International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

Face Gender Recognition Based on Neural Networks and Open CV

K. Yamini Saraswathi [*1], Dr. M. Sailaja[*2]
[*1] M. Tech Student, [*2] Professor,
Department of ECE, JNTU Kakinada, India

Article Info
Volume 9, Issue 5
Page Number : 555-562

Publication Issue
September-October-2022

Article History
Accepted : 10 Oct 2022
Published : 28 Oct 2022

ABSTRACT

Automatic gender recognition has now pertinent to an extension of its usage in various software and hardware, particularly because of the growth of online social networking websites and social media. However, the performance of already exist system with the physical world face pictures, images are somewhat not excellent, particularly in comparison with the result of task related to face recognition. Within this paper, we have explored that by doing learn and classification method and with the utilization of Convolutional Neural Networks (CNN) technique, a satisfied growth in performance can be achieved on such gender classification. The tasks that is a reason why we decided to propose an efficient convolutional network architecture which can be used in extreme case when the amount of training data used to learn CNN architecture. We examine our related work on the current unfiltered image of the face for gender recognizing and display it to dramatics outplay current advance updated methods. In this application we successfully proved CNN gives better results.

Keywords: Computer Vision, CNN, Classification, Unfiltered images, Gender recognition

I. INTRODUCTION

There has been rising interest in the problem of gender classification of text, especially in the social media and marketing domains. Much of this is due to the growing sources of user information, ranging from short tweets and comments to longer blog post and online novels. Existing systems mainly use features such as words, word classes, and POS (part-of-speech) n-grams for classification learning. However, none use deep learning, nor extend their models to a variety of sources, for example from the blogging to the media sphere. Our effort represents two contributions. First, we apply deep learning to datasets spanning several media (e.g. blogs and literature from several centuries). In this approach, our model directly learns to predict gender based on the surrounding context of words. This technique was proposed by Lai et al for a wide variety of text classifications, and we extend their previous work to the topic of gender classification. Second, we develop a model based on the RCNN developed for gender classification. Our model obtains an accuracy score comparable to the state-of-the-art models without much fine-tuning.

In this paper, we have used Convolution Neural Network (CNN) for gender recognition so in order to improvise the previously used method and to obtain much accurate result. Depicts how to process face image through CNN and find the pattern, extract feature to recognize gender from image accurately. The advantage of using CNN is it automatically extract the feature from an image and give output, we don't require to use feature descriptors like Histogram Oriented Gradient (HOG) and Support vector machine (SVM), eigenvector to extract the feature from image manually so as to do further recognition task or classification task. Gender Recognition was started with the problem in psychophysical studies to classify gender from human face; it concentrates on the efforts of perceiving human visual processing and recognizing relevant features that can be used to distinguish between female and male individuals. Exploration has proved that the discrepancy between a female face and male face can be used effectively to improvise the result of face recognition software in bio-metrics devices. With following recognition of trait-like gender, age, human expression, facial disease etc.

With this human-computer intercommunication, supervision, vigilance device, and digital vision system and much more that will work on whole human presence. Nevertheless, in a physical world scenario, the challenge is how to do work with the face image which influences various factors like illumination, pose facial expression, age estimation, occlusion, and background instruction data, and noise, error. It is a kind of motivation to do something new in the evolution of a boisterous face-based gen-der recognition application that has extreme detection accuracy. The Conventional Neural Network (CNN) technique used in face recognition, involving face dependent gender recognition, age-based recognition, comprises the phases of accepting the image as input and then transforming input images for further processing, dimension reduction, feature extraction, feature procurement, and classification, in this sequence.

Initial knowledge of these technique realms is needed to find out the finest extractor of feature for design. In extension to which, recognition method performance is highly vulnerable to the specified classifier used, which completely relies on the pattern retrieval technique applied to the method which we have used in the research work related to this paper. It is most difficult to find such a classifier that aggregate the finest among the chosen feature extractor so excellent recognition result can be obtained. The profound Convolutions neural system (CNN) is a neural system varies with the number about convolutional layers utilized to be compatible with sub-sampling layers and end with you quit offering on that one or additional completely joined layers (Fully connected layer) in the calibre multilayer perceptron. A convincing gain of the CNN over another traditional method in feature recognition technique is its capability to simultaneously perform following tasks like features extraction, reducing data dimension, and classification in the particular organize network structure. This kind of model is described and can speed up recognition process and provide the result with high accuracy and minimum cost.

