INTERNATIONAL RESEARCH JOURNAL OF ENGINEERING AND TECHNOLOGY (IRJET) e-ISSN: 2395-0056
VOLUME: 08 ISSUE: 12 | DEC 2021 WWW.IRJET.NET p-ISSN: 2395-0072

STUDY OF GENDER PREDICTION FROM IMAGES USING DEEP LEARNING TECHNIQUES

Priyanka Pal[1], Jaimala Jha[2]

[1,2]Madhav Institute of Technology & Science
[2]Professor, Department of Computer Science & Engineering, Gwalior, India

---***---

Abstract— Mechanized face estimation is an essential test task with a wide range of potential applications, particularly as social phases evolve and internet networking becomes more widespread. Hearty face confirming structures are in high demand to help overcome mental illness and trauma. Different apps give customer confirmation to provide a higher degree of security by gaining control of both physical and virtual places. For example, the development, beauty care items and eyeglasses, insurgencies, appearances, camera see centers. In this research the prediction of gender can be significantly increased by using deep learning for extracting information. The application of profound learning with deep learning methodologies has led to state-of-the-art performance. The gender image estimate is measured using extensive tests using Gender Labels - IMDB-WIKI dataset in the largest available public datasets for facial photographs.

Keywords— Gender Prediction, Deep Learning Techniques, MATLAB Etc.

I. INTRODUCTION

The person's face contains relatively large amounts of information and characteristics, including expression, ethnicity, gender and age[1]. For example, most people are able to discern human qualities such as gender and can tell whether the person is a male of a girl by merely looking at his/her face. They can also determine an individual's age and tell them whether they are a child or an adult. On the other hand, it is a challenge to build applications to identify people in their face and gather information from ages and gender, because of the necessity of creating a model that works for all human subjects, to computer visions, on which modernity depends on many important aspects of our everyday life [1]. Indees, computer vision and pattern recognition are the foundation of most automatic facial character classification systems. Computer vision comprises methods and techniques for comprehension, analysis and extraction of picture information. It is a science, in other words, that aims to construct a system that can show our reality and show it. In contrast, pattern recognition is a technology that allows robots to identify and detect patterns, which might be form, speech signal, fingerprint picture, handwritten words, environment or a human face [2], and to understand how they are recognized. The design of a pattern recognition system consists of three key components: preprocessing, extraction features and classification (see Figure 1). Preprocessing and collection of data include the selection and division into training sets, a suitable database, standardizing samples, isolating superfluous samples. In this phase, you may also make certain image alterations such as image dimension, segmentation, light modification or 2 other repairs. In terms of extracting functions, good characteristics are found, model categories and representation are defined, and the data by measuring these characteristics are reduced. In the classification portion, a classificatory splits space into areas and assigns a category to a pattern is very critical. It is highly critical. There is also a fourth opted component in the design of an evaluation designed to evaluate the performance of the system based on speed, accuracy and cost. [2]

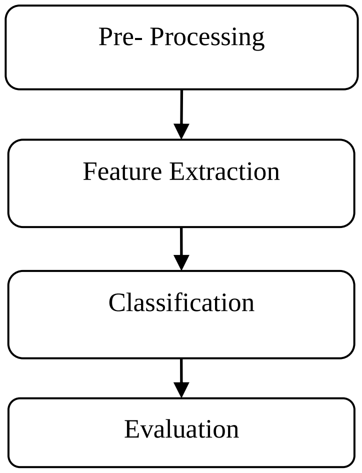

Figure 1. The Main Parts of a Pattern Recognition System[2]

Even though computer vision and model acknowledgement have been an active study field since the 1970s, age and gender prediction from facial photos are investigated more lately. Two visual difficulties are investigated in this thesis: Classification of gender and estimate of age[3]. Gender prediction from face photos is an appealing study issue and an important work for the computer, but the need for and present performance are still lacking. The variety in lighting, resolution, expression, pose etc. makes this deficit widespread. Face images, in turn, can increase the performance of a wide range of applications including the interaction of the human and computer, customer information measurement and access control. It may also have a significant impact on a number of areas, including security systems, biometric authentication, medical imaging systems, surveillance systems and interfaces, content based 3 and demographic research [3]. [3]. Most of the available solutions depend only on face classification algorithms in order to resolve the issue of gender prediction. These algorithms may usually be used on conventional databases with aligned frontal faces with good resolution. On the other hand, it is more difficult to prediction of an individual's facial image than to anticipate his or her gender due to the wide diversity of face appearance such as humanity's variety, stances and facial expressions [3]. Gender prediction is very beneficial in several applications, such as demographic profiling, forensic technology, gender-specific interfaces for human computing, security checks, gender-based advertising and Electronic Customer Relationship Management (ECRM)[3].

