INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

E-attendance system using face recognition

Aditya Ladage
aditya.ladage@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Pooja Pache
pooja.pache@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Sahil Maniar
sahil.maniar@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

Sheetal Pereira
sheetalpereira@somaiya.edu
K. J. Somaiya College of Engineering, Mumbai, Maharashtra

ABSTRACT

Traditionally, the attendance of the students has been a major concern for the colleges and the faculties have to spend quite some time of their lectures for taking the attendance manually. In this paper, we are introducing a new way of attendance monitoring by making use of smartphones available with the teachers. We have suggested the use of YOLO algorithm for face detection and Siamese network for face recognition. This system will automatically mark the attendance of the students and thereby save the time and efforts for the faculties. The designed system will be quite efficient and reliable as Siamese network has proven to render high accuracies in face recognition.

Keywords— *YOLO (You Only Look Once), Siamese, CNN (Convolutional Neural Network), R-CNN (Region based Convolutional Neural Network), DNN (Deep Neural Networks), Darknet, CUDA*

1. INTRODUCTION

One of the considerably important issues for any college or university today is the attendance of the students. Attendance monitoring has been a tedious task for the faculties as they have to take attendance manually to avoid malpractices. At the end of each month, it is also the faculties' duty to generate attendance report of every student. We are introducing a new paradigm for the attendance system which will automate the process eventually reducing the work of the faculties. We have suggested the use of cameras in the classroom to capture the images of the students subsequently detecting and recognizing the students in the image and marking their respective attendance. For detecting the faces, we are making use of YOLO algorithm which is actually used for object detection but we will tune it for detecting faces. The recognition of the faces will be done using Siamese network which works on facial features of the detected faces for identification. The identified students will be marked and reports will be generated for every student. At the end of the month, our system will be notifying the students, their parents and their respective mentors about their attendance. In this way, the attendance monitoring system can be made with very little human interference.

2. RELATED WORK

The biometric system for attendance is well known and widely used these days but the problem with the biometric system is that it is delay based system where people have to form a queue to scan their fingerprints and get their attendance marked. Thus, we intend to remove this delay factor by making use of image processing.

In recent years, image processing has been used to process important information from the image which involves interpretation of the image for extracting useful information from the image. Also with the increasing amount of popularity of smartphones among the common people, there has been a demand to exploit the mobility of these devices and create easily accessible systems. This is why we have decided to create out attendance system which would make use of smartphones rather than any dedicated setup for attendance. The face of a person is unique but in a manner that it possesses a set of features that might resemble with some other face in one or the other way. Our system will make use of this unique set of features possessed by each student to recognize faces and mark attendance.

There have been many systems created for face detection and subsequently recognition for attendance system but there have been certain drawbacks in each of them. One way is using Viola-Jones algorithm and Back Propagation Neural Network (BPNN). In BPNN, there are two weighted propagations. In the forward propagation, the input is fed through the network to generate output activation of the propagation and in backward propagation, a feedback network is formed by feeding the output as the input in order to generate a difference between target and actual output. Hence, while training, two propagations are required for every epoch instead of one as in almost all other

networks. The Viola-Jones algorithm makes use of Haar feature selection for matching commonalities in human faces which has higher chances of inaccuracies.

Another approach for face recognition based system is noise filtering, histogram normalization and skin classification using MATLAB. There will be a problem with such a system as the system would be restricted to only the inbuilt features of MATLAB which would also make it difficult to train models for recognition. Another different type of network named Deep Neural Network (DNN) could be used for face recognition. Deep Face is a model using DNN which tends to be a human level accurate by taking a large set of images for training. The major issue with this model is that it requires an enormously large dataset and it is not possible to provide and store such a huge set of images for every student.

We will be using a unique combination of YOLO algorithm and Siamese network for our system.

3. SYSTEM DESCRIPTION

For the purpose of ease of development, we have divided our entire attendance system into multiple modules. The first module is the development of a face detection component for the system. Next step is to create a training module which can train itself based on the input images for a user. This will be followed by the face recognition module. The entire system will be hosted on a Django based web application with a MySQL database. As mentioned earlier in the introduction that the face detection and recognition will be done with the help of YOLO algorithm and Siamese network respectively, it is important to note that these are a part of Convolutional Neural Networks (CNN) and a deep understanding of CNN will help in easy implementation of these above-mentioned modules.

3.1 Convolutional Neural Networks

A Convolutional Neural Network is a Deep Learning algorithm which takes an input image, assigns importance to various aspects/features in the image and differentiates the various objects in the image from one another. CNN's have been inspired by the connectivity pattern of the human nervous network. They are made up of neurons with learnable weights and biases. Several inputs are received by each neuron and a weighted sum of these inputs is passed to an activation function and this results to the output.