The CNN performs both the work of feature classification and feature ex-traction and inside a single network structure through training a neural network on the collection of huge known data which is called training data normal face image dataset. The CNN has the capability of extraction of numerous different properties from an un-processed input image that requires either no or little pre-processing needed. The CNN gives halfway resistance and boosts to geometric transformations and deformation and 2-dimensional changes in shapes. Hence, the CNN is specially made to overcome, lacking the other existing feature extractor that is described by having static behaviour. The benefit obtains with the use of CNN is that they are comparatively easy to assist the network layer (input, hidden, output) in learning parameter, weight, (loss through BPNN). They have less number of parameters in comparison to fully connected multiple layer perceptron neural networks with the similar count of hidden layers used between input and output layer. Therefore, the CNN has shown an excellent successful result in a huge range of applications such as tracking human in mob i.e., human tracking system (HTS), surveillance system that deals with object/article/human, traffic signal recognition (TSR), optical character recognition (OCR), face recognition (FR), and many others application of CNN and obviously computer vision numerous applications.

II. RELATED WORKS

Discriminating Gender on Twitter: Accurate prediction of demographic attributes from social media and other informal online content is valuable for marketing, personalization, and legal investigation. This paper describes the construction of a large, multilingual dataset labelled with gender, and investigates statistical models for determining the gender of uncharacterized Twitter users. We explore several different classifier types on this dataset. We show the degree to which classifier accuracy varies based on tweet volumes as well as when various kinds of profile metadata are included in the models. We also perform a large-scale human assessment using Amazon Mechanical Turk. Our methods significantly out-perform both baseline models and almost all humans on the same task.

Improving Gender Classification of Blog Authors: The problem of automatically classifying the gender of a blog author has important applications in many commercial domains. Existing systems mainly use features such as words, word classes, and POS (parts-of speech) n-grams, for classification learning. In this paper, we propose two new techniques to improve the current result. The first technique introduces a new class of features which are variable length POS sequence patterns mined from the training data using a sequence pattern mining algorithm. The second technique is a new feature selection method which is based on an ensemble of several feature selection criteria and approaches. Empirical evaluation using a real-life blog data set shows that these two techniques improve the classification accuracy of the current state-of the-art methods significantly.

Gender Attribution: Tracing Stylometric Evidence Beyond Topic and Genre: Sociolinguistic theories (e.g., Lakoff (1973)) postulate that women's language styles differ from that of men. In this paper, we explore statistical techniques that can learn to identify the gender of authors in modern English text, such as web blogs and scientific papers. Although recent work has shown the efficacy of statistical approaches to gender attribution, we conjecture that the reported performance might be overly optimistic due to non-stylistic factors such as topic bias in gender that can make the gender detection task easier. Our work is the first that consciously avoids gender bias in topics, thereby providing stronger evidence to gender-specific styles in language beyond topic. In addition, our comparative study provides new insights into robustness of various stylometric techniques across topic and genre.

Improving word representations via global context and multiple word prototypes: Unsupervised word representations are very useful in NLP tasks both as inputs to learning algorithms and as extra word features in NLP systems. However, most of these models are built with only local context and one representation per word. This is problematic because words are often polysemous and global context can also provide useful information for learning word meanings. We present a new neural network architecture which 1) learns word embedding's that better capture the semantics of words by incorporating both local and global document context, and 2) accounts for homonymy and polysemy by learning multiple embedding's per word. We introduce a new dataset with human judgments on pairs of words in sentential context, and evaluate our model on it, showing that our model outperforms competitive baselines and other neural language models

Recurrent Convolutional Neural Networks for Text Classification: Text classification is a foundational task in many NLP applications. Traditional text classifiers often rely on many human-designed features, such as dictionaries, knowledge bases and special tree kernels. In contrast to traditional methods, we introduce a recurrent convolutional neural network for text classification without human-designed features. In our model, we apply a recurrent structure to capture contextual information as far as possible when learning word representations, which may introduce considerably less noise compared to traditional window-based neural networks. We also employ a max-pooling layer that automatically judges which words play key roles in text classification to capture the key components in texts. We conduct experiments on four commonly used datasets. The experimental results show that the proposed method outperforms the state-of-the-art methods on several datasets, particularly on document-level datasets.