II. GENDER PREDICTION

It takes multiple processes to predict gender from the picture. We must first detect and remove the face from the image. We can also recognize facial traits like eyes, mouth and nose and align the discovered face with them. We next feed through our model the detected and aligned face to generate the prediction.

1) Face detection

Faces on pictures are identified by the computer in vision detection. For uncontrolled face detection, there are several ways. Some previous systems depend on the edges of the face being detected. The algorithm most frequently used [3]. The method separates the image input into rectangular areas. Each section passes through a cascade of weak classifiers which detect the existence of simple hair-like characteristics. The hairlike characteristic is the difference in the sum of pixel intensity in several neighboring locations. The segment is recognized as containing face when it passes through all the phases of the waterfall, otherwise it is refused to do so. For varying sizes of rectangular portions, this process is repeated. AdaBoost is used for the classification. The key advantage of the algorithm is that the time required for calculation of hairlike characteristics is constant owing to the usage of integral pictures. N an integral picture or a summary area table is a data structure in which the total of the cells above and left for each cell represents the value. In a single pass, the summed area table can be calculated. In many famous libraries like OpenCV1, hair cascade face detection is employed Another approach uses HO G or a descriptor histogram of oriented gradients. It was first invented[2] to detect individuals. The pixel intensity gradient for each pixel is calculated from the algorithm. Next, the picture is divided into smaller areas. A histogram of the steps is constructed in each location. Only the leading gradient is thus saved[3].

Multitask Cascading Coevolutionary Neural Networks (MTCNN), presented at first, is used as a profound learning approach for face detection. In one pass, the system detects facial and face points. It uses a cascade of neural networks. The image of the input first changes to make a pyramid of the picture. A multi-scale depiction of an image is an image pyramid. It allows you to find items at various scales in photos. Often paired with the sliding window, the objects can be found in many places. The pyramid is then feed into the Proposition Network, the first coevolutionary neural network (P-Net). The candidate bounding boxes [4] are found in this network. The candidates are the entries for the next network known as Refine Network (R-Net). Bound boxes are fused with non-maximum deletion at both stages. Non-maximum removal is a strategy used to remove unnecessary edges. It deletes all but the local maximum gradient values. Last stage is refined and basic facial location is carried out by the Output Network (O-Net). The MTCN N will only locate 5 points of reference, compared to other systems of facial detection: both eyes, both mouth and nose tips [4].

The gender and classification problem is directly connected with face detection. GENDERNET was established as one of the earliest methods for gender rating. The samples of 30 x 30 pixels were categorized by a completely linked neural network, consisting of 3 layers of 900, 40 and 900 neurons. Another approach used to classify gender is the Support Vector Machines. The process uses a 21 x 21 pixel sampled image and classifies it using an SVM with a radial function kernel. The issue of gender classification was recently dealt with as a gender prediction sub-problem. For gender prediction, current approaches are using the same design of the neural network. Gender is a binary classification problem that may readily be described. If gender is patterned using regression, the resulting number can be viewed as trust [5].

2) Facial Landmarks Detection

Detecting facial markings is a key step in aligning the face. The facial sightings of facial structures such as nose, eyebrows, eyes, mouth and jaw are employed to locate. A very rapidly created algorithm is used in the Dlib library. It uses a regression cascade in which each tree updates a facial mark co-ordinates vector. Each tree node is selected based on the pixel intensity differences. There was a n alternate method that employs profound learning. They are creating a new neural network called the Face Alignment Network (FAN). The network comprises four stacked 'Hourglass networks' with a bottleneck block updated. It invented the Hourglass network. It is designed for the estimate of human poses. A network of Hourglass reduces the image down to a very low resolution and then combines its characteristics in different resolution [5].

