Gender Prediction using Deep Learning Algorithms and Model on Images of an Individual

G Rajendra¹, Kodali Sumanth², Chandragiri Anjali³, Gudavalli Pardhasai⁴, Madhumanchi Supraja⁵

Assistant Professor, Department of IT, Lakireddy Bali Reddy College of Engineering, Mylavaram-521230, India.

Students MCA Department, Lakireddy Bali Reddy College of Engineering, Mylavaram-521230, India³,⁴,⁵.

¹rajendra.raju@gmail.com,²kodalisumanth143@gmail.com,³chandragirianju@gmail.com,⁴gudavalli98@gmail.com,⁵suprajamadhumanchi143@gmail.com.

Abstract: The classification of gender based on the biometric is an old and traditional approach which are treated as a sub-branch of “Soft Computing”. But in this proposed work, works on the celebrity images dataset obtained from the Kaggle repository. The major focus of this to extract the important features from the images by training the neural network by using an autoencoder and decoder. The system extracts the feature maps to visualize the extracted features and makes some sort of interpretations to analyze the data. Gender classification using iris information is a relatively old technique, and most gender classification methods classify all iris texture features. These features are fed to the classifier, this data might be proper or improper, and if it is improper it leads to more generalization results in less accuracy. The proposed algorithm to create increase the size of the dataset implements the Data Augmentation Technique using basic image operations. Then it applies a convolution neural network for predicting the gender of the user. The model has proved practically that increasing the size of the dataset and reducing the features improves the classifiers learning performance.

Keywords: Feature Extraction, Data Augmentation, Convolution Neural Network, Feature Maps.

1. Introduction

The traditional mechanisms stated that with the analysis of iris data, the system can successfully predict the gender of a person with high accuracy rates than the normal image processing techniques, which takes a lot of time to build the model. These types of applications are really important in places where gender selection [1] plays a prominent role. Suppose an organization wants to celebrate women’s day for its employees where the entry of men is strictly restricted. So, at the entry points, detectors [2] are arranged to allow the persons based on their gender. So, to extract the attributes based on iris, the analysis of iris texture is performed and by obtained the iris data from sensors. In the last decades, image processing combined with machine learning algorithms are used to build solutions for these types of applications by segmenting the image into various parts, and from among them, the system extracts the region of interest by applying the necessary filters and normalizing the matrix values obtained. But with the advancement of artificial intelligence, computer vision, and deep learning techniques performing all these actions automatically in the name of neural network layers.
2. Related Work

In [3] Tapia, J suggested an unsupervised method because the dataset contains the iris information without having any labels attached to it. The images are stored in the encrypted to provide security to the iris information, which is very sensitive information but the model finds it very difficult to predict the gender-based on this information. So, it designed a DBN algorithm to divide the images into checked and unchecked groups based on their similarities, which takes the help of the soft biometrics concept. Then the model is compared with a supervised neural model known as “Lenet-5”, by identifying the Normalized NIR iris images. Then it found that unsupervised methods combined with soft computing give better performance than the traditional neural networks.

In [4] Argyris, Y et al explored the importance of visual congruence in social media applications. In this era, to gain popularity and to become overnight stars people are filling their storylines and timelines with attractive videos and photos. This model finds the similarities based on the context, which is the strong driving force that makes people connect with others without even knowing their faces and other details. Most of the time, influencers on social media post either images or images with quotes without writing any description or including any hashtags. So, this model tried to find the patterns based on the themes posted by the influencers and their followers. From these extracted patterns, the model has expanded this system by considering the actuators that explain the process of their interconnection. Then to identify the key factor for this connectivity it uses the concept of social influence, which works on the propagating and behavioural aspects of the influencers. To classify the visualize the themes; it considers fine-tuning of three proposed CNN models to convert into transfer learning. Using the statistical measures and operational measures it counts the follower’s and influencers’ engagement, visual convergence, and controls with their measurements.

In [5] Haider, K et al proposed a deep learning network integrated with big data to handle this huge amount of images. This model designed an app, which identifies gender-based on facial points. This mainly works with uncontrolled conditions and designs an efficient algorithm to handle the heavy load in the smartphone during the retrieving of images from the cloud. A computer vision technology handles lighting conditions to overcome any darkening and blur effects. A new convolution neural network “Revolutionary Neural Network”, using ImageNet as a pre-trained model and finally compared it with a backpropagation network. These align the recognized faces by converting them into local binary patterns. These can also handle the 3D-aligned faces by defining the huge factors for the neural network.

In [6] Thomas, V et al proposed the prediction model based on the iris image. An ethnicity-based AdaBoost algorithm is proposed and obtained an accuracy of 85.95% but the precision values are very high, which states that the true prediction values are correctly classified. This model combines the two datasets one with Asian and another with Non-Asian, it also improves the resolution of images which are poor to enhance the capability of the model. It first designed an ensemble model that combines RBF, DT, and SVM to compare the front camera face images. To obtain the accuracy of 96%, employ a group of RBF networks and inductive decision trees and also use ICA and Support Vector Machines to match face images for gender identification of gender in the forefront face images.