A gender recognition system using shunting inhibitory convolutional neural networks: Automatic gender recognition has now pertinent to an extension of its usage in various software and hardware, particularly because of the growth of online social networking websites and social media. However, the performance of already exist system with the physical world face pictures, images are somewhat not excellent, particularly in comparison with the result of task related to face recognition. Within this paper, we have explored that by doing learn and classification method and with the utilization of Convolutional Neural Networks (CNN) technique, a satisfied growth in performance can be achieved on such gender classification tasks that is a reason why we decided to propose an efficient convolutional network architecture which can be used in extreme case when the amount of training data used to learn CNN architecture is limited. We examine our related work on the current unfiltered image of the face for gender recognition and display it to dramatics outplay current advance updated methods.

III. Methodology

In proposed system we use open computer vision and convolutional neural networks to overcome the difficulties obtained in the existing system. By using CNN, we can reduce the time and feature extraction process. By using Neural networks will automatically extracts the features from the images.

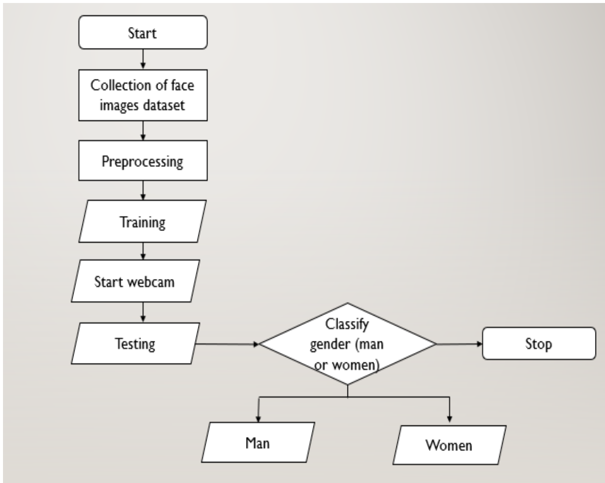


Figure 1 : Block diagram of proposed method

IV. Implementation

The project has implemented by using below listed algorithm.

Conventional Neural Network (CNN):

A convolutional neural network consists of an input layer, hidden layers and an output layer. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution. In a convolutional neural network, the hidden layers include layers that perform convolutions. Typically this includes a layer that performs a dot product of the convolution kernel with the layer's input matrix. This product is usually the Frobenius inner product, and its activation function is commonly ReLU. As the convolution kernel slides along the input matrix for the layer, the convolution operation generates a feature map, which in turn contributes to the input of the next layer. This is followed by other layers such as pooling layers, fully connected layers, and normalization layers.

Convolutional layers

In a CNN, the input is a tensor with a shape: (number of inputs) x (input height) x (input width) x (input channels). After passing through a convolutional layer, the image becomes abstracted to a feature map, also called an activation map, with shape: (number of inputs) x (feature map height) x (feature map width) x (feature map channels).

Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. Although fully connected feed forward neural networks can be used to learn features and classify data, this architecture is generally impractical for larger inputs such as high resolution images. It would require a very high number of neurons, even in a shallow architecture, due to the large input size of images, where each pixel is a relevant input feature. For instance, a fully connected layer for a (small) image of size 100 x 100 has 10,000 weights for each neuron in the second layer. Instead, convolution reduces the number of free parameters, allowing the network to be deeper. For example, regardless of image size, using a 5 x 5 tiling region, each with the same shared weights, requires only 25 learnable parameters. Using regularized weights over fewer parameters avoids the vanishing gradients and exploding gradients problems seen during back propagation in traditional neural networks. Furthermore, convolutional neural networks are ideal for data with a grid-like topology (such as images) as spatial relations between separate features are taken into account during convolution and/or pooling.

Pooling layers

Convolutional networks may include local and/or global pooling layers along with traditional convolutional layers. Pooling layers reduce the dimensions of data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. Local pooling combines small clusters, tiling sizes such as 2 x 2 are commonly used. Global pooling acts on all the neurons of the feature map. There are two common types of pooling in popular use: max and average. Max pooling uses the maximum value of each local cluster of neurons in the feature map, while average pooling takes the average value.

Fully connected layers

Fully connected layers connect every neuron in one layer to every neuron in another layer. It is the same as a traditional multi-layer perceptron neural network (MLP). The flattened matrix goes through a fully connected layer to classify the images.