III. DEEP LEARNING

The growth of high-performance computing facilities made deep education techniques that employ deep neural networks popular. The capacity to analyze a huge number of characteristics in unstructured data ensures a high level of power and flexibility in profound learning. The data are transmitted via numerous layers of a deep learning algorithm; each layer can extract the features gradually and transfer them to the next layer. The first layers remove low level characteristics while the subsequent levels combine characteristics to form a full depiction [6].

1) Evolution of Deep Learning

Artificial Neural Networks'(ANN) First generation consisted of neural layer perceptron's that were limited in computing. The error rate was determined for the second generation and the error reported. Limited Boltzmann machine has overcome the background propagation restriction, which facilitated learning. Then there will be the development of other networks [6]. The calendar shows the development and the traditional model of deep models. The performance of deep learning classifiers improves on a big scale and an increasing amount of data compared with standard methods of learning. It shows how typical algorithms of machinery learning and profound study algorithms perform [6]. When the threshold of the training data is reached, typical machine learning algorithms are steady, while deep learning increases their performance with an increased data volume.

2) Deep Learning Approaches

In supervised learning, unattended learning, strengthened learning and hybrid learning, deep neural networks succeed.

a) Supervised Learning

The input variables represented by X in supervised learning are mapped to output variables represented by Y by means of an algorithm for mapping.

$$Y = f(X) \quad .....1.1$$

The objective of the study algorithm is to approximate the mapping capability to forecast a new input output (Y) (X). To correct the output, the error of the trained predictions can be employed. When all inputs are trained to achieve the specified result[6], learning can be stopped. Regression for the resolution of problems with regression [6], Classification support vector machines [6], Classification random forest, and regression problems [6].

b) Unsupervised Learning

We have only the input data and no corresponding output to the map in an unchecked learning. The purpose of this learning is to learn about data using data modelling. The intriguing structure present in the data may be found in the algorithms. Uncontrolled learning uses clustering problems and association problems. Included in the difficulties of clustering are the unattended learning algorithms, for example K-means algorithm [6].

c) Reinforcement Learning

Increased learning uses an algorithmic education system of reward and punishment. The algorithm or agent will learn from its surroundings in this way. The agent is rewarded for good performance and for wrong performance. In the case of an automatic car, for example, the agent is awarded a reward for safe transport to the destination and a penalty for off-road transport. Likewise, the reward state can win the game and the penalty for checking in the case of a software for the play of chess. The agent attempts to maximize the price and reduce the fine. The algorithm is not explained in reinforcement learning how the learning can be carried out; yet, it operates by its own means [7].

d) Hybrid Learning

Architectures that use generative (unsupervised) and discriminatory (supervised) parts are hybrid learning. A hybrid deep neural network can be designed using the integration of multiple architectures. They are used to recognize people with action bank characteristics and are projected to yield far better results [6].

IV. LITERATURE SURVEY

Alex Darborg [2020] Alex Darborg the technique of identifying and verified people in the photograph by their face is typically defined as gender prediction. Recently, researchers have paid more attention to this area and improved the basic models regularly. The aim of this project is to implement a one-shot learning system for real-time facial recognition. "One shot" indicates to learn from one or a few samples of training. This study examines various ways of solving this issue. In order to achieve acceptable accuracy, converging neural networks know to require extensive datasets. This study presents a means of resolving this difficulty by limiting to one the number of training examples and by applying the concept of transference learning[7] to achieve an accuracy of almost 100%.

Ramalakshmi K [2020] Recognition of gender is a process of deeply learned recognition of the gender in a person's face. Some elements that affect the recognition of faces include variety, lighting and occlusion. These are minimized by improved predictive accuracy.

