RESEARCH ARTICLE OPEN ACCESS

Enhanced Human Age and Gender Detection Using Convolutional Neural Networks

Santosh Kumar, Harshvardhan Tailor, Hemant Singh Jadoun, Mandeep Kumar Biloniya, Aryan Jangid

Department of CS/IT/AI&DS, Global Institute of Technology, Jaipur
Department of AI&DS, Global Institute of Technology, Jaipur
Department of AI&DS, Global Institute of Technology, Jaipur
Department of AI&DS, Global Institute of Technology, Jaipur
Department of AI&DS, Global Institute of Technology, Jaipur

ABSTRACT

The primary motivation behind developing an automated method for gender and age detection in humans stems from its crucial role in pattern recognition and computer vision. Beyond age and facial emotion detection, it holds significance in various applications within computer vision, particularly in non-verbal communication methods such as facial expressions and hand gestures, which are integral to human-computer interaction.

While considerable research has successfully advanced computer modeling of age, gender, and emotions, it still lags behind the capabilities of the human vision system. In this project, we propose an architecture for age and gender classification utilizing Convolutional Neural Networks (CNNs). Our model demonstrates superior accuracy in both age and gender detection compared to other classifier-based methods. Additionally, for modeling human emotions, we aim to predict emotions using deep CNNs to enhance future marketing strategies.

We employ the Viola-Jones pre-processing algorithm to extract features from images, which are then fed as input to the CNN. The results of the detection process are presented to the user through a well-designed user interface.

Keywords — Face Detection, Viola-Jones, Deep CNN

I. INTRODUCTION

Age detection from images plays a very vital role in human and computer vision which has wide range of applications like forensics or social media. It can detect other biometrics of human and such as age, gender, and emotions. Wide researches are conducted already to detect age using facial features. Various public standard datasets can be used for a realtime age detection which helps public performance comparison of desired methods. As a result, a loads of active researches has been done, with several recent works using the concept of Convolutional Neural Networks (CNN) for extraction of features.

Facial expressions can be recognized using non-verbal communication between humans, along with the interpretation of facial expressions is being widely studied. Facial expressions plays an important role in human interaction, Facial Expression Recognition (FER) algorithm with the help of computer vision helps in applications such as human-computer interactions and data analytics.

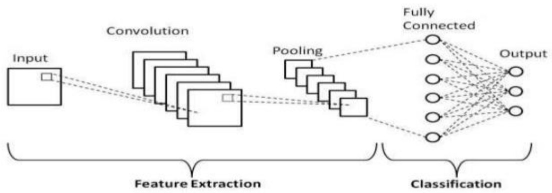

Figure 1: - CNN architecture

II. LITERATURE SURVEY

Our exploration into the background commenced by delving into research papers and online blog posts relevant to our topic. One notable research paper introduces a novel framework for facial expression recognition utilizing an attentional convolutional network. Attention proves crucial in detecting facial expressions, enabling neural networks with fewer than 10 layers to rival much deeper networks in emotion recognition tasks. Despite successful face recognition, challenges persist due to factors such as illumination, pose variation, facial expressions, and facial components like eyebrows, nose, and mouth length. To address these, the researchers utilized the Dlib library in OpenCV to handle face recognition tasks, employing methods such as Adaboost for feature selection and the Viola-Jones algorithm for Haar-like feature extraction, which are then fed into a CNN model for processing. Their deep model underwent training on a vast dataset comprising four million images for face recognition purposes. This model serves as the foundation for our facial attribute recognizers, which are further fine-tuned for tasks including apparent age estimation, gender recognition, and emotion recognition. We collected images from various sources, amassing over four million images encompassing more than 40,000 individuals for facial recognition. Each image is labeled with gender, and the data is annotated with emotions. These images undergo a semi-automated trimming process, with human annotators involved to ensure accuracy. Subsequently, the images undergo preprocessing to extract and align the faces.

