International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Real Time Attendance System using Convolutional Neural Networks(CNN)

Ch. Raghava Prasad, Mallavelli Jaya SriNandan, Veerubhotla Surya Atchyuth, Siddana Phaneendra

Abstract: ***The management of the attendance can be an incredible weight on the instructors in the event that it is completed in registers. Determining this issue, keen and automatic attendance marking system by using the executive's framework is being used. In any case, verification is a significant problem in this framework. Brilliant attendance framework is implemented commonly along with the assistance of soft biometrics. Acknowledgment of face is one of the updated biometric techniques this framework got to be enhanced. Being a principle element of biometric confirmation, facial acknowledgment feature has become most utilized enormously in a few such applications, similar to video observing and surveillance-based CCTV film framework, a connection between PC and people and admittance frameworks existing inside and in network security. By using this structure, the issue present in along with intermediaries, understudies also have been checking on the present despite the fact that they are not physically present can without much of a stretch be illuminated. The primary usage steps utilized regarding this sort of framework are facial discovery and perceiving the distinguished the different face of the people. This term paper recommends a perfect model for actualizing a computerized attendance the board framework in order to make understudies for a class by utilizing the procedure of acknowledgment-based face detection procedure, by means of utilizing Convolutional Neural Network (CNN), Max pooling.***

Keywords— Max pooling, SoftMax function, ReLu rectified linear unit, CNN.

I. INTRODUCTION

The Person identification is one of the most decisive constructing integrant for new age smart interactions. Amongst all the present identification of person methods, facial recognition is known to be the most efficient one to recognize a person, since the face recognition technique is the best condition that uses to recognize people in daily lives. Although there are other different methods, such as identification of a person fingerprint, proved that can provide a better performance, those are not appropriate for their smart interactions due to their disruptive nature.

Biometrics are the sensational thing happened in the society to identify the unique persons in with difference fingerprints.

We humans have many ways to identify people that each and every human in the world is different. So the use of biometrics undoubtedly have an authentication to identify the right person. But in advancement happened in technology to make people more convenient and comfortable without wasting their time just by detecting their faces we can make person is present for the day. In contrast of biometrics, with manual face recognition provides static identification that is the person to be recognized does not need to cooperate or make any specific posture or hold on.Face recognition had come into picture at the mechanism became more efficient and intelligent and had the advancement to improve the abilities and senses of human. Explanations for facial recognition technology consideration on Applicability in various applications not only simple video of the person but comprising in content-context based video processing, law enforcement system and in security systems and Automatic detection and attendance system.Attendance using Face Recognition in real time which provides flexibleness to recognize several students at the same time distinctly rather than identifying once. To increase the accuracy in identification, efficiency in finding the right person and reliability of the recognition of unique people uniquely, algorithms are needed. To get the accuracy on several students at a time we must have perfect accurate detection algorithm in order to find the different id numbers at a time.

Universities are the place where researches and development and advancement of many machines, things and in life too but in all aspects the universities may be advanced but in case of attendance monitoring all these corporate universities are still using the old attendance system like taking attendance in book or website manually and later biometrics attendance system was introduced which is a revolution. At times this trend continues but in most of the universities biometrics was confined to faculty and workers not for the students. But now trend has changed with the new averaging technology.

Basically, in this conducted experiment is targeted for employing a system that is efficient of recognizing the students in an institution, taking their presence of the students daily and maintaining the everyday report. Therefore, face detection technique is employed to mark the presence of students.

Revised Manuscript Received on December 30, 2019.
***** Correspondence Author
Ch. Raghava Prasad,* Assoc.Professor (Electronics and Communications Engineering) K L E F Guntur, India chrp@kluniversity.in
Mallavelli Jaya SriNandan, B.Tech Student (Electronics and Communications Engineering)K L E F Guntur, India nandanmallavelli99@gmail.com
Veerubhotla Surya Atchyuth, B.Tech Student (Electronics and Communications Engineering) K L E F Guntur, India suryaav123@gmail.com
Siddana Phaneendra, B.Tech Student(Electronics and Communications Engineering K L E F Guntur, India siddanaphaneendra@gmail.com

So, in a college or a university webcams which generally acknowledge its frequent students when students are arriving into the classroom.

II. LITERATURE SURVEY

The core intention which we had done in this paperwork is developing a smart attendance marking management and maintained system using facial identification that will take attention of the hurdles which are being faced in other modern-day system which are operated automatically system on the basis of soft biometrics. The major methodology which is needed to be followed is totally a truly a current image of a student to that to be stored as a student database, which further than be used to observe the presence of students and point the attendance if the pictures of the students in database similar to the real-time image.