Coevolutionary neural network is the network used for training the system (CNN). The faces are recognized and cut from the image to improve their accuracy. Open CV is applied to recognize the face by the frontal features of the face. Face detection This takes place during network training. Included photos are the data set utilized for the training. Without compromise, the suggested approach forecasts the genre of the individual [8]

Mandar Khatavkar [2020] Supriya Khatavkar Face discovery, recognition and gender evaluation are one of the largest investigation areas in PC vision, not only because they are tested as an item, but because of countless applications that require facial recognition, follow-up and recognition. While many critical types of exams on facial recognition, recognition, and gender evaluation issues have been performed independently over the last couple of years, a continuous video for individually recognizable evidence does not combine any specific examinations of facial identification, recognition and gender assessment. We feel that this kind of massive scrutiny should work in this way. The key obligations of our work are segregated into three areas, in particular on the face of individual recognized evidence, recognition and genderual orientation. We use the Histogram (LBPH) technique and the Convolution Neural Network (CNN) as a Locally Binary Beispiel in our research in order to remove the facial highlights of face images whose computerization is low. By measuring the neighboring pixels and convolution levels of local dual patterns histogram (LBPH), we remove a strong facial component for recognition of face recognition and gender evaluation. We demonstrate the exploratory results for perceiving face and genderual orientation for individual identity[9] using these strategies.

Chaturvedi Prague [2020] Prague Lung cancer has emerged as one of the most reported diseases in all people. It is also one of the most often reported cancer deaths. Lung cancer cases are rapidly growing. In India there are approximately 70,000 instances every year. In its first stages the condition tends to be asymptomatic, making it almost impossible to identify. Therefore, early identification of cancer plays a vital role in life-saving. A patient can have a better chance of recovery from an early detection. In efficient cancer detection, technology plays a key role. Many scientists have proposed various ways based on their studies. Recently, numerous CAD techniques and system have been presented, developed and developed to handle this problem with the use of computer technology. The systems use several techniques for machine learning and deeper learning, and multiple methods have been developed to forecast malignant malignancies based on image processing techniques. The focus of this study is on listing, discussing, comparing and analyzing a number of approaches for picture segmentation, extraction of characteristics, as well as numerous methods for early stage classification and detection of lung cancer[10].

Be. Matus Námesny [2019] picture gender prediction is a major computer vision application. Many ways of solving this challenge are available. Three distinct strategies have been evaluated. We use a wide set of available public data sets and one manually labeled dataset to train the best technique. We expand data further by adding an image color channel and training the optimal way. We show that a network training with a big data set enhances performance, but it does not improve additional color channels. Based on our findings, we design a gender classification application[11].

Unconstrained real-world face photos are classified by **Olatunbosun Agbo-Ajala et al. [2020]** gender predictions in non-filtered faces as predetermined gender. In this field of research significant advancements were produced since it is useful for smart applications in the real world. However, the usual methods for unfiltered benchmarks demonstrate their inability to deal with substantial degrees of variation. Recently, approaches based on Convolutional Neural Network (CNNs) for classification were widely adopted, due to their excellent facial analysis ability. We propose a new end-to-end CNN strategy in this piece, to ensure that unfiltered realities become resilient and gendered. The two-tier CNN architecture comprises extraction and classification of features themselves. The extraction feature extracts gender feature, and the classification classifies the face images to the right age and gender. Especially with a solid pretreatment picture technique, which prepares and processes those faces before being fed into the CND model, we discuss the enormous variances in the unfiltered realm. Our network is technically pre-trained using a noisy IMDb-WIKI, then tailored to the MORPH-II and ultimately to the OIU (original) dataset training set. The results reveal that our model achieves the state-of-the-art performance in the genre classification when analyzed for the accuracy of classification using the same OIU additive criterion. It enhances in group classification the best published results by 16.6% (exact accuracy) and 3.2% (one-off accuracy), and also increases gender categorization to 3.0% (exact accuracy) [12].

V. RESULT DISCUSSION

- The data set WIKI and IMDB considered to have gender data pictures of Age and Gender.
- We extract images from both datasets using the WIKI. mat and IMDB. mat file format.
- Specify the male and female different photos in a separate directory with Final.csv and opencv.
- Now prepare pictures with the necessary resolution and delete defective photographs (32x32x3).

