WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Age, Gender and Emotion Prediction from Face Using Deep Learning

Kotla Pramodhini[1], Mr. G. Dharmaraju[2], Kuramana Shalini[3], Koyyala Sahithi[4], Kottakota Kavyasree[5]

[1,3,4,5]UG Scholar, Department of CSE, GMRIT, Rajam, Andhra Pradesh, India
[2]Assistant Professor, Department of CSE, GMRIT, Rajam, Andhra Pradesh, India
kotlapramodhini@gmail.com[1], dharmaa.surya@gmail.com[2], shaliniKurmana@gmail.com[3], koyyalasahithi@gmail.com[4], kavyasree.kottakota@gmail.com[5]

ABSTRACT

Automatic prediction of emotion, age and gender from face images has drawn a lot of attention recently, due it is wide applications in various facial analysis problems. The study aims to explore the relationship between personality traits, intelligence, and facial images. In this study, deep learning based on the convolutional neural networks (CNN) to predict emotion, gender and age group of facial images with high accuracy rate. Deep learning (DL) based detection gives performance better than traditional methods with image processing. The model leverages the power of deep learning to accurately predict these attributes from facial images. By training a model on a large dataset of labeled images, the model can learn to recognize patterns and make accurate predictions. The modified convolutional neural network (M-CNN) for feature extraction and a Multi Support Vector Machine (M-SVM) for classification are used. The proposed model achieved high accuracy in emotion (97.47%), age (98.43%), and gender (96.65%) classification. This model is trained on a popular face age and gender dataset, and achieved promising results.

Keywords: facial age, gender, emotion recognition, on Deep Convolutional Neural Networks (DCNN), M-SVM

Introduction

Age, gender, and emotion prediction from facial images using deep learning techniques have gained significant attention in recent years. Researchers have explored various approaches to extract features from human face images for these prediction. Different models on Deep Convolutional Neural Networks (DCNN) have been proposed to improve accuracy and efficiency. It is important in various applications such as social media, security control, advertising, and entertainment. It plays a vital role in interpersonal communication and is a significant area of research in computer vision. The system evaluated using classification rate, precision, and recall on datasets and real-world images. Every human face is different and has essential characteristics that can reveal information about age, gender, race, personality, and other aspects of the face, such as emotion. The contours of the eyes, head, lips, nose, jaw line, eyebrow, eyelashes, wrinkles, and so forth are among the most significant facial landmarks that can be identified. These particular face traits are present in every human being. These subtle differences between males and females are present in individuals, and the face muscles contract with age. Furthermore, computers and other technologies find it extremely difficult to recognise and decode these face clues, even if it comes naturally to humans. Our emphasis was on identifying and capturing features from the detected head. Consequently, we collected features from the ear, hair, and face. we employed a head detection technique utilizing YOLOV3 (You Only Look Once Version 3). This involved creating a head detector that utilizes YOLOV3 to identify the head within the entire image. Once the head was detected, we utilized it for recognizing emotions, age, and gender, along with their corresponding accuracies. Unlike traditional approaches that solely focus on the face, our methodology emphasized the importance of the entire head. Consequently, we gathered features not only from the face but also from the hair and ears. The introduction of the YOLO version 3 network played a key role in effectively detecting the head from the test image in this process.

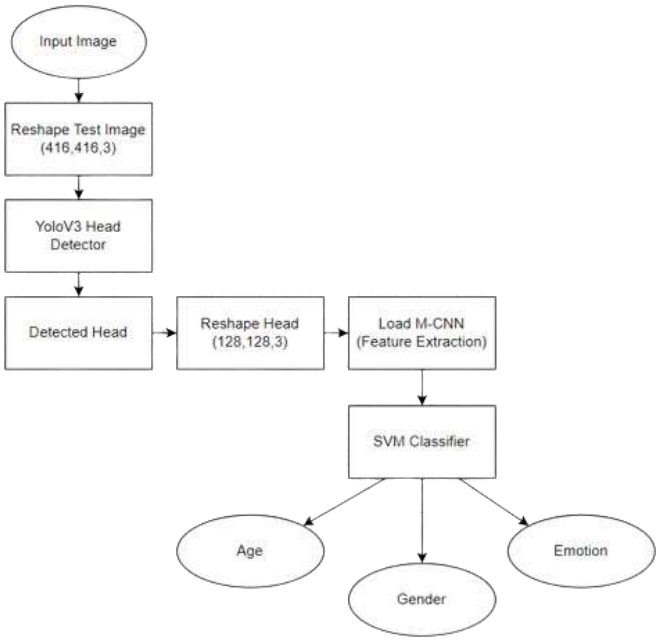
It uses deep learning algorithms, such as deep convolutional neural networks (DCNN) and support vector machines (SVM), to predict age and gender from face images. Different methodologies have been used to identify gender based on Biometric traits, such as facial features and structures, have been utilized to identify gender. These traits include the shape of

