Gender and Age Detection using Deep Learning

Utkarsha Kumbhar*, Prof. A. S. Shingare

Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

ABSTRACT

For the past few years, gender and age detection has been an active area of study and researchers have been putting a lot of effort to contribute quality research in this area. Starting from preprocessing of data to building a model which gives high precision results is tedious task for researchers. There is a immense dormant field of study as it can be used in monitoring, surveillance, human-computer interaction and security. However, there is still a lack of the performance of existing methods on real live images. Many difficult tasks such as computer vision, speech recognition, and natural language processing are easily solved with deep learning. Therefore, the approach of deep learning remarkably growing and this also takes place in image classification. Therefore, to analyses and focuses on comparative study of different algorithms for gender and age recognition system to give elevated degree of precision is required.

Keywords : CNN, Adience dataset, Feature extraction, neural network, deep learning

I. INTRODUCTION

In every sector, age and gender have played a significant role. It appears that a great deal of effort has been done in every department. Unreliable data was gathered from the media, healthcare, and FMCG (rapid moving consumer goods) industries. Due to today's broad market of consumers and job hopefuls, diversity appears to be more vital than ever. Employees have been pushed to reassess standard staffing and company methods as a result of the competitive economy. A varied workplace made up of men and women of all ages is extremely important in today's world for a variety of reasons.

In the current scenario, developing systems based on age and gender for various platforms such as mobile, applications that use social networks such as Facebook, Instagram, Snapchats, and various matrimonial and dating websites such as shadi.com, happen, tinder, bumble, and others is a critical task for human security concerns.

The important task is to use this data intelligently and to build helpful solutions for human needs. An intelligent algorithm must be designed to overcome this problem.

Deep learning is an artificial intelligence (AI) function that mimics the human brain's processing of data and pattern creation in order to make decisions.

Deep learning is an artificial intelligence subset of machine learning that uses neural networks to learn unsupervised from unstructured or unlabelled data. Deep neural learning or deep neural network are other terms for the same thing. Deep learning gives numerous ways and algorithms to solve the age and gender detection dilemma. Choosing which algorithms to use is a complex job. Data scarcity is also a major issue in determining gender and age.

Deep learning and machine learning can help solve a variety of problems. In this paper, a convolution neural network is utilised to extract features from the Audience dataset, and a few classifiers are utilised to improve accuracy.

II. RELATED WORK

There are numerous systems on the market that can determine gender and age using various technologies, approaches, and techniques. Deep Learning-based techniques are an increasingly prominent set of approaches used by software engineering academics to automate development chores (DL). The popularity of these techniques is due to their automated feature engineering capabilities, which helps in the modelling of software artefacts. However, because deep learning techniques are being adopted at such a quick rate, it is difficult to distil the current research landscape's successes, failures, and prospects. Deep learning (DL) is a machine learning field with a lot of potential.

The state-of-the-art studies in several scientific areas, including as computer vision, object recognition, speech recognition, and natural language processing, have led to the artificial intelligence waking up from its deep slumber in the recent decade. Many researchers are currently using DL approaches to try to solve a variety of challenges in various domains.

The deep learning technology has substantially increased the quality of picture categorization and object recognition over the last seven years. The essential components of a gender recognition system before applying deep learning methods are feature descriptor and classification approach. Many techniques have been used to extract discriminative features from facial images, which can be broadly classified into geometric and appearance-based methods. Geometric features such as the distance between eyes, the length of the eyes and ears, the length and width of the face, and so on have been used in the former category. On the other hand, in appearance-based approaches, the entire image is taken into account rather than just a portion of it.

For the encoding of facial images, Dyadic Wavelet Transform (DyWT) and Local Binary Pattern (LBP) were previously utilised. LBP is a state-of-the-art texture descriptor, and DyWT depicts a facial picture at multiple scales. DyWT decomposes an image into distinct sub-bands at different scales, making analysis simple. The DWT transform has been utilised for face description, however because it is not translation invariant, it does not have the best potential for feature extraction. DyWT is a superior choice for face description since it is translation invariant. LBP, on the other hand, is superior at capturing local detail. In a new method, DyWT and LBP are combined. For the challenge of gender recognition, we propose a description of facial images. Deep learning provides a technique to dissecting an image into subcategories rather than dissecting it into subcategories. Convolutional neural networks (CNNs) revolutionised the field of computer vision. It not only improves image classification accuracy with time, but it also helps with generic feature extraction such as scene categorization, object detection, image retrieval, image caption and semantic segmentation.