Receptive field

In neural networks, each neuron receives input from some number of locations in the previous layer. In a convolutional layer, each neuron receives input from only a restricted area of the previous layer called the neuron's receptive field. Typically the area is a square (e.g. 5 by 5 neurons). Whereas, in a fully connected layer, the receptive field is the entire previous layer. Thus, in each convolutional layer, each neuron takes input from a larger area in the input than previous layers. This is due to applying the convolution over and over, which takes into account the value of a pixel, as well as its surrounding pixels. When using dilated layers, the number of pixels in the receptive field remains constant, but the field is more sparsely populated as its dimensions grow when combining the effect of several layers.

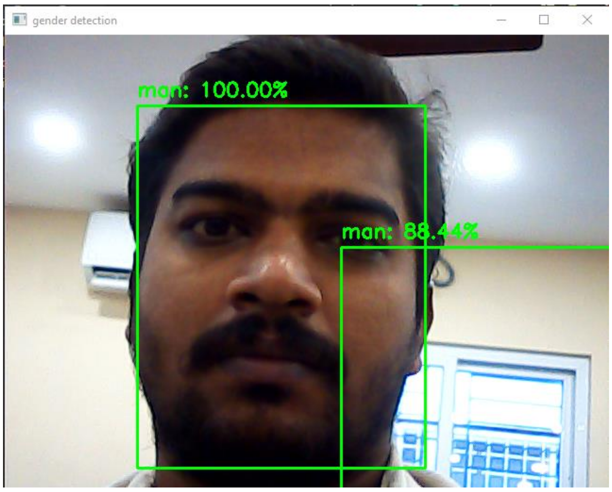
Weights

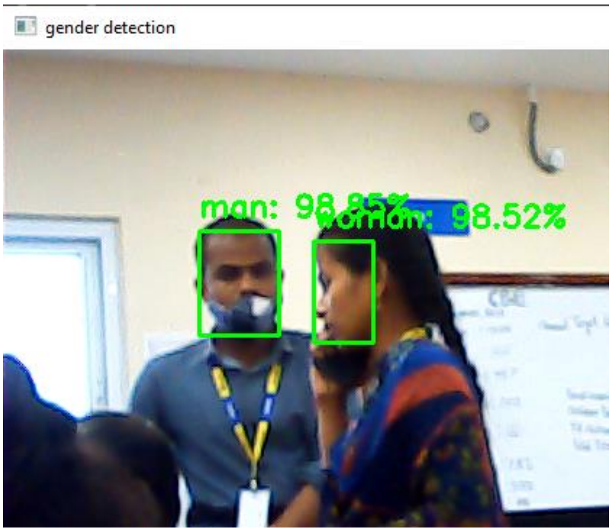
Each neuron in a neural network computes an output value by applying a specific function to the input values received from the receptive field in the previous layer. The function that is applied to the input values is determined by a vector of weights and a bias (typically real numbers). Learning consists of iteratively adjusting these biases and weights.
The vector of weights and the bias are called filters and represent particular features of the input (e.g., a particular shape). A distinguishing feature of CNNs is that many neurons can share the same filter. This reduces the memory footprint because a single bias and a single vector of weights are used across all receptive fields that share that filter, as opposed to each receptive field having its own bias and vector weighting.

V. Results and Discussion

The following images will visually depict the process of our project.

Here we can see the few images about identification of gender classification.





VI. Conclusion

In conclusion, we introduce the task of Visual Question Answering (VQA). Given an image and an open-ended, natural language question about the image, the task is to provide an accurate natural language answer. We provide a dataset containing over 250K images, 760K questions, and around 10M answers. We will set up an evaluation server and organize an annual challenge and an associated workshop to facilitate systematic progress. We demonstrate the wide variety of questions and answers in our dataset, as well as the diverse set of AI capabilities in computer vision, natural language processing, and common sense reasoning required to answer these questions accurately. The questions we solicited from our human subjects were open-ended and not task-specific. For some application domains, it would be useful to collect task-specific questions. For instance, questions may be gathered from subjects who are visually impaired, or the questions could focused on one specific domain (say sports). Bighametal. Created an application that allows the visually impaired to capture images and ask open-ended questions that are answered by human subjects. Interestingly, these questions can rarely be answered using generic captions. Training on task-specific datasets may help enable practical VQA applications.