Convolution is basically sliding a filter over an image and take a dot product along the way while sliding. Hence the result is a scalar quantity. In CNN, we take the input as tensor which is a multi-dimensional matrix of number. The main building blocks of CNN are the convolution layers over the tensor input. Each layer is dependent on an independent filter and is obtained by convolving the filter with the image. These filters are randomly initialized and we can make them be our parameters by learning the network over multiple epochs. The filters in the initial layers depict some low-level features such as edges, colors, gradients, etc. As we go deep in the convolutional layers we can find out high-level features giving us a network having an overall understanding of the image.

3.2 YOLO algorithm for face detection

The algorithms for object detection broadly classified into two groups. The first group is of algorithms based on the classification that works in two stages. The first stage involves the selection of interesting regions from the image and in the second stage, we classify those regions with the help of CNN. Some common examples of such algorithms are Region based Convolutional Neural Networks (R-CNN) and Fast R-CNN. The second group is of algorithms is based on regression. In such algorithms, instead of selecting interesting parts in the image, we try to predict classes and bounding boxes for the entire image in a single stage.

R-CNN and its family do not look on the entire image but rather they look on parts of the image which may have a greater chance of containing an object. While R-CNN is very accurate, the major problem with it is that it is very slow and also not a complete end-to-end object detector. The single stage detectors generally tend to be not as accurate as of the two-stage detectors, but are significantly faster. YOLO is a good paradigm of a single stage detector. While R-CNN can process only 5 FPS on a GPU, the YOLO algorithm can process up to 45 FPS on a GPU. Firstly, YOLO takes the input image. The image is divided into square grids by the YOLO framework. Then image classification and localization are applied to each of the square grids. Finally, YOLO predicts the bounding boxes and their class probabilities to detect the object if an object is found.

There are currently three main YOLO implementations – Darknet, AlexeyAB/darknet and Darkflow each with its own advantages and disadvantages. Darknet is the official implementation that has been created by the people who were behind this algorithm. Darknet has been written in C programming language with the help of CUDA which is a parallel computing architecture. This enhances the computing performance of Darknet by harnessing more power from the GPU.

We are using Darknet's pre-trained YOLO model. One can also create their own model and train it by providing labelled images of the object they want to detect. We are running YOLO with CUDA and OpenCV on NVIDIA 1050Ti graphic card. In general, YOLO can be used for detecting an object but for our attendance system, we are specifically interested only in face detection. So, we have written a script that trains and detects only human faces.

In figure 1, we can see the image provided to YOLO for detection of the human face and all the faces detected by YOLO are shown in figure 2.

Fig. 1: Input image for YOLO algorithm

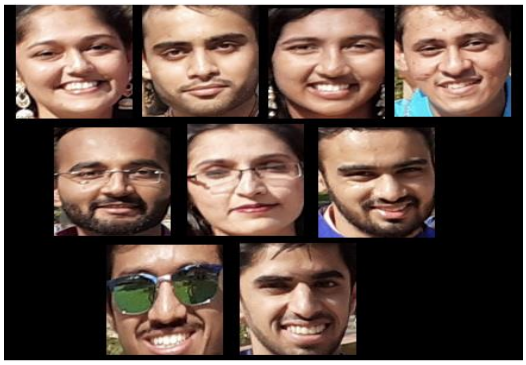

Fig. 2: Faces detected by YOLO algorithm

3.3 Siamese network for face recognition

In deep learning, generally, we require a huge amount of data and the model works well as the size of the dataset for training increases. However, it is always preferred if our model may learn from few data as it is not always possible to have a large set of data for the specific applications. Hence, for this purpose, we are using One-Shot Learning where our model can learn from a small set of data.

Siamese network works on the algorithm of One Shot Learning by taking only one image of the person for training. It is neither convenient nor feasible for the system to keep training the model every time an image of a new person is added to the database. Hence, it was necessary to use a neural network based on One Shot Learning algorithm for the face recognition part of our system. Siamese network finds a similarity between two faces by comparing the distance between the faces based on a set of parameters.

When a labelled image is provided for training to the network, the Siamese network creates an encoding also called as the feature vector representing the image after multiple layers of convolution. This encoding is based on facial features in the image like the distance between the eyes, the distance between the nose and lips and many such features which are unique to every person. The encoding is a 128×1 matrix with a certain weightage provide to each feature of the face. This encoding (or feature vector) is stored in the database for the respective user instead of storing the actual image of the face. To find the distance or similarity between any two images, the Siamese network uses a triplet loss function. In triplet loss, we have encodings for three images - an anchor image, a positive image and a negative image. To detect the faces to be similar, the distance between the faces should be less. For this, the distance between the anchor and positive image encodings should be low and the distance between the anchor and negative image encodings should be high. The difference between these two distances should be less than a negative value denoted byα. This will ensure proper recognition even in the case where the above distances are equal.