the face, jawline, eyebrows, and other facial landmarks, behaviors. Age prediction is based on age-specific characteristics and skin changes in the face image. modified convolutional neural network (M-CNN) is used for feature extraction. Support Vector Machine (SVM),CNN methods are used. The proposed model is trained on the datasets namely Age Recognition (AR-2022), UTKFace, CK+ Dataset has shown 98.43% accuracy for age and 96.65% accuracy for gender estimation and 97.47%. Accuracy for Emotion Prediction.

Literature Survey

Adenowo discusses the use of deep learning, specifically Convolutional Neural Networks (CNNs), for age, gender, and emotion recognition tasks in facial images. He compares the performance of the trained model on different datasets, including Adience, FER2013, and a custom dataset called "blackfaces"[1]. Abdolrashidi, Amirali proposed a comprehensive image classification model was developed using a multi-task learning approach, where the same CNN architecture is shared across age, gender, and emotion recognition tasks. This allows the model to jointly learn features that are useful for all three tasks, potentially improving overall performance[2]. Uddin, Md Jashim implements a general Convolutional Neural Network (CNN) for real-time face detection, gender classification, and emotion classification simultaneously. It involves dataset preparation, preprocessing, and a powerful classification model[3]. Sharma, Neha highlights the use of deep architectures in CNN algorithms and it does not compare the proposed model with a wide range of existing algorithms[4]. Mellouk, Wafa authors review a variety of deep learning architectures that have been used for facial emotion recognition, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures that combine CNNs and RNNs[5]. Haseena captures a person's face and predicts their age and gender for nutritional advice. Real-world images captured in real-time are also used for evaluation. It captures the face and classifies the age and gender of an individual only[6]. Katna, Rishabh applied natural language processing (NLP) and machine learning techniques for author profiling. Machine learning classifiers including logistic regression (LR), random forest (RF), decision tree (DT), and support vector machine (SVM) were used in combination with the NLP techniques[7]. Agbo-Ajala propose a model that uses CNN architecture to predict the age group and gender of human's faces from unfiltered real-world environment. The model's suitability for age group and gender classification is confirmed by the robust image preprocessing algorithm, which manages part of the variability seen in typical unfiltered real-world faces. It uses a novel end-to-end CNN approach with a two-level architecture for age and gender classification of unfiltered real-world faces[8]. Juel, Juel Sikder utilizes a modified convolutional neural network (M-CNN) for feature extraction and a Multi Support Vector Machine (M-SVM) for classification. They introduced a modified convolutional neural network (M-CNN) for feature extraction, followed by a Multi Support Vector Machine (M-SVM) for classification. Additionally, the paper incorporates a head-detecting technique using the You Only Look Once Version 3 (YOLOV3) network, which extracts features from the entire head instead of just the face[9]. Levi, Gil proposes a simple convolutional net architecture that utilizes deep-convolutional neural networks (CNN) to improve performance on age and gender classification tasks. And emphasize the generality of their network design, which performs well across different, related problem[10]. Madinah proposes a hybrid feature extraction and facial expression recognition method that utilizes Viola-Jones cascade object detectors and Harris corner key-points for face and facial feature extraction, and principal component analysis, linear discriminant analysis, histogram-of-oriented gradients (HOG) feature extraction, and support vector machines (SVM) for emotion classification[11]. Dehghan, Afshin describes Sighthound's fully automated age, gender, and emotion recognition system using deep convolutional neural networks with lower error rates[12]. jyoti Sharma discusses a methodology and algorithm for real-time age, gender, and emotion estimation from face images on a webcam. Several techniques such as PCA, LBP, SVM, VIOLA-JONES, and HOG are mentioned as methods used for human emotion, gender, and age detection in various conditions[13]. Singh, Arjun uses a deep CNN-based WideResNet approach for age and gender prediction, and for Emotion Recognition we have used conventional CNN. It is a multi-task learning approach that allows the model to jointly learn features that are useful for all three tasks, potentially improving overall performance[14]. Savoiu focuses on recognizing human emotions through visual recognition of facial expressions using deep learning methods, specifically convolutional neural networks (CNNs)[15].