A model been stipulated by Naveed et al [4], which is a combination of set which has two databases. One of the databases contains the faces of the students with their id numbers and the other one is commonly used for marking the attendance of the students. The picture before the phases of detection and identification, the camera is used to connect the face picture of the student and performs background check and make the noise to be removed.

Hongsheng Li [5] present highly efficient algorithms for performing forward and reverse propagating techniques of Convolutional Neural Network (CNN) for pixelwise organization on images data. For pixelwise classification or organization task works, which are known as image segmentation and image detection, surrounding image patches are moved into Neural Network which is convolved for determining the classes of centered pixels via frontward propagation and for updating CNN parameters via reverse propagation. Notwithstanding, forward and in backward spread of propagation was initially intended for entire picture grouping or classification.

Rui Zhao [2] Direct employing of image data into pixelwise classification in a patch-by-patch scanning manner is exceptionally ineffective, since corresponding patches of pixels have large overlapping condition, which lead to a lot of unnecessary computation. The proposed algorithm disseminates all the repetitive calculation in convolution and pooling on pictures by presenting novel d-normally sparse kernels. It creates the very same outcomes as those by fix by-fix checking. Convolution and pooling activities with such kernels can consistently get to memory and can run proficiently on GPUs.

Xiaogang Wang [6] A small number of patches of intrigue can be looked over each preparation picture for in propagation which is backward by applying a mask to the error map at the end CNN layer. Its calculation intricacy is consistent regarding the quantity of patches examined from the picture. Tests have shown that our proposed calculations accelerate normally utilized patch by-patch looking over 1500 times in both forward and in reverse propagation. The speedup increments with the spans of pictures and fixes. Source code of GPU usage is fit to be discharged to the general people.

III. METHODOLOGY

A. Image Acquisition

Image Acquisition is the chief stage which fuses catching of tainted leaf pictures to fabricate database. The RGB shading pictures of persons are taken from the digital camera. The database of 30 pictures are assembled. During this movement, we perform the max pooling technique. From database five picture is taken and prepared further for testing.

B. Convolutional Layers

Contribution done by each and every convolution layer, similar to that of conventional neural system, is the yield of the topmost layer and tangled by a few convoluted bits. The convoluted pieces are utilized more than once in every tangible domains of the whole area, and the final outcome of convolution establishes a component guide of the information picture. Convolutional layers should operate a convolution operation to the input, moving the result got by present layer to its following layer consecutively. The convolution of the neural system emulates the output response of an individual neuron to visual stimuli. In different way of expressing, it determines the eliminating or exploding gradient problem in training of these outdated multi-layer neural networks in presence of multiple layers which are made to be hidden by the use of technique which is called backpropagation. What's more, the convolution pieces are the substance which must to be discovered by the convoluted layer, involving the weighted matrix w and the bias b. In this research, we considered the size of convolution of kernel is indicated as 10 x 10. w is instated by the "xavier" calculation, b is introduced with value zero(0), they will be at last dictated by convoluted training process of the network.

$$x_j^l = f\left(\sum_{i \in M_j^{l-1}} x_i^{l-1} k_{ij}^l + b_j^l\right) - - - eq1$$

Where representation of the above equation l the level, k is the convolution kernel, b is the bias, and M j indicates the feature map.

C. Convolution filters

Let’s consider that the shine of flashlight covers a area of 3 x 3. Presently, we should envision this flashlight sliding over every one of the zones of the information picture. In terms of deep learning, this flashlight is said to be known as **filter/kernel** and the region which is shining over is simply known **receptive field.** A significant note is that the depth of this channel filter must be equivalent to the depth of the info. The filters are then convoluted with the input volume to get intended activation maps. Activation maps specified as 'actuated' regions. The real evaluations of the kernel matrix is known to be the change with each learning cycle over the pre-training set.

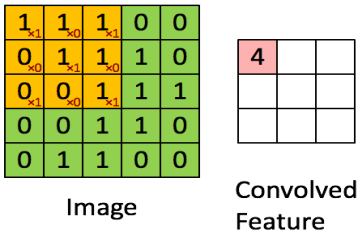


Fig1: Convoluving of feature image

D. Max-Pooling :

The foremost relevant category of pooling technique is a pooling layer of kernel size 2x2 which actually uses the MAX function. What it capable to do is, it could have the capability to take the size maximum of each 4*4 matrix of the original picture. The yield includes maps got after the computation of the convolution layer are commonly very little decreased in measurement. On the off chance that the measurement doesn't undergo any type of change over, there is going to be a lot of calculation need to do, and the learning process of network will turn out to be extremely troublesome, bound for get hold of a sensible outcome. The pooling layer is commonly a level to decrease the component of the element guide, and the technique is known as non-straight down-sampling technique. In the present data, each element diagram which must to be placed into the pooled layer is examined, after that the quantity of yield highlight maps is unaltered, yet the size of each element map will be littler.