In image processing tasks, convolutional neural networks (CNNs) are one of the most powerful kinds

of deep neural networks. It's a powerful tool that's widely utilised in computer vision applications. Convolution layers, subsampling layers, and full connection layers are the three types of layers in a convolution neural network.

III. PRAPOSED METHOD AND MATERIALS

In proposed system, Adience dataset [15] is used for solving the problem. Adience benchmark, which is used to categorise people based on their age and gender. Images from smart-phone devices are automatically uploaded to Flickr in the Adience set. Because these photographs were posted without deliberate filtering, as is common on media webpages (e.g., photographs from the LFW collection [25]) or social websites (the Group Photos set [14]), viewing circumstances in these photographs are a little different.

As a result, Adience photos record extreme differences in head attitude, lighting quality, and more. There are around 26K photos in the Adience collection, with 2,284 subjects. The collection is divided by age groups. A typical five-fold, subject-exclusive cross-validation methodology [15] is used to test for both age and gender classification. These images are used more than recent alignment approaches to emphasise the performance improvement ascribed to network architecture rather than better pre-processing. Deep Learning is used to correctly determine a person's gender and age from a single photograph of their face. The anticipated gender and age ranges are (0-2), (4-6), (8-12) (15-20) (25-32) (38-43) (48-53) (60-100).

Deep learning, also known as deep neural networks or neural learning, is a type of artificial intelligence (AI) that aims to mimic the brain's functions. It's a type of machine learning that uses functions to make decisions in a nonlinear way. Deep learning occurs when decisions are made without supervision on unstructured data. Recognition of objects, speech recognition, and language translation are some of the tasks performed through deep learning.

Deep learning is a subset of machine learning that analyses data using hierarchical neural networks. Within these hierarchical neural networks, comparable to the human brain, neuron codes are linked together. The hierarchical nature of deep learning, unlike other standard linear programmes in machines, allows it to take a nonlinear approach, processing data over a number of layers, each of which will integrate subsequent tiers of additional information.

There are three different types of neural networks: ANN, RNN, and CNN. Each has its own importance in a number of situations. Feature engineering is a critical aspect in building a network-based model. Feature engineering consists of two steps: feature extraction and feature selection. A convolutional neural network (CNN) is employed in the proposed system to extract a feature from the dataset. CNN models are employed in a wide range of applications and domains, but they're especially common in image and video processing.
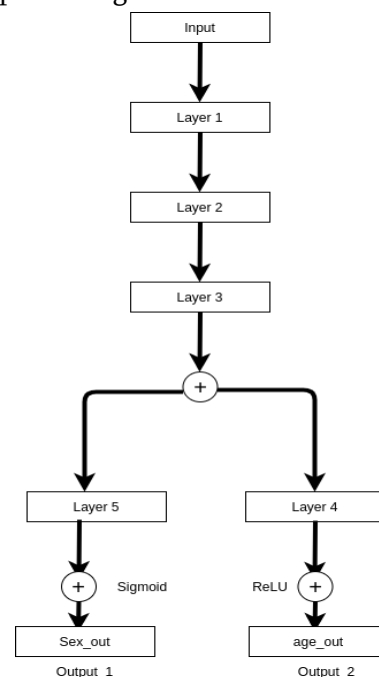

Figure 1. Proposed system neural architecture.

For both age and gender classification, we used our proposed network design throughout our studies. Only three convolutional layers and two fully-connected layers with a modest number of neurons make up the network. This is in compared to the much bigger structures used in [16] and [17]. Our decision to use a smaller network was influenced by both our goal to avoid overfitting and the nature of the problem. Figure 1 depicts our network architecture in its entirety. More information is available in the text. one of the issues we're trying to resolve: On the Adience set, age is classified into eight categories, but gender is only divided into two. In contrast to, for example, this is in contrast to the ten thousand identity classes used to train the facial recognition network in [48]. The network processes all three colour channels directly. Images are initially rescaled to 256 256 pixels and then sent to the network with a crop of 227 227 pixels. Following that, the three convolutional layers are defined as follows.

1. In the first convolutional layer, 96 filters with a size of 3707 pixels are applied to the input, followed by a rectified linear operator (ReLU), a max pooling layer with a maximum value of 3 3 regions with two-pixel strides, and a local response normalisation layer [28].

2. The second convolutional layer, which contains 256 filters of size 96 5 5 pixels, processes the preceding layer's 96 28 28 output. The same hyper parameters as before are used for ReLU, a max pooling layer, and a local response normalisation layer.

3. Finally, the third and final convolutional layer applies a set of 384 filters of size 256 3 3 pixels to the 256 14 14 blobs, followed by ReLU and a max pooling layer.