- Male and female format of 15676 pictures*.jpg are total photos.
- Load image Datastore data and number for instance two classes for each label (male and female).
- Split data for training and testing into 80 percent and 20 percent and randomize it.
- In our research we utilize VGG16 and SVM with transfer education to define neural network architecture.
- Resnet50 weight and multi-class SVM for training use with transfer learning.
- We employ 10 convolution layer with neuron in neural network design (128,128,128,64,32,32,16,32,64,64), a max-pooling layer, a drop-out layer (0.3), a relu activation function, a similar padding function, three FCN layers (500,100,2).
- Optimizer training parameters ADAM, gpugtx usage 1050ti, learning rate 0.0001, factor drop 0.3, epoch 100, batch size 128.
- Finally train and save the neural network to legendernet.mat.
- 13, 32 gigabytes of memory, core i5 10, 240 gbssd, and gpu is used for GTX 1050TI system configurations.

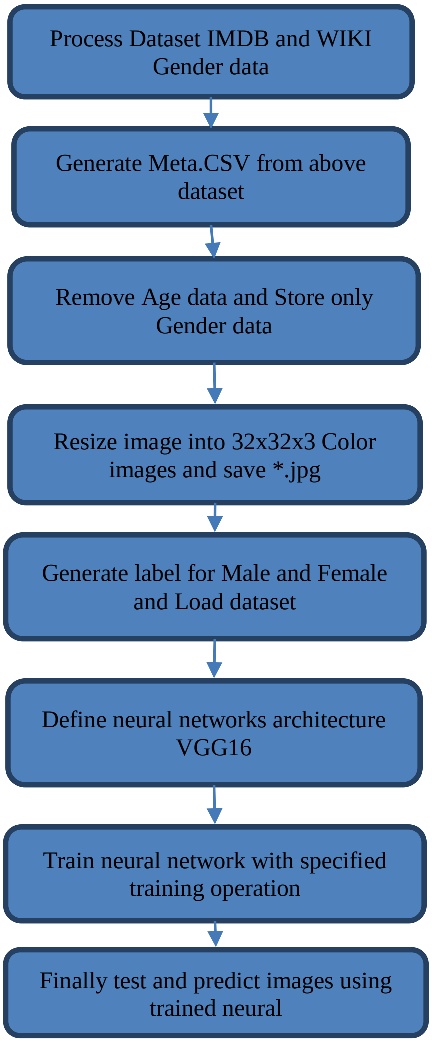

Figure 2: Flow chart

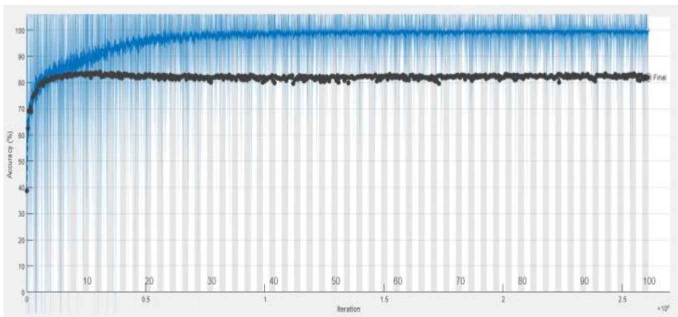

Figure 3: Training and Validation Accuracy

Training accuracy refers to the trained model identifying independent images that were not used in training, whereas test accuracy refers to the trained model identifying independent images that were not used in training. The percentage of anticipated values (Y_Predict) that match actual values is calculated by accuracy (Y_True). If the anticipated value matches the actual value, the record is called correct. We then calculate Accuracy by dividing the number of accurately predicted records by the total number of records.

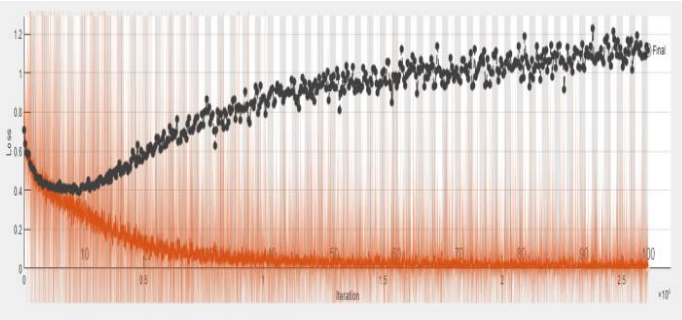

Figure 1.4 Training and Validation Loss

A learning curve is a graph of model learning performance as a function of time or experience. In machine learning, learning curves are a common diagnostic tool for algorithms that learn progressively from a training dataset. After each update during training, the model can be tested on the training dataset and a holdout validation dataset, and graphs of the measured performance can be constructed to display learning curves. Examining model learning curves during training can help diagnose learning issues, such as an under fit or Over fit model, as well as whether the training and validation datasets are sufficiently representative. In this post, you'll learn about learning curves and how to use them to diagnose machine learning models' learning and generalization behavior, with examples of plots illustrating common learning issues.

