International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 03 | March - 2022 ISSN: 2582-3930

Age and Gender Detection System using CNN

Hemleena Barik

School of Computer Science and IT – Jain (Deemed-to-be-University)

Abstract : *At the present time, the gender detection systems play a very important role in any person's day-to-day life. Human gender detection which is a part of facial recognition has received extensive attention because of its different kind of application. 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, I have explored that by doing learn and classification method and with the utilization of Convolution Neural Network. An approach using a convolutional neural network (CNN) is proposed for realtime gender classification based on facial images. The proposed CNN architecture exhibits a much-reduced design complexity when compared with other CNN solutions applied in pattern recognition. The number of processing layers in the CNN is reduced to only four by fusing the convolutional and subsampling layers. Unlike in conventional CNNs, I replace the convolution operation with cross-correlation, hence reducing the computational load. The network is trained using a second-order backpropagation learning algorithm with annealed global learning rates. Performance evaluation of the proposed CNN solution is conducted on two publicly available face databases of GitHub. The neural network is able to process and classify a 32 × 32-pixel face image in less than 0.27ms, which corresponds to a very high throughput of over 3700 images per second. Training converges within less than 20 epochs. These results correspond to a superior classification performance, verifying that the proposed CNN is an effective real-time solution for gender recognition.*

1. INTRODUCTION

Gender classification was first perceived as an issue in psychophysical studies; it focuses on the efforts of understanding human visual processing and identifying the key features used to categorize between male and female individuals. The conventional approach applied in face recognition, including face-based gender recognition, typically involves the stages of image acquisition and processing, dimensionality reduction, feature extraction, and classification, in that order. Prior knowledge of the application domain is required to determine the best feature extractor to design. The convolutional neural network (CNN) is a neural network variant that consists of a number of convolutional layers alternating with subsampling layers and ends with one or more fully connected layers in the standard multilayer perceptron (MLP). A significant advantage of the CNN over conventional approaches in pattern recognition is its ability to simultaneously extract features, reduce data dimensionality, and classify in one network structure. The CNN performs both feature extraction and classification within a single network structure through learning on data samples. Feature selection is also integrated into the training process by learning the weights held responsible for extracting features. The CNN also has the ability to extract topological properties from a raw input image with no or minimal pre-processing required. In addition, a certain degree of invariance is achieved while preserving the spatial topology of input data. The CNN also provides partial resistance and robustness to geometric distortions and transformations, and other 2D shape variations . Hence, the CNN is specifically designed to cope with shortcomings of the traditional feature extractor that is characterized by being static, is designed independently of the trainable classifier, and is not part of training procedure.

2. LITERATURE REVIEW

In the field of image processing and machine learning, a lot of research work has been done on human gender estimation. In this section, a brief overview of previous work on human gender estimation has been presented. Lian HC obtained an accuracy of 94.81% applying local binary pattern (LBP) and SVM with polynomial kernel on the CAS-PEAL face database. According to this method, a good accuracy can be achieved if the block size for the LBP operator is correctly selected, which is really a difficult task. Li et al. performed the classification of gender utilizing only five facial features (eyes, nose, mouth, brows, forehead). One drawback of this research is that the feature extraction method they have used is affected by complex backgrounds. Saeed Mozaffari, Hamid Behravan and Rohollah Akbari used geometric based feature for male female classification where they have used AR and Ethnic dataset containing 126 frontal images in each dataset. Here they have achieved 80.3% and 86.6% accuracy respectively. In a texture based local binary pattern has been used for feature extraction and as classification algorithm naive Bayes, ANN and linear SVM has been applied. They achieved 63% accuracy with only 100 face images that has been collected from Nottingham Scan database which is quite low. Sajja, T. K., Kalluri, H. K. have worked on gender classification from face images using LBP, SVM and Back Propagation. In this research they have used ORL dataset which contains 400 images and Nottingham Scan database which contains 100 images. After implementation they gained 100% accuracy for ORL dataset and 71% accuracy for Nottingham Scan database respectively. The work in showed a high classification accuracy of 99.30% using SUMS face database. In this work, the researchers applied 2D-DCT feature extraction, Viola and Jones face detection and the K-means nearest neighbour (KNN) algorithm as classifier. Being a compute-intensive algorithm, 2D-DCT is not suitable for real time applications. 4 Using principal component analysis (PCA), researchers in processed the face image to reduce the dimensionality. After that, a good subset of eigenfeatures has been selected using genetic algorithm (GA). Here, they reported an average error rate of 11.30%. The main drawback of this method is that, the GA exhibits high computational complexity. Althnian et al. used Human Gender Detection from Facial Images 191 hand crafted and fused features for face gender recognition where they have used both SVM and CNN and gained best accuracy 86.60% using CNN which can be improved further. Serna et al. worked on gender detection using VGG and ResNet where they analyzed how bias affects deep learning. They divided the images into 3 ethnic groups and also experimented on an unbiased group. Here they achieved best average accuracy 95.27%, LDA and SVM were used in this paper and they gained 96% and 97% accuracy for MBGC and FRGC dataset respectively. In they investigated human gender from hand shape from a small dataset containing 40 images and they achieved 98% accuracy. As classification algorithm Score-level fusion and LDA have been applied here. Hong 'Lim et al. presented a novel method for gender classification using finger nail with 80 samples donated by 40 people. With the use of PCA and SVM as classification algorithm, they showed about 90% accuracy in this research. So considering the whole literature review, it is clear that an improvement in gender classification is needed. The main disadvantages of the above gender classification research works are that, the feature extraction and the classification are performed separately. To obtain an optimum pre-processing and feature extraction design, prior knowledge is needed here. In case of CNN which is a multilayer neural network model, it can optimize filters through automated learning where it is independent of prior knowledge which demonstrate a superior performance can be achieved using CNN