III. DEVELOPMENT PROCESS

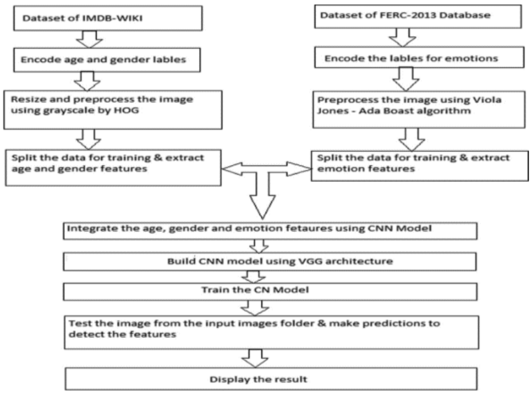

Figure 1: - Proposed System Architecture

For Age and Gender Detection is used for age estimation which can be seen in image classification and object detection fueled by deep learning. From the deep learning concept, we learn four key ideas that we apply to our solution:

- The deeper the neural networks (by a sheer increase of parameters/model complexity) the better the capacity to model highly non-linear transformations - with some optimal depth on current architectures.
- The larger and more diverse the datasets used for training, the better the network learns to generalize and the more effective it becomes to over-fitting.
- The alignment of the object in the input image impacts the overall performance.
- Use When the training dataset is small that is a network pre-trained for comparable inputs which will help us from the transferred knowledge.

Our procedure consistently initiates by rotating the input image at various angles to identify the face exhibiting the highest score. We then align the face using the angle and crop it for further steps. This is a simple and effective procedure that does not involve facial landmark detection. We use deep VGG-16 architecture for our Convolutional Neural Network (CNN). We start from pre-trained CNNs on the large ImageNet dataset to classify images such that it helps us by discriminating 1000 object categories in images by the representations learned, and to obtain a meaningful representation and a smooth and warm start for further fine-tuning on relatively smaller face datasets.

Number Managing the CNN on facial images along with age annotations is a very important step for better performance as we know the CNN adapts to best fit the particular data distribution and perform effective age detection. Because of shortage of facial images with apparent age annotations, we go for the benefits of adjusting over crawled internet face images. We compute 523,050 face images from the IMDb and Wikipedia websites to form IMDB-WIKI our new dataset.

While we address age estimation within the realm of regression, we extend our approach by framing age estimation as a multi-class classification of age categories, followed by a refinement of expected values using SoftMax.

Our main contributions are as follows:

1. The UTK Face dataset is the largest dataset with real age and gender annotations.
2. A new regression approach is employed, combining deep classification with subsequent refinement of expected values.
3. FER-2013 (Kaggle) dataset for emotion detection.

- The We introduced the IMDB-WIKI dataset for age detection, offering a detailed analysis of the projected DEX system. We applied the method and reported the results on standard age estimation datasets. Additionally, for Emotion Detection & Classification, we extensively evaluated and tested various pre-processing techniques and model architectures. Through our experimentation, we successfully developed a custom Convolutional Neural Network (CNN) model that achieved an accuracy of 70.47% on the FER-2013 test set.
- During pre-processing, we experimented with centered and scaled data, discovering that subtracting the mean significantly aids in aligning the training distribution across all sets before training and evaluation. To further enhance our model's robustness, we implemented data augmentation techniques such as random rotation, shifting, flipping, cropping, and sheering of training images. This approach led to approximately a ten percent reduction in inaccuracies.
- We implemented several CNN architectures sourced from different papers for emotion detection across various datasets. Ultimately, our custom-developed CNN architecture yielded the best performance. However, we acknowledge that detecting errors in neural networks can be challenging.
- Through our analysis, we observed errors across different emotion classes and conducted visual inspections of images classified correctly and incorrectly. An early observation was our difficulty in accurately classifying certain emotions, particularly those reliant on fine details in images, such as small facial features or curves. To address this, we increased the number of layers and reduced filter sizes in our network to enhance its capacity to capture intricate details.
- However, this adjustment led to overfitting issues, which we mitigated by implementing dropout, early stopping at approximately 100 epochs, and augmenting our training set. We noted that we could only effectively learn training set noise after achieving approximately 70% accuracy on

the development set, as evidenced by the accuracy plot during training.

- Moving forward, we suggest further exploration into enabling increased parameterization of the network for improved performance.Do not confuse "imply" and "infer".

- During real-time classification using OpenCV's Haar cascades to detect and extract a face region from a webcam video feed, we found that it's optimal neither to subtract the training means nor to normalize the pixels within the detected face region before classification.

- In real-time classification, our model demonstrated strengths in detecting neutral, happy, surprised, and angry emotions. However, illumination emerged as a crucial factor influencing the model's performance. This suggests that our training set may not accurately represent the distribution of emotions under lower brightness conditions on the screen. Further investigation into this aspect could enhance the robustness of our model.

IV. IMPLEMENTATION DETAILS

A. Dataset Download: We start by acquiring image sets from two datasets: UTK Face, which provides gender and age labels, and FER-2013.