Base result

Table 1: Base result

Results	Mini-batch Accuracy	Validation Accuracy	Mini-batch Loss	Validation Loss
Base result	84.38%	78.32%	0.3247	0.5432

Classes	Precision	Specificity	Sensitivity	Recall	F-score
Male	0.7283	0.8498	0.6721	0.6721	0.6991
Female	0.8122	0.6721	0.8498	0.8498	0.8306

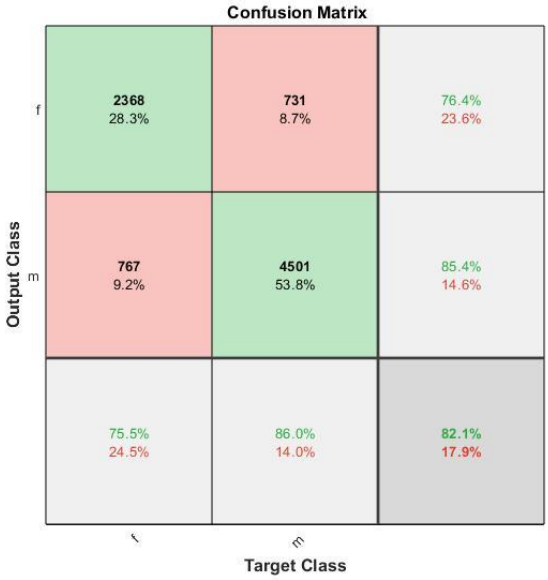

Figure 4: Confusion Matrix of Male and Female

An N x N matrix is used to evaluate the performance of a classification model, where N is the number of target classes. The matrix compares the actual goal values to the machine learning model's predictions. This provides us with a comprehensive picture of how well our classification model is working and the types of errors it makes. For a binary classification task because we have only two classes, we'd use a 2 x 2 matrix with four values, as seen below:

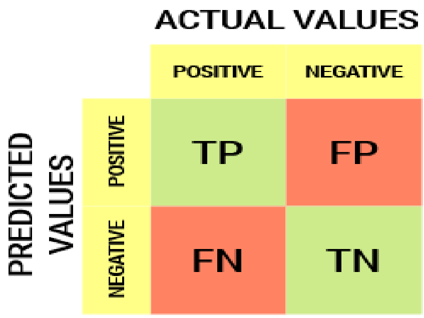

Figure 5: Confusion Matrix of Male and Female

There are two possible values for the target variable: positive or negative. The target variable's real values are represented in the columns. The rows represent the target variable's expected values.

Propose Results using VGG16

Table 2: Propose VGG16

Results	Mini-batch Accuracy	Validation Accuracy	Mini-batch Loss	Validation Loss
Propose VGG16	**97.66%**	78.70%	0.0395	0.2484

Classes	Precision	Specificity	sensitivity	Recall	F-score
Male	0.7641	0.8603	0.7553	0.7553	0.7597
Female	0.8544	0.7553	0.8603	0.8603	0.8573

Table 3: SVM with transfer learning result

Classes	Accuracy	Precision	Specificity	Sensitivity	Recall	F-score
Male	0.7525	0.7304	0.8566	0.6484	0.6484	0.6870
Female	0.7525	0.8026	0.6484	0.8566	0.8566	0.8287

One of the most often used supervised machine learning approaches is the Support Vector Machine (SVM). It can be used for classification and regression. This notebook focuses on classification. In its fundamental form, SVM is a linear classifier. It establishes a decision boundary as a hyperplane. To deal with nonlinear instances, nonlinear kernels are used. Data is translated into a higher-dimensional space and is intended to be linearly separable in this manner. The kernels in a scikit-learn implementation can be linear, poly, rbf, sigmoid, or custom.