In the CNN feature extraction strategy, just the top layer of the network is trained, and the rest of the network remains constant. In this method, there are

a variety of instances to evaluate. If the new dataset is greater than the old one, the weight should be adjusted; this is the fine-tuning strategy.
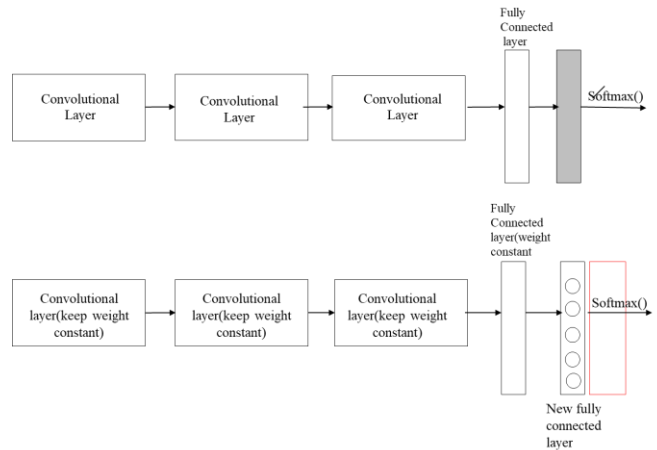
Figure 2. small data with exactly similar to original dataset.

In the figure.2, the transfer learning function is used on small data that are identical to the original dataset. When the targeted data is small, new fully connected layer is generated by keeping previous weights constants.

IV. RESULTS AND DISCUSSION

In the proposed system, CNN is employed to extract features, and various classifiers are employed to achieve high accuracy.

The proposed system is organised into three primary parts. Images are captured utilising a live webcam as well as images from a saved location in the first two modules. With the help of the OpenCV module, it is determined whether the image is male or female and what age range it belongs to after receiving it as an input.

Figure 3. Image captured from web cam

Because of elements such as cosmetics, lighting, obstacles, and facial expressions, determining an exact age from a single image is extremely difficult. As a result, rather than utilising regression to solve an issue, it is preferable to use classification. Classification is a task that necessitates the application of machine learning algorithms to learn how to assign a class label to problem domain instances. The term "man" or "female" was used here. In terms of modelling, classification necessitates a training dataset with a large number of examples of inputs and outputs from which to learn.

The results of classification predictive modelling algorithms are examined. Classification accuracy is a common metric for evaluating a model's performance based on projected class labels. Although classification accuracy isn't ideal, it's a solid place to start for a lot of classification problems.

**Figure 4.** Image getting from saved location

There are various classification algorithms in machine learning. SVM, MLP, Random Forest, decision tree Extratree, adaboost, gradient boosting classifiers are used to calculate accuracy for the proposed system.
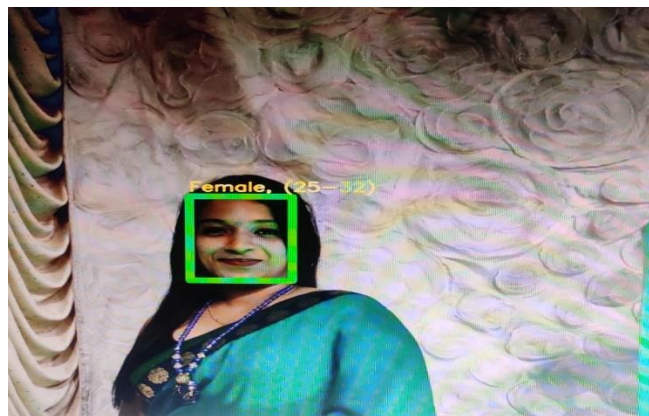
Figure 5. gender and age detection

TABLE I
ACCURACY OF DIFFERENT CLASSIFIERS

CLASSIFIER	ACCURACY
SVM	97.79%
DECISION TREE	97.00%
ADABOOST	97.16%
GRADIENT BOOSTING	97.003%
RANDOM FOREST	97.31%
MLP	97.03%
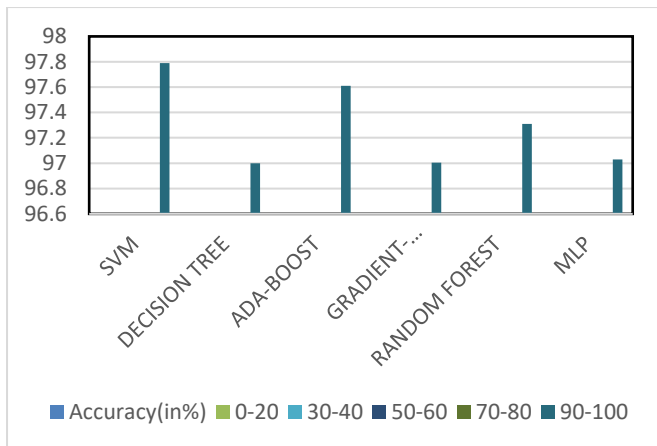
Figure 6. Graph for accuracy of different classifiers

Table 1 and figure 6 presents the results after applying different classifiers respectively. For the feature extraction CNN (figure 2) is used whereas accuracy is calculated by classifiers.