B. Label Encoding Function: Once downloaded, we encode labels using one-hot encoding. This is necessary because Convolutional Neural Networks (CNNs) require numeric values for processing.

C. For Resizing and Pre-processing: Images are resized to a standard size of 256x256 pixels and converted to grayscale using the Histogram of Oriented Gradients (HOG) algorithm. For emotion detection, the Viola-Jones - AdaBoost algorithm is employed to extract facial features optimized for emotion recognition, reshaping the image to focus solely on facial features.

D. Feature Extraction and Integration: Facial features such as eyebrow shape, nose, mouth, and facial contours are extracted using the DLIB library in OpenCV. These features, along with age, gender, and emotional attributes, are integrated for training.

E. Training and Testing: The CNN model is constructed based on the VGG-16 architecture and trained over multiple epochs. The loss Gauss function is utilized to filter out distorted or irrelevant images. During testing, users input images, and the model predicts age, gender, and emotion by comparing them with trained images.

F. Output: Finally, fresh images undergo the same feature extraction process, with the extracted features fed into the trained ML algorithm to predict the corresponding labels.

V. REQUIREMENT ANALYSIS

A. SOFTWARE:

Prerequisites are:

- Keras2 (with TensorFlow backend)
- OpenCV
- Python 3.5
  (TensorFlow not supported in higher versions)
- Num-Py
- Tensor-Flow
- h5py (for Keras model serialization)

B. HARDWARE:

Intel core processor with high GPU power & frequency.

C. DATASET:

a. UTK Face – for age and gender.
b. FER-2013 – for emotion detection.

VI. REFERENCES

[1] Shervin Minaee, Amirali Abdolrashidi, Deep - Emotion: Facial Expression recognition using attentional convolutional network, 2019 3rd IEEE International Conference on Cybernetics (CYBCONF)doi:10.1109/cybconf.2017.7985780.

[2] Alen Salihbasi and Tihomir Orehovacki presented a paper titled "Development of Android Application for Gender, Age, and Face Recognition Using OpenCV" at MIPRO 2019 on 20 May 2019. The paper was published with ISSN: 1057-7149.

[3] Jiu-Cheng Xie, Chi-Man Pun, 'Age and Gender Classification using Convolutional Neural Networks, IEEE Workshop on Analysis and Modelling of Faces & Gestures', IEEE Transactions on Information Forensics and Security, ISSN: 1556-6013,07 March 2019.

[4] Gil Levi and Tal Hassner,' Facial Emotion Analysis using Deep Convolutional Neural Network', 2017. The International Conference on Signal Processing and Communication (ICSPC) took place in Coimbatore in 2017.

[5] Mostafa Mohammadpour, Seyyed Mohammed. R Hashemi, Hossein Khaliliardali, Mohammad. M Alyan Nezhadi,' Facial Emotion Recognition using Deep Convolutional Networks', 2017 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), 22 December 2017.

[6] Dehghan, Afshin & G. Ortiz, Enrique & Shu, Guang & Zain Masood, Syed. (2017). DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Network.

[7] 2018-Sepidehsadat Hosseini, Seok Hee Lee, Hyuk Jin Kwon, Hyung Il Koo Nam Ik Cho, "Age and Gender Classification Using Wide Convolutional Neural Network and Gabor Filter", IEEE2018.

[8] Rekha N, Dr.M.Z.Kurian "Face Detection in Real Time Based on HOG" international journal of advanced Research in computer engineering & tchnology volume 3 issue 4,april2014 ISSN:2278-1323

[9] Ajit P. Gosavi, S. R. Khot "Facial Expression Recognition Using Principal Component Analysis" ISSN: 2231-2307, Volume-3, Issue-4, September 2013

[10] A. Agarwal, R. Joshi, H. Arora and R. Kaushik, "Privacy and Security of Healthcare Data in Cloud based on the Blockchain Technology," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 87-92.

[11] G. K. Soni, D. Yadav, A. Kumar and L. Sharma, "Flexible Antenna Design for Wearable IoT Devices," IEEE 2023 3rd International Conference on

Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, pp. 863-867, 2023.

[12] Pradeep Jha, Deepak Dembla & Widhi Dubey , "Implementation of Transfer Learning Based Ensemble Model using Image Processing for Detection of Potato and Bell Pepper Leaf Diseases", International Journal of Intelligent Systems and Applications in Engineering, 12(8s), 69–80, 2024.