Predicted images

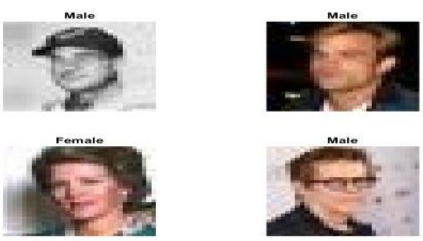

Figure 6: Predicted images

Figure 6 show gender classification system determines a person's gender (male/female) based on the face of the individual in each image. Many other applications, such as facial recognition and smart human-computer interface, can benefit from a good gender classification approach.

VI. CONCLUSION

Since many earlier methods tackled gender classification challenges, much of this work has been centered on restricted photos taken in laboratory contexts until recently. Such settings fail to reflect accurately the appearance changes on social sites and online repositories typical to real-world photos. However, internet images aren't just more difficult, they are numerous, too. Modern systems with efficient limitless training data allow easy availability of vast photographic collections, however these data are not always adequately designated for supervised learning. With limited etiquette, we provide outcomes with a lean deep learning architecture to avoid over fitness. In comparison with certain current network architectures, our network is "shallow," lowering the number of parameters and overlap. The results are tested against the existing unfiltered image benchmark. First of all, profound education can be employed to produce better outcomes in the classification of gender, even in view of the far smaller spectrum of modern unrestricted images. Secondly, the simplicity of our models implies the ability to significantly improve outcomes by using more training data. Our main focus is on decreasing computing power and enhancing system accuracy. Possibility for the face image processing work, gender classification, etc. to get the latest results. The results showed that the neural networks are strong. By allowing deeper architecture to be constructed, the availability of higher hardware and larger data sets can considerably improve the outcomes.

REFERENCES

1) Mirza Mohtashim Alam "Gender Detection from Frontal Face Images" School of Engineering and Computer Science Department of Computer Science & Engineering, 2015.
2) AJITA RATTANI "CONVOLUTIONAL NEURAL NETWORKS FOR GENDER PREDICTION FROM SMARTPHONE-BASED OCULAR IMAGES" 19 FEBRUARY 2018.
3) Amitha Mathew1, P.Amudha2 and S.Sivakumari "Deep Learning Techniques: An Overview" Department of Computer Science and Engineering, School of Engineering, 2016.
4) Erhan Sezerer, Ozan Polatbilek, Özge Sevgili, and Selma Tekir "Gender Prediction from Tweets with Convolutional Neural Networks" 2018.
5) Amirali Abdolrashidi, Mehdi Minaei, Elham Azimi, Shervin Minaee4" Age and Gender Prediction From Face Images Using Attentional Convolutional Network" 2017.

6) Sebastian Lapuschkin “Understanding and Comparing Deep Neural Networks for Age and Gender Classification” 2017.
7) Alex Darborg “Real-time face recognition using one-shot learning A deep learning and machine learning project” 2020, Department of Information and Communication Systems.
8) Ramalakshmi K , T. Jemima Jebaseeli , Venkatesan R3 “Prediction of Gender from Facial Image Using Deep Learning Techniques” Vol.-15, No.-2, February (2020) pp 118-128 ISSN (Print) 0973-8975.
9) Supriya Mandar Khatavkar “Gender And Age Detection Using Deep Learning Techniques” International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 09 | Sep 2020
10) Pragya Chaturvedi1 , Anuj Jhamb1 , Meet Vanani1 and Varsha Nemade1 “Prediction and Classification of Lung Cancer Using Machine Learning Techniques” ASCI-2020 IOP Conf. Series: Materials Science and Engineering 1099 (2020) 012059 IOP Publishing doi:10.1088/1757-899X/1099/1/012059
11) Be. Matus Námesny “Gender and Age Classification in Camera Data” Brno, Spring 2019
12) Mohamed Hamed N. Taha, Aboul Ella Hassanien, and Hamed Nasr Eldin T. Mohamed "Deep Iris: Deep Learning for Gender Classification Through Iris Patterns” 2019.