VII. REFERENCES

[1]. S. Antol, C. L. Zitnick, and D. Parikh. Zero-Shot Learning via Visual Abstraction. In ECCV, 2014. 2, 3

[2]. J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh. VizWiz: Nearly Real-time Answers to Visual Questions. In User Interface Software and Technology, 2010. 1, 2, 8

[3]. K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge. In International Conference on Management of Data, 2008. 2

[4]. A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. R. H. Jr., and T. M. Mitchell. Toward an Architecture for Never-Ending Language Learning. In AAAI, 2010. 2

[5]. X. Chen, H. Fang, T.-Y. Lin, R. Vedantam, S. Gupta, P. Dollar, and ´ C. L. Zitnick. Microsoft COCO Captions: Data Collection and Evaluation Server. arXiv preprint arXiv:1504.00325, 2015. 3

[6]. X. Chen, A. Shrivastava, and A. Gupta. NEIL: Extracting Visual Knowledge from Web Data. In ICCV, 2013. 2

[7]. X. Chen and C. L. Zitnick. Mind's Eye: A Recurrent Visual Representation for Image Caption Generation. In CVPR, 2015. 1, 2

[8]. J. Deng, A. C. Berg, and L. Fei-Fei. Hierarchical Semantic Indexing for Large Scale Image Retrieval. In CVPR, 2011. 2

[9]. J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell. Long-term Recurrent Convolutional Networks for Visual Recognition and Description. In CVPR, 2015. 1, 2

[10]. D. Elliott and F. Keller. Comparing Automatic Evaluation Measures for Image Description. In ACL, 2014. 1

[11]. A. Fader, L. Zettlemoyer, and O. Etzioni. Paraphrase-Driven Learning for Open Question Answering. In ACL, 2013. 2

[12]. A. Fader, L. Zettlemoyer, and O. Etzioni. Open Question Answering over Curated and Extracted Knowledge Bases. In International Conference on Knowledge Discovery and Data Mining, 2014. 2

[13]. H. Fang, S. Gupta, F. N. Iandola, R. Srivastava, L. Deng, P. Dollar, ´ J. Gao, X. He, M. Mitchell, J. C. Platt, C. L. Zitnick, and G. Zweig. From Captions to Visual Concepts and Back. In CVPR, 2015. 1, 2

[14]. A. Farhadi, M. Hejrati, A. Sadeghi, P. Young, C. Rashtchian, J. Hockenmaier, and D. Forsyth. Every Picture Tells a Story: Generating Sentences for Images. In ECCV, 2010. 2

[15]. H. Gao, J. Mao, J. Zhou, Z. Huang, and A. Yuille. Are you talking to a machine? dataset and methods for multilingual image question answering. In NIPS, 2015. 2

[16]. D. Geman, S. Geman, N. Hallonquist, and L. Younes. A Visual Turing Test for Computer Vision Systems. In PNAS, 2014. 1, 2

[17]. S. Guadarrama, N. Krishnamoorthy, G. Malkarnenkar, S. Venugopalan, R. Mooney, T. Darrell, and K. Saenko. YouTube2Text: Recognizing and Describing Arbitrary Activities Using Semantic Hierarchies and Zero-Shot Recognition. In ICCV, December 2013. 2

[18]. M. Hodosh, P. Young, and J. Hockenmaier. Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics. JAIR, 2013. 1

[19]. A. Karpathy and L. Fei-Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions. In CVPR, 2015. 1, 2

[20]. S. Kazemzadeh, V. Ordonez, M. Matten, and T. L. Berg. ReferItGame: Referring to Objects in Photographs of Natural Scenes. In EMNLP, 2014. 2

[21]. R. Kiros, R. Salakhutdinov, and R. S. Zemel. Unifying VisualSemantic Embeddings with Multimodal Neural Language Models. TACL, 2015. 1, 2

[22]. C. Kong, D. Lin, M. Bansal, R. Urtasun, and S. Fidler. What Are You Talking About? Text-to-Image Coreference. In CVPR, 2014. 2

[23]. A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In NIPS, 2012.

Cite this article as :

K. Yamini Saraswathi, Dr. M. Sailaja, "Face Gender Recognition Based on Neural Networks and Open CV", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 5, pp. 555-562, September-October 2022.

Journal URL : https://ijsrst.com/IJSRST229595