[13] H. Arora, G. K. Soni, R. K. Kushwaha and P. Prasoon, "Digital Image Security Based on the Hybrid Model of Image Hiding and Encryption," IEEE 2021 6th International Conference on Communication and Electronics Systems (ICCES), pp. 1153-1157, 2021. doi: 10.1109/ICCES51350.2021.9488973.

[14] Pradeep Jha, Deepak Dembla & Widhi Dubey, "Deep learning models for enhancing potato leaf disease prediction: Implementation of transfer learning based stacking ensemble model", Multimedia Tools and Applications, Vol. 83, pp. 37839–37858, 2024.

[15] G. K. Soni, A. Rawat, D. Yadav and A. Kumar, "2.4 GHz Antenna Design for Tumor Detection on Flexible Substrate for On-body Biomedical Application," 2021 IEEE Indian Conference on Antennas and Propagation (InCAP), pp. 136-139, 2021. doi: 10.1109/InCAP52216.2021.9726323.

[16] Babita Jain, Gaurav Soni, Shruti Thapar, M Rao, "A Review on Routing Protocol of MANET with its Characteristics, Applications and Issues", International Journal of Early Childhood Special Education, Vol. 14, Issue. 5, pp. 2950-2956, 2022.

[17] P. Upadhyay, K. K. Sharma, R. Dwivedi and P. Jha, "A Statistical Machine Learning Approach to Optimize Workload in Cloud Data Centre," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 276-280, doi: 10.1109/ICCMC56507.2023.10083957.

[18] Pradeep Jha, Deepak Dembla & Widhi Dubey , "Crop Disease Detection and Classification Using Deep Learning-Based Classifier Algorithm", Emerging Trends in Expert Applications and Security. ICETEAS 2023. Lecture Notes in Networks and Systems, vol 682, pp. 227-237, 2023.

[19] Gori Shankar, Vijaydeep Gupta, Gaurav Kumar Soni, Bharat Bhushan Jain and Pradeep kumar Jangid, "OTA for WLAN WiFi Application Using CMOS 90nm Technology", International Journal of Intelligent Systems and Applications in Engineering (IJISAE), vol. 10, no. 1(s), pp. 230-233, 2022.

[20] Unmasking Embedded Text: A Deep Dive into Scene Image Analysis, Maheshwari, A., Ajmera.R., Dharamdasani D.K., 2023 International Conference on Advances in Computation, Communication and Information Technology, ICAICCIT 2023, 2023, pp. 1403–1408

[21] Internet of Things (IoT) Applications, Tools and Security Techniques, Kawatra, R., Dharamdasani, D.K., Ajmera, R,et.al. 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2022, 2022, pp. 1633–1639

[22] P. Jha, D. Dembla and W. Dubey, "Comparative Analysis of Crop Diseases Detection Using Machine Learning Algorithm," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 569-574, doi: 10.1109/ICAIS56108.2023.10073831.

[23] P. Jha, R. Baranwal, Monika and N. K. Tiwari, "Protection of User's Data in IOT," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 1292-1297, doi: 10.1109/ICAIS53314.2022.9742970.

[24] P. Jha, T. Biswas, U. Sagar and K. Ahuja, "Prediction with ML paradigm in Healthcare System," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2021, pp. 1334-1342, doi: 10.1109/ICESC51422.2021.9532752.

[25] Mehra, M., Jha, P., Arora, H., Verma, K., Singh, H. (2022). Salesforce Vaccine for Real-Time Service in Cloud. In: Shakya, S., Balas, V.E., Kamolphiwong, S., Du, KL. (eds) Sentimental Analysis and Deep Learning. Advances in Intelligent Systems and Computing, vol 1408. Springer, Singapore. https://doi.org/10.1007/978-981-16-5157-1_78

[26] Gaur, P., Vashistha, S., Jha, P. (2023). Twitter Sentiment Analysis Using Naive Bayes-Based Machine Learning Technique. In: Shakya, S., Du, KL., Ntalianis, K. (eds) Sentiment Analysis and Deep Learning. Advances in Intelligent Systems and Computing, vol 1432. Springer, Singapore. https://doi.org/10.1007/978-981-19-5443-6_27

[27] P. Jha, D. Dembla and W. Dubey, "Implementation of Machine Learning Classification Algorithm Based on Ensemble Learning for Detection of Vegetable Crops Disease", International Journal of Advanced Computer Science and Applications, Vol. 15, No. 1, pp. 584-594, 2024.

[28] Bhatia, Pramod; Garg, Vivek et al. (Patent No: 20 2022 102 590.8) (2022) Intelligent seating system based on IoT and machine learning. .