Methodology



- The main architecture of the paper is a two-phase system consisting of a Head Detection stage and a Classification stage.
- The Head Detection stage utilizes the You Only Look Once Version 3 (YOLOV3) network to detect the head in the input image. YOLOV3 provides bounding boxes to each region on the image and selects the one with the highest probability.
- The detected head is then reshaped into a (128,128,3) dimension to match the dimension of the Modified Convolutional Neural Network (M-CNN) classifier used in the Classification stage.

Classification Phase :

The classification phase in the proposed system involves predicting the emotions, gender, and age of the input image.The system uses a modified convolutional neural network (M-CNN) for feature extraction and a Multi Support Vector Machine (M-SVM) for classification .The input image is reshaped to (416,416,3) dimensions to match the head detector model dimension. The classification phase in the paper involves predicting the categories of emotion, age, and gender based on the features extracted from the head using the modified convolutional neural network (M-CNN) .

Dataset Description :

Three datasets are used in this study: AR-2022, FER-2022, and GR-2022. The FER-2022 dataset contains 14,216 images in 7 emotion categories: 'Disgust', 'Angry',

'Happy', 'Fear', 'Neutral', 'Surprise', and 'Sad'. The GR-2022 dataset contains 9,893 images in 3 gender categories: 'Male', 'Female', and 'Other'. The AR-2022 dataset contains 22,597 images in 11 age categories: '0-2', '3-6', '7-12', '13-19', '20-26', '27-33', '34-40', '41-50', '51-60', '61-75', and '75'' As the dataset and corresponding ground truth are loaded then we reshape the dataset images into (416,416,3) dimensions. We use this dimension as if we can get an optimal weight and better detection of the head. Next, align the pictures with the appropriate background.

Dataset Preprocessing :

The datasets used in this study are preprocessed before training the proposed model. Data augmentation technique is used to increase the number of images in the datasets. This technique involves image rotation and zooming. The preprocessed images are then used for feature extraction using a modified convolutional neural network (M-CNN). The M-CNN consists of five convolutional layers with relu, batch normalization, max-pooling, and dropout functions.

Multi Support Vector Machine (M_SVM) :

The proposed system utilizes a Multi Support Vector Machine (M-SVM) algorithm for classification purposes. M-SVM is used as a classifier in the modified convolutional neural network (M-CNN) for feature extraction. Support Vector Machine (SVM) is a widely used algorithm for regression and classification tasks . it specifically refers to the use of Support Vector Machines for multi-class classification tasks. In multi-class classification, the goal is to categorize data points into more than two classes or categories.

Modified Convolutional Neural Network (M_CNN):

It's a deep learning architecture used for various computer vision tasks, such as image classification, object detection, and segmentation. The "multi-column" design implies that the network consists of multiple parallel columns or branches, each processing an input image differently. The M-CNN is designed to extract features not only from the face but also from the head, including hair and ears. The M-CNN is trained on datasets such as FER-2022, GR-2022, and AR-2022 for emotion, gender, and age classification respectively. Thus introduced a modified convolutional neural network (M-CNN) for feature extraction in the recognition of emotions, gender, and age .

we focused on a modified method to recognize emotion along with gender and age using a multi-support vector machine (M-SVM), modified convolutional neural network (M-CNN) and YOLO version 3 network. Instead of face detection, our prediction process incorporates a head-detection technique. We have developed a system utilizing the You Only Look Once Version 3 (YOLOV3) network to identify heads. This approach allows us to extract features from the entire head rather than just focusing on the face.

The proposed model divides into two phases:

• Head detection stage

• Classification stage

Head detection stage -

Initially, the system loads image datasets. Reshape the dataset images into (416,416,3) dimensions. YOLOV3 network is defined to detect the head from the input image. Train the YOLOV3 model to successfully detect the head. The evaluation process of YOLOV3 is elaborated, and the head detector is defined.

Classification stage –

After detecting the head, the images are preprocessed for Classification. The classification phase involves using the extracted features from the M-CNN and feeding them into the M-SVM for predicting the emotion, age, and gender categories.

The M-CNN architecture is tailored for feature extraction from the detected head. It consists of five convolutional layers with relu activation, batch normalization, max-poolin

g, and dropout functions. The model is trained on the diverse datasets to effectively capture features relevant to age, gender, and emotion. The extracted features from the M-CNN are used to train an M-SVM for the classification tasks. SVMs are well-suited for handling multi-class classification problems and are trained to predict age groups, gender categories, and emotional states based on the learned features. The proposed approach achieves remarkable accuracies of 97.47%, 98.43%, and 96.65% for emotion, gender, and age classification, respectively. These high accuracies demonstrate the effectiveness of the multi-stage deep learning approach in facial analysis. The combination of YOLOV3 for head detection, M-CNN for feature extraction, and M-SVM for classification proves to be a powerful strategy, achieving impressive accuracies in predicting emotion, gender, and age categories.