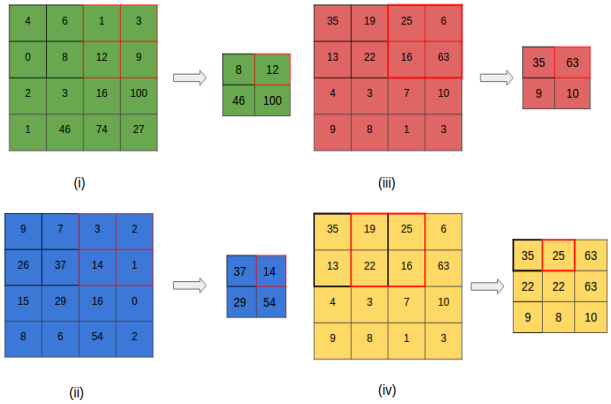


Fig2: Max pooling techinque

E. Activation functions :

We have to check whether the information that is received by neuron is familiar or same for the given data or else it should be ignored. The activation function is known to be non-linear transformation which we actually have to be done over the input signal. Eventually, the output which undergo transformation is sent to the following next layer of neurons as input information.

$$y = activation\left(\sum(weights \times inputs) + bias\right)$$

There are many activation functions available, but we use ReLu Rectified Linear Unit, SoftMax function.

F. Rectified Linear Unit Layer (ReLu)

A rectified linear unit has got the output as 0 .We have to look through whether the input is less than 0 or not, and if the raw output is not viable. Which means, if the input of the linear unit is happened to be greater than 0 then, the output obviously equivalent to the same as input. ReLu s' machinery is more like a real neuron in your body.

$$g(z) = max\{0, z\}$$ ----eq2

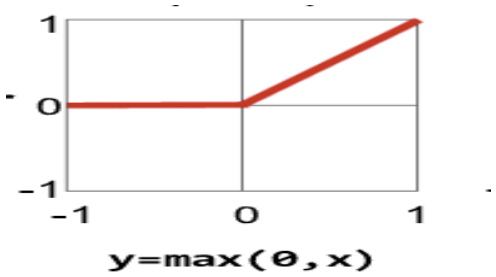


Fig3: Rectified Linear Unit(ReLu)

G. SoftMax Function:

To calculate the probabilities of each objective of every class present over all viable objective classes by the use of SoftMax function . Later the probabilities of those classes which are classified and calculated will be helpful for verifying the objective class for the given inputs. The most important benefit of getting the output probabilities range is satisfied by the use of SoftMax function. The range is in between **0 to 1**, and the sum of all the probabilities considered to be **equal to one**.

$$P\left(y = j|\theta^{(i)}\right) = \frac{e^{\theta^{(i)}}}{\sum_{j=0}^{k} e^{\theta_k^{(i)}}}$$ -----eq3

$$where\ \theta = w_0 x_0 + w_1 x_1 + \cdots + w_0 x_k = \sum_{i=0}^{k} w_i x_i = w^T x$$

IV. TRAINING

We must have to train our convolutional layers on the 100-class raw dataset of students taken from the digital camera. It is a database of more than 30 members and each candidate has his/her own a training set of 30 images of them along with .jpg file and .xml file. For the process pretraining we utilize the 23 convolutional levels or layers which are hidden pursued by a normal pooling layer and a completely associated layer. We prepared this network for around a week and accomplish a solitary yield exactness of 88% on the understudy face to detect by utilizing max pooling . We at that point have to transform the paradigm which has to perform identification. It views us that including both the layers of convolutional and associated to be as pretrained networks which can improve execution procedure. And being subsequent to their model, we include 4 convolutional layers and two completely associated layers with haphazardly modified loads. We normalize the rebounding box measurement and tallness varied by means of the picture width and stature with the goal that they fall between 0 and 1.

We parametrize the jumping envelope x and y directions to be offsets of a specific framework area so the values are also limited in between 0 and 1.