In [7] Singh, M. et al designed a prototype based on the unknown feature learning mechanism especially while it combines the emerging features with the help of deep learning techniques. These types of architectures are very popular for handling the limited size data and very biometrics, iris, and other soft computing techniques fails. Even they provide security for the unsecured data by combining popular encryption techniques. An architecture is designed for handling the unlabelled features and autoencoders are implemented for handling the labelled data. The GANS are one of the important topics in this module which can increase the size of the data with the help of a generator and
discriminator module, which is a common technique for the applications where real-time images cannot be captured but still the system needs to increase its size.

3. Proposed Methodology

The proposed algorithm works on the celebrity dataset obtained from the Kaggle repository, which contains both images and a CSV file, which has 40 features to predict the gender of the celebrity. The proposed paper works on images to predict males or females. The algorithm consists of three modules, which are illustrated in the below sections. A sample image from the dataset is displayed in figure 1.

![Figure 1: Image from Celebrity Dataset](image-url)

3.1. Data Augmentation

All the existing systems suffered from limitation of data, which raises overfitting problems. So, the proposed algorithm implemented basic image manipulation techniques like rotation, scaling, transformation, zooming, rescaling, and other operations to create new images. Data augmentation can be performed either online or offline. But the offline mode is very expensive and as well as it consumes more space. So, in this paper, the model has chosen online mode to reduce the cost of the model. The manipulation values are represented in table 1.

| S.No | Operation Name       | Operation Value       |
|------|----------------------|-----------------------|
| 1    | Rescaling            | 1/255                 |
| 2    | Rotation             | 40 degrees            |
| 3    | Width and height shift| 0.2                   |
| 4    | Shear                | 0.3                   |
| 5    | Zoom                 | 0.2                   |
| 6    | Horizontal Flip      | Enabled               |

These operations create new images to handle situations in real-time, where the quality of the image is poor due to blur or noisy data. Samples of five images are displayed as output in figure 2.
3.2 Feature Map

The results obtained from CNN are automatic and are generated by the pre-trained models but they don’t clearly explain the way how and why they are chosen. One of the important layers in neural networks is “Feature Extraction”. To extract the features from the images and to display the selected features clearly, CNN implements the concept of filter techniques, which helps in the process of feature mapping. This process initially starts with getting the summary of the neural network model. The layers can be accessed to know their properties like weight, to extract the shape of the image. The filters obtained by executing the model can be illustrated in figure 3, which consists of 4 rows and 12 columns. The rows represent filter values and columns represent the channel value. To visualize the feature maps, the activation maps play a key role in identifying the preserved features among the given input image.

The feature map acts as the output of the filter which is obtained from the previous layer. During this phase, the filters cover every pixel one at a time, this makes the neurons at this point get activated. All these activated neurons are collected as “Feature Map”. The output of the feature map after the first activation layer is represented in figure 4.
3.3 Convolutional Neural Network

The proposed algorithm contains a neural network with one two-dimensional convolutional layer, one- two hidden layers, and one dense output layer. During the process of classification, initially, random weights are assigned to the input layer. The images which are taken as input has high dimensions, to reduce their dimensionality, all the layers are interconnected with the max-pooling layer. This layer helps in down sampling the images. A drop-out layer is added to the network after input and hidden layers. This addition will not only prevent the overfitting problem but also makes the network learn the robust features which help in generating more random subsets. It increases the number of iterations to train the model but for each epoch, the time taken to train the model becomes less. To convert the generated vectors of the image into a one-dimensional vector, a flatten layer is included before the output layer. The last layer of the network is a fully connected layer (or) dense layer, which gets the inputs from all the previous layers. The block diagram of the proposed system is shown in figure 5.

4. Experimental Results

The performance of the neural network is measured as the accuracy of the training and test data as well as the loss value of the training and test dataset. To view these measures simultaneously, the neural network provides a predefined function known as “history()” and it is plotted as line graphs. The visualization of history parameters for the proposed algorithm is represented in figure 6.
The proposed algorithm uses a sequential model to generate the output which is understandable to the user by using the summary, a predefined function that holds the information about the four components. The first component represents the name of the layers in sequence, the second component represents the shape of each layer after training, the third component weights in each layer and the fourth component describes the total weights in the model. The entire summary for the created model is described in figure 7.

| Layer (type)                  | Output Shape    | Param #  |
|-------------------------------|----------------|----------|
| conv2d (Conv2D)               | (None, 150, 150, 16) | 448      |
| max_pooling2d (MaxPooling2D)  | (None, 75, 75, 16)    | 0        |
| dropout (Dropout)             | (None, 75, 75, 16)    | 0        |
| conv2d_1 (Conv2D)             | (None, 75, 75, 16)    | 4640     |
| max_pooling2d_1 (MaxPooling2) | (None, 75, 75, 16)    | 0        |
| dropout_1 (Dropout)           | (None, 75, 75, 16)    | 0        |
| conv2d_2 (Conv2D)             | (None, 75, 75, 16)    | 4640     |
| max_pooling2d_2 (MaxPooling2) | (None, 18, 18, 64)    | 0        |
| dropout_2 (Dropout)           | (None, 18, 18, 64)    | 0        |
| flatten (Flatten)             | (None, 20736)       | 0        |
| dense (Dense)                 | (None, 512)         | 10617344 |
| dropout_3 (Dropout)           | (None, 512)         | 0        |
| dense_1 (Dense)               | (None, 1)           | 513      |

