hard images based on a deep learning approach," vol. 23, no. 1, p. 502, 2023. doi: 10.3390/s23010502

[8] K. Takahashi, "The resolution of immunofluorescent pathological images affects diagnosis for not only artificial intelligence but also human," 2022. doi: 10.34172/jnp.2021.26

[9] S. Kumar, S. Singh, and J. Kumar, "Gender classification using machine learning with multi-feature method," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 2019, pp. 0648-0653: IEEE. doi: 10.1109/CCWC.2019.8666601

[10] A. Greco, A. Saggese, and M. Vento, "Digital signage by real-time gender recognition from face images," in *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT*, 2020, pp. 309-313: IEEE. doi: 10.1109/MetroInd4.0IoT48571.2020.9138194

[11] A. Tursunov, Mustaqeem, J. Y. Choeh, and S. J. S. Kwon, "Age and gender recognition using a convolutional neural network with a specially designed multi-attention module through speech spectrograms," vol. 21, no. 17, p. 5892, 2021. doi: 10.3390/s21175892

[12] H. Feng, "Face-based gender recognition with small samples generated by DCGAN using CNN," in *Fifth International Conference on Computer Information Science and Artificial Intelligence (CISAI 2022)*, 2023, vol. 12566, pp. 634-640: SPIE. doi: 10.1117/12.2668011

[13] V. S. Kumar, N. B. Reddy, and U. Kiirran, "Predictive Analytics on Gender Classification using Machine Learning," in *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2023, pp. 1634-1639: IEEE

[14] W. M. D. W. Zaki, M. M. Daud, S. R. Abdani, A. Hussain, H. A. J. C. m. Mutalib, and p. i. biomedicine, "Automated pterygium detection method of anterior segment photographed images," vol. 154, pp. 71-78, 2018. doi: 10.1016/j.cmpb.2017.10.026

[15] M. A. Zulkifley, S. R. Abdani, N. H. J. M. T. Zulkifley, and Applications, "Pterygium-Net: a deep learning approach to pterygium detection and localization," vol. 78, pp. 34563-34584, 2019. doi: 10.1007/s11042-019-08130-x

[16] H. M. Ahmad and S. R. Hameed, "Eye diseases classification using hierarchical MultiLabel artificial neural network," in *2020 1st. Information Technology To Enhance e-learning and Other Application (IT-ELA*, 2020, pp. 93-98: IEEE. doi: 10.1109/IT-ELA50150.2020.9253120

[17] A. Song, Z. Wu, X. Ding, Q. Hu, and X. J. F. I. Di, "Neurologist standard classification of facial nerve paralysis with deep neural networks," vol. 10, no. 11, p. 111, 2018. doi: 10.3390/fi10110111

[18] T. Wang, S. Zhang, L. A. Liu, G. Wu, and J. J. I. A. Dong, "Automatic facial paralysis evaluation augmented by a cascaded encoder network structure," vol. 7, pp. 135621-135631, 2019. doi: 10.1109/ACCESS.2019.2942143

[19] Y. Liu, Z. Xu, L. Ding, J. Jia, and X. Wu, "Automatic assessment of facial paralysis based on facial landmarks," in *2021 IEEE 2nd International Conference on Pattern Recognition and Machine Learning (PRML)*, 2021, pp. 162-167: IEEE. doi: 10.1109/PRML52754.2021.9520746

[20] G. S. Parra-Dominguez, C. H. Garcia-Capulin, and R. E. J. D. Sanchez-Yanez, "Automatic facial palsy diagnosis as a classification problem using regional information extracted from a photograph," vol. 12, no. 7, p. 1528, 2022. doi: 10.3390/diagnostics12071528

[21] A. S. Amsalam, A. Al-Naji, A. Y. Daeef, and J. J. J. o. T. Chahl, "Computer Vision System for Facial Palsy Detection," vol. 5, no. 1, pp. 44-51, 2023. doi: 10.51173/jt.v5i1.1133

[22] J. Barbosa, W.-K. Seo, and J. J. B. M. I. Kang, "paraFaceTest: an ensemble of regression tree-based facial features extraction for efficient facial paralysis classification," vol. 19, no. 1, pp. 1-14, 2019. doi: 10.1186/s12880-019-0330-8

[23] M. Grandini, E. Bagli, and G. J. a. p. a. Visani, "Metrics for multi-class classification: an overview," 2020. doi: 10.48550/arXiv.2008.05756