3. METHODOLOGY

A neural network is a system of interconnected artificial "neurons" that exchange messages between each other. The connections have numeric weights that are tuned during the training process, so that a properly trained network will respond correctly when presented with an image or pattern to recognize. The network consists of multiple layers of feature-detecting "neurons". Each layer has many neurons that respond to different combinations of inputs from the previous layers. As shown in Figure 2, the layers are built up so that the first layer detects a set of primitive patterns in the input, the second layer detects patterns of patterns, the third layer detects patterns of those patterns, and so on. Typical CNNs use 5 to 25 distinct layers of pattern recognition.

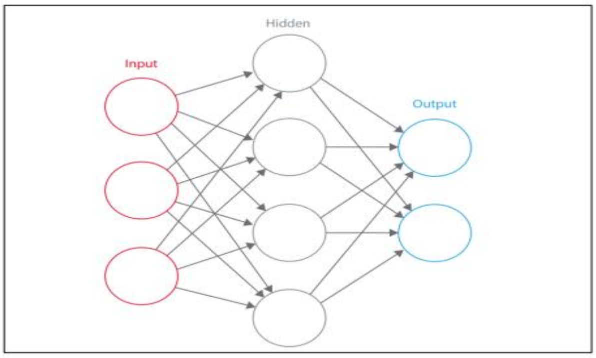

Figure 2. An artificial neural network

Training is performed using a "labelled" dataset of inputs in a wide assortment of representative input patterns that are tagged with their intended output response. Training uses general-purpose methods to iteratively determine the weights for intermediate and final feature neurons. Figure 3 demonstrates the training process at a block level.

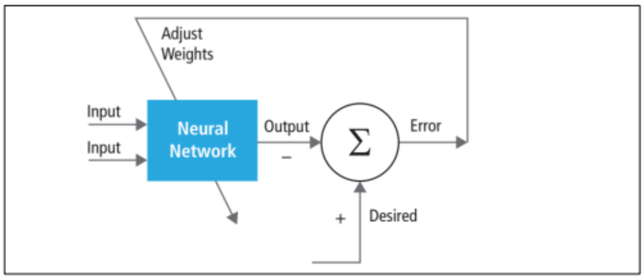

Figure 3. Training of neural networks

Neural networks are inspired by biological neural systems. The basic computational unit of the brain is a neuron and they are connected with synapses.