Figure 7: Summary of the Neural Network Model

The proposed model uses different special layers like maxpooling to reduce the problems due to normalization. Dropout layer to overcome the problem of overfitting. The results in figure 7 explains the number of parameters are reduced from 200,45,253 to 10,641,441, which in turn helps to reduce the complexity of the model.
Conclusion:

The proposed algorithm implements a simple and novel convolutional network whose accuracy has improved better than the existing systems. The designed algorithm tries to predict the class labels using 3 important features namely brown hair, smiling, and wearing glasses. The reason for the proposed algorithm to choose a deep learning algorithm than the machine learning algorithms is the generation of the categorical class label from an input image is efficiently obtained by the neural networks. The limitation of the proposed system is limited training. The overall training of the model can be improved either by using the pre-trained models like ImageNet, Alex Net, Google Net, and others or by using the data augmentation techniques which has two phases of implementation, generator to create new images and discriminator to distinguish between the real image and created images during the training phase.

References:

[1] Kowsari K., Heidarysafa M., Odukoya T., Potter P., Barnes L.E., Brown D.E. (2021) Gender Detection on Social Networks Using Ensemble Deep Learning. In: Arai K., Kapoor S., Bhatia R. (eds) Proceedings of the Future Technologies Conference (FTC) 2020, Volume 1. FTC 2020. Advances in Intelligent Systems and Computing, vol 1288. Springer, Cham. https://doi.org/10.1007/978-3-030-63128-4_26

[2] Tunç A., Taşdemir S., Köklü M. (2021) Fuzzy Based Noise Removal, Age Group and Gender Prediction with CNN. In: Kahraman C., Cevik Onar S., Oztaysi B., Sari L., Cebi S., Tolga A. (eds) Intelligent and Fuzzy Techniques: Smart and Innovative Solutions. INFUS 2020. Advances in Intelligent Systems and Computing, vol 1197. Springer, Cham. https://doi.org/10.1007/978-3-030-51156-2

[3] Tapia, J., & Aravena, C. (2017). Gender Classification from NIR Iris Images Using Deep Learning. Advances in Computer Vision and Pattern Recognition, 219–239. doi:10.1007/978-3-319-61657-5_9

[4] Argyris, Y. A., Wang, Z., Kim, Y., & Yin, Z. (2020). The effects of visual congruence on increasing consumers’ brand engagement: An empirical investigation of influencer marketing on Instagram using deep-learning algorithms for automatic image classification. Computers in Human Behaviour, 106443. doi: 10.1016/j.chb.2020.106443.

[5] Haider, K. Z., Malik, K. R., Khalid, S., Nawaz, T., & Jabbar, S. (2017). Deep gender: real-time gender classification using deep learning for smartphones. Journal of Real-Time Image Processing. doi:10.1007/s11554-017-0714-3

[6] Thomas, V., Chawla, N. V., Bowyer, K. W., & Flynn, P. J. (2007). Learning to predict gender from iris images. 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems. doi:10.1109/btas.2007.4401911

[7] Singh, M., Nagpal, S., Vatsa, M., Singh, R., Noore, A., & Majumdar, A. (2017). Gender and ethnicity classification of Iris images using deep class-encoder. 2017 IEEE International Joint Conference on Biometrics (IJCB). doi:10.1109/btas.2017.8272755.
[8] A Mallikarjuna Reddy, Vakulabharanam Venkata Krishna, LingamguntaSumalatha and AvukuObulesh, “Age Classification Using Motif and Statistical Features Derived on Gradient Facial Images”, Recent Advances in Computer Science and Communications (2020) 13: 965. https://doi.org/10.2174/2213275912666190417151247.

[9] Padmaja Grandhe, Dr. E. Sreenivasa Reddy, Dr.D. Vasumathi. (2016). An Adaptive Cluster Based Image Search and Retrieve for Interactive ROI to MRI Image Filtering, Segmentation, and Registration (Vol. 94, No.1). Journal of Theoretical and Applied Information Technology. Retrieved from http://www.jatit.org/volumes/Vol94No1/24Vol94No1.pdf

[10] Jyoti Patil, M.S., Pradeepini, G., Development of deep learning algorithm for brain tumour segmentation, International Journal of Engineering and Advanced Technology, 2019, 9(1), pp. 2800–2803

[11] Padmanabhnuni, Ms. S. S. (2020). AN EXTENSIVE STUDY ON CLASSIFICATION BASED PLANT DISEASE DETECTION SYSTEM. JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES, 15(5). https://doi.org/10.26782/jmcs.2020.05.00002

[12] A. M. Reddy, V. V. Krishna, L. Sumalatha and S. K. Niranjan, "Facial recognition based on straight angle fuzzy texture unit matrix," 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, 2017, pp. 366-372, doi: 10.1109/ICBDACI.2017.8070865.