We have used the FaceNet model for our network. When the user provides an image for recognition, an encoding is created for the image using several convolutional layers. The distance between this encoding and each encoding in the database is found and the user encoding with which our image has the least distance is provided as the result of our recognition. We have set a threshold for the maximum distance that is acceptable for a user to be recognized. If the distance between two images is greater than this threshold, the result would be that no user is recognized. This ensures the elimination of false positive if the distance is high.

Fig. 3: User found in the database for the Siamese network

In figure 3, the face on the left was registered in the database for the user Sahil and that on the right was provided as an input and set the distance threshold to 0.6 resulting in the user being detected.

Fig. 4: User not found in the database for Siamese network

The figure 4, shows a user who was not registered in the database and hence could not be found as his distance from other users was much more than the specified threshold. This illustrates that the system is resistant to false positives.

3.4 Final system model

As face detection is an integral part of face recognition, we have to integrate the YOLO algorithm with Siamese network. When a faculty/professor clicks a photo of the students and submits the photo to our attendance system, YOLO algorithm is applied for detecting all the faces in the image. For every face detected, we are creating a separate image of that face. Each such image of the faces detected will be provided as an input to the Siamese network one by one and the result of the recognition phase will be stored in a file. This file would be later read to mark the attendance of the students. Also, once the face is recognized, the intermediate image that was the output of YOLO algorithm will be deleted by our system so as not to create any kind of storage overheads.

4. CONCLUSION

The attendance management system is a necessary tool for taking attendance in any environment where attendance is critical. Most of the existing approaches are time consuming, intrusive and require manual work from users. In this project, we have tried to eliminate the above challenges by successfully demonstrating the use of face recognition in student's e-attendance management system which will be marking attendance by acquiring the pictures with the help of a cell phone camera after which the faces in the image will be detected and compared with the already enrolled pictures in the database. Finally, based on the results of the acquired image, a student will be marked as absent or present. The proposed system is designed to be cost effective with no specific vendor hardware and software required for deployment.

5. FUTURE WORK

To extend the scope of this system, we can try to create a system that can not only just mark the attendance but also analyze it and create reports at the end of each month. We can also add a mailing system which will notify the students and their parents about the students' attendance report. Also, the accuracy of the system can be improved by training the system to correctly detect faces with scarves, sunglasses and beard. An extension of this system may also be used to detect malpractices during examinations like copying or talking. Another way to extend the system is could be faculties using it to know how they can make teaching more interesting and interactive by tracking down the student gestures, thus enhancing the quality of education.

6. REFERENCES

[1] Florian Schroff, Dmitry Kalenichenko, James Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), October 2015.

[2] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, "Siamese Neural Networks for One-shot Image Recognition", Department of Computer Science, University of Toronto 2015.

[3] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Recognition", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), September 2014.

[4] Sumit Saha 2018, "A Comprehensive Guide to Convolutional Neural Networks—the ELI5 way", Medium – Towards Data Science, viewed 17 March 2019, <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

[5] Harsh Pokharna 2016, "The best explanation of Convolutional Neural Networks on the Internet!", Medium, viewed 19 March 2019, <https://medium.com/ technology made easy/the-best-explanation-of-convolutional-neural-networks-on-the-internet-fbb8b1ad5df8>

[6] Adrian Rosebrock 2018, "YOLO object detection with OpenCV", PyImageSearch, viewed 20 December 2018, <https://www.pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/>.

[7] Michal Maj 2018, "What is object detection? Introduction to YOLO algorithm", Appsilon Data Science, viewed 18 December 2018, <https://appsilon.com/object-detection-yolo-algorithm/>.

[8] Enrique A. 2018, "Object detection with YOLO: implementations and how to use them", Medium, viewed 27 December 2018, < https://medium.com/@monocasero/object-detection-with-yolo-implementations-and-how-to-use-them-5da928356035>.

[9] Navesh Sallawar, Shubham Yende, Vaibhav Padgilwar, Vishal kale, Parag Gorlewar, Gaurav Varma, "Automatic attendance system by using face recognition", International Research Journal of Engineering and Technology (IRJET) Volume 4 Issue 4, April 2017.

[10] Smriti Tikoo, Nitin Malik, "Detection of Face using Viola Jones and Recognition using Back Propagation Neural Network", International Journal of Computer Science and Mobile Computing Volume 5 Issue 5, May 2016.