Results and Discussion

S.no	Method	Accuracy		
		Age	Gender	Emotion
1	M-CNN, M-SVM	98.43%	96.65%	97.47%
2	Multi-task CNN, HOG	65.5%	96.94%	63.71%
3	Deep CNN	96.2%	76.1%	91%
4	HOG+viola-jones algorithm	71.53%	98.90%	97%
5	wide-resnet based CNN approach	96.26%	96.26%	69.2%

Histogram of Oriented Gradients (HOG) features are combined with Multi-task CNN , although it does a good job of classifying gender, its accuracy in estimating age and recognising emotions is lower. Using a Deep Convolutional Neural Network (Deep CNN) to obtain high accuracy in age estimation and emotion identification, but comparably poor accuracy in gender categorization. Using a Convolutional Neural Network with a Wide-R esNet foundation, classifies people according to their age and gender with high accuracy but shows less accuracy in identifying emotions. Employing Convolutional Neural Networks (M-CNN) and Support Vector Machines (M-SVM), achieves high accuracy across all tasks, especially in age estimation.

Conclusion

Age, gender and emotion recognition from facial images are important tasks in facial recognition research. Face age, gender, and emotion recognition are powerful tools that have a wide range of potential applications. However, it is important to use these tools responsibly and to consider the ethical implications of their use. With careful planning and implementation, facial analysis can be a valuable tool for improving security, marketing, healthcare, and human-computer interaction. It uses three datasets namely FER-2022 (Face Emotion Recognition), AR-2022 (Age Recognition), and GR-2022 (Gender Recognition). It utilizes a modified convolutional neural network (M-CNN) for feature extraction and a Multi Support Vector Machine (M-SVM) for classification. The M-CNN is trained on three datasets named AR-2022, FER-2022 and GR-2022. The M-SVM is trained on the extracted features from the M-CNN. The proposed approach achieves an accuracy of 97.47%, 98.43% and 96.65% for emotion, gender and age classification.

References

1. Haseena, S., et al. "Prediction of the age and gender based on human face images based on deep learning algorithm." *Computational and Mathematical Methods in Medicine* 2022 (2022).
2. Abdolrashidi, Amirali, et al. "Age and gender prediction from face images using attentional convolutional network." *arXiv preprint arXiv:2010.03791* (2020).
3. Uddin, Md Jashim, et al. "A convolutional neural network for real-time face detection and emotion & gender classification." *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN* (2020): 2278-2834.
4. Sharma, Neha, Reecha Sharma, and Neeru Jindal. "Face-based age and gender estimation using improved convolutional neural network approach." *Wireless Personal Communications* 124.4 (2022): 3035-3054.
5. Mellouk, Wafa, and Wahida Handouzi. "Facial emotion recognition using deep learning: review and insights." Procedia Computer Science 175 (2020): 689-694.
6. Adenowo, Adetokunbo A., and Adekunle Olu Adewole. "Performance Evaluation of Deep Learning Method for Age, Gender and Emotion Recognition Tasks." *African Journal of Computing & ICT* 13.4 (2020).
7. Katna, Rishabh, et al. "Machine learning based approaches for age and gender prediction from tweets." *Multimedia Tools and Applications* 81.19 (2022): 27799-27817.
8. Agbo-Ajala, Olatunbosun, and Serestina Viriri. "Deeply learned classifiers for age and gender predictions of unfiltered faces." *The Scientific World Journal* 2020 (2020).
9. Juel, Juel Sikder. "An Approach Based on Deep Learning for Recognizing Emotion, Gender and Age." (2022).

10. Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2015.

11. Madinah, Saudia Arabia. "Emotion detection through facial feature recognition." International Journal of Multimedia and Ubiquitous Engineering 12.11 (2017): 21-30.

12. Dehghan, Afshin, et al. "Dager: Deep age, gender and emotion recognition using convolutional neural network." *arXiv preprint arXiv:1702.04280* (2017).

13. jyoti Sharma, Dibya, Sachin Dutta, and Dibya jyoti Bora. "REGA: Real-Time Emotion, Gender, Age Detection Using CNN—A Review." (2020).

14. Singh, Arjun, et al. "Age, gender prediction and emotion recognition using convolutional neural network." *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*. 2021.

15. Savoiu, Alexandru, and James Wong. "Recognizing facial expressions using deep learning." *Recognizing Facial Expressions Using Deep Learning* (2017): 1-8.