V. CONCLUSION

We presented the accuracy of many classifiers in this paper. CNN is primarily employed as a learning function for feature extraction, which allows it to determine age and gender. We use 5 hidden layers and 5-fold cross validation to produce a more accurate result.

Combination of feature extractions like frequency, and color with different classifiers will add accuracy to these applications.

VI. FUTURE SCOPE

The implementation can be used as exorbitant model and covers different field such as retail management sector, cosmetic industry and security area prominently. Face recognition is utilised in a variety of ways in retail, from security to ads. From sending personalised adverts to identifiable customers to gender identification. Also, Security, operational problems, and safety are all important considerations in retail venues, both indoors and out. With state-of-

the-art adaptive video analytics platforms, each camera can be turned into an alarm sensor, sending push notifications to a central monitoring station that can help cut expenses and enhance operations.

VII. REFERENCES

[1]. Arora, Shefali, and M. P. S. Bhatia, "A Robust Approach for Gender Recognition Using Deep Learning," In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6, 2018.

[2]. Levi, Gil, and Tal Hassner, "Age and gender classification using convolutional neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 34-42, 2015.

[3]. Wang, Xiaofeng, Azliza Mohd Ali, and Plamen Angelov, "Gender and age classification of human faces for automatic detection of anomalous human behaviour." In 2017 3rd IEEE International Conference on Cybernetics (CYBCONF), pp. 1-6, 2017.

[4]. Haseena, S., S. Bharathi, I. Padmapriya, and R. Lekhaa, "Deep Learning Based Approach for Gender Classification." In 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 1396-1399, 2018.

[5]. Ramdhani, B., Djamal, E.C. and Ilyas, R., "Convolutional Neural Networks Models for Facial Expression Recognition. In 2018 International Symposium on Advanced Intelligent Informatics (SAIN), pp. 96-101, August 2018.

[6]. Tiagrajah V. Janahiraman1 and Prasantth Subramaniam (2019) Gender Classification Based on Asian Faces using Deep Learning. 2019 IEEE 9th International Conference on

System Engineering and Technology (ICSET), 7 October 2019, Shah Alam, Malaysia

[7]. Ericsdotter C and Ericsson A M 2001 Gender differences in vowel duration in read Swedish: Preliminary results Working Papers - Lund University Department of Linguistics 34 – 37.

[8]. L. Best-Rowden, H. Han, C. Otto, B. F. Klare, A. K. Jain, Unconstrained face recognition: Identifying a person of interest from a media collection, Information Forensics and Security, IEEE Transactions on 9 (2014) 2144–2157.

[9]. J. Orozco, O. Rudovic, J. Gonz`alez, M. Pantic, Hierarchical on-line appearance-based tracking for 3d head pose, eyebrows, lips, eyelids and irises, Image and Vision Computing 32 (2014) 14–26

[10]. Arora, Shefali, and M. P. S. Bhatia, "A Robust Approach for Gender Recognition Using Deep Learning," In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6, 2018.

[11]. Levi, Gil, and Tal Hassner, "Age and gender classification using convolutional neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 34-42, 2015.

[12]. Wang, Xiaofeng, Azliza Mohd Ali, and Plamen Angelov, "Gender and age classification of human faces for automatic detection of anomalous human behaviour." In 2017 3rd IEEE International Conference on Cybernetics (CYBCONF), pp. 1-6, 2017.

[13]. Haseena, S., S. Bharathi, I. Padmapriya, and R. Lekhaa, "Deep Learning Based Approach for Gender Classification." In 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 1396-1399, 2018.

[14]. Ramdhani, B., Djamal, E.C. and Ilyas, R., "Convolutional Neural Networks Models for Facial Expression Recognition. In 2018 International Symposium on Advanced Intelligent Informatics (SAIN), pp. 96-101, August 2018.

[15]. E. Edinger, R. Enbar, and T. Hassner. Age and gender estimation of unfiltered faces. Trans. on Inform. Forensics and Security, 9(12), 2014

[16]. A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Neural Inform. Process. Syst., pages 1097–1105, 2012.

[17]. K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman. Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531, 2014

Cite this article as :