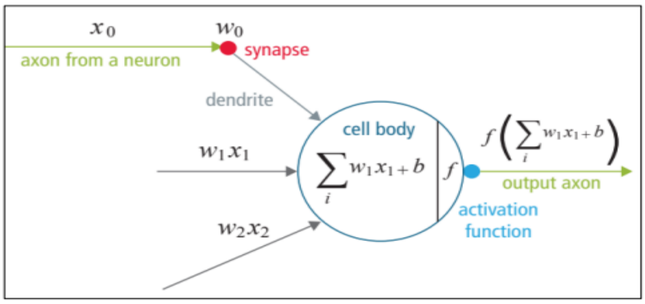

Figure 4. A biological neuron with a basic mathematical model

A CNN is a special case of the neural network described above. A CNN consists of one or more convolutional layers, often with a subsampling layer, which are followed by one or more fully connected layers as in a standard neural network. The design of a CNN is motivated by the discovery of a visual mechanism, the visual cortex, in the brain. The visual cortex contains a lot of cells that are responsible for detecting light in small, overlapping sub-regions of the visual field, which are called receptive fields. These cells act as local filters over the input space, and the more complex cells have larger receptive fields. The convolution layer in a CNN performs the function that is performed by the cells in the visual cortex. In typical CNN, each feature of a layer receives inputs from a set of features located in a small neighbourhood in the previous layer called a local receptive field. With local receptive fields, features can extract elementary visual features, such as oriented neural network that classifies feature vectors into classes. In a CNN, convolution layers play the role of feature extractor. But they are not hand designed. Convolution filter kernel weights are decided on as part of the training process. Convolutional layers are able to extract the local features because they restrict the receptive fields of the hidden layers to be local. In our proposed system, It utilizes a CNN (Convolutional Neural Network) architecture. CNN which is a deep learning algorithm is capable of distinguishing images from their characteristics. CNN is generally used for image analysis, image segmentation, image classification, medical image analysis, image and video recognition, etc. In this research, at first, applied an image processing technique as pre-processing on images to transform the raw data into an efficient and useful format. Later, the CNN architecture has been applied. CNN has 5 building blocks -

Kernel: A filter that used to extract the features from the input image. It is a dot matrix that moves over the input data perform dot product with the sub region of input data and get the output as matrix dot product. It moves on the input data by the stride value.

Stride: The amount of movement between applications of the filter to the input image is referred to as the stride.

Padding: Padding edge pixel. Padding can fix the border effect problem. It helps to keep all the content of image border.

Pooling: Pooling is required to down sample the detection of features in feature maps. Pooling summarizing the presence of features in the paths of the feature map. There are 2 methods available in pooling. 1. Max Pooling It summarized the most activated presence of features respectively. 2. Average Pooling It summarized the average pooling.

Flatten: Once the pooled featured map is obtained, the next step is flatten. It involves transforming the entire pooled feature map matrix into a single column which is the fed to the neural network for processing.

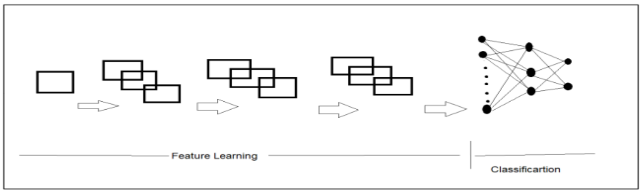

Figure 5. Work flow of CNN

4. CONCLUSION

Though many previous methods have addressed the problems of age and gender classification, until recently, much of this work has focused on constrained images taken in lab settings. Such settings do not adequately reflect appearance variations common to the real-world images in social websites and online repositories. Internet images, however, are not simply more challenging: they are also abundant. The easy availability of huge image collections provides modern machine learning based systems with effectively endless training data, though this data is not always suitably labelled for supervised learning. Two important conclusions can be made from our results. First, CNN can be used to provide improved gender classification results, even considering the much smaller size of contemporary unconstrained image sets labelled for gender. Second, the simplicity of our model implies that more elaborate systems using more training data may will be capable of substantially improving results beyond those reported here.

5. REFERENCES

[1] Tivive FHC, Bouzerdoum A. A gender recognition system using shunting inhibitory convolutional neural networks. In: International Joint Conference on Neural Networks; 2006; Vancouver, Canada. New York, NY, USA: IEEE. pp.5336–5341.

[2] Khalajzadeh H, Mansouri M, Teshnehlab M. Face recognition using convolutional neural network and simple logistic classifier. Stud Comp Intell 2013; 223: 197–207.

[3] Strigl D, Kofler K, Podlipnig S. Performance and scalability of GPU-based convolutional neural networks. In: 18th Euromicro International Conference on Parallel, Distributed and Network-Based Processing; 17–19 February 2010;Pisa, Italy. New York, NY, USA: IEEE. pp. 317–324.

[4] LeCun Y, Bottou L, Bengio Y, Haffner. P. Gradient-based learning applied to document recognition. P IEEE 1998;86: 2278–2324.

[5] Duffner S. Face Image Analysis with Convolutional Neural Networks. Munich, Germany: GRIN Verlag, 2009.

[6] Fan J, Wei X, Ying W, Yihong G. Human tracking using convolutional neural networks. IEEE T Neural Network 2010; 21: 1610–1623.