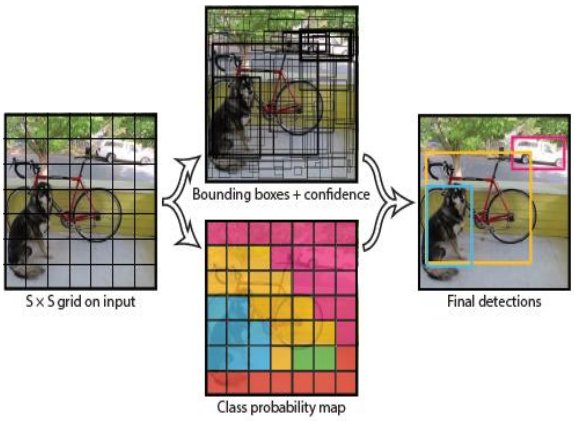


Fig4:Trained and Detected picture

We finally convoluted all the hidden layers and trained the data and recognition of students and attendance is registered and by using SoftMax function and ReLu Rectified Linear Unit with the activation the testing data is tested with less error probability.

V.TABULAR COLUMNS

Real Time Detectors	Train	mAP	FPS
Fast YOLO	2007+2012	51.7	149
30Hz DPM	2007	29.6	28
YOLO	2007+2012	63.6	39

Table 1. Parameters of Real Time Detectors

Less than Real Time Detectors	Train	mAP	FPS
Fastest DPM	2007	30.4	15
CNN	2007+2012	71.4	0.5
R-CNN	2007+2012	60.8	18
YOLO VGG-16	2007+2012	65.3	4.14

Table 2. Parameters of Real Time Detectors

VI. RESULTS

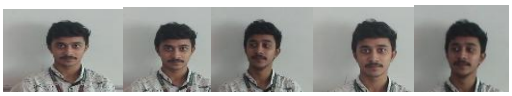
Trained Data

Trained Id: 160040942

Trained Id: 160040505

Trained Id: 160040827

Trained Id: 160040807



Trained Id: 160040174

Tested Outputs:

All the trained data is stored as .jpg files as database .xml file is used as testing database and we trained the data with 30 images of each students with different angles and with help of activations functions ReLu and SoftMax function we get the detected image as exact id numbers of the students.

VII. CONCLUSION

This facial recognition technique which we have experimented on the students faces is centered on convolution neural network (CNN) which is submitted in this paper. And the network which comprises of twenty-three hidden layers. The digital camera is used to take the video data and databases are created with the data which consists of student's facial images by doing the processes training and testing. We had done our research work on the database of tested data of faces of student database, the recognition rate is 89.59% correspondingly, all the experiments has been done on fully complete data of student's facial database, the recognition rate are 87.82% respectively. So, CNN based facial recognition is most efficient way for the facial recognition-based attendance system.

REFERENCES

1. Donahue, Jeff, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. "Decaf: A deep convolutional activation feature for generic visual recognition." In *International conference on machine learning*, pp. 647-655. 2014.
2. Zhao, Rui, Wanli Ouyang, and Xiaogang Wang. "Unsupervised salience learning for person re-identification." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3586-3593. 2013..
3. Sadeghi, Mohammad Amin, and David Forsyth. "30hz object detection with dpm v5." In *European Conference on Computer Vision*, pp. 65-79. Springer, Cham, 2014.
4. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.
5. Wang, Limin, Yuanjun Xiong, Zhe Wang, and Yu Qiao. "Towards good practices for very deep two-stream convnets." *arXiv preprint arXiv:1507.02159* (2015).
6. Kang, K., Ouyang, W., Li, H. and Wang, X., 2016. Object detection from video tubelets with convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 817-825).
7. Rao, G.A., Syamala, K., Kishore, P.V.V. and Sastry, A.S.C.S., 2018, January. Deep convolutional neural networks for sign language recognition. In *2018 Conference on Signal Processing And Communication Engineering Systems (SPACES)* (pp. 194-197). IEEE.
8. Prasad, M. V. D., B. Jwala Lakshmamma, A. Hari Chandana, K. Komali, M. V. N. Manoja, P. Rajesh Kumar, Ch Raghava Prasad, Syed Inthiyaz, and P. Sasi Kiran. "An efficient classification of flower images with convolutional neural networks." *International Journal of Engineering and Technology* 7, no. 11 (2018): 384-391.
9. Prasad, M. V. D., G. Jaya Sree, K. Gnanendra, P. V. V. Kishore, and D. Anil Kumar. "Fire Detection using Computer Vision Models in Surveillance Videos." (2006).
10. Bojja P., Sanam N., Design and development of artificial intelligence system for weather forecasting using soft computing techniques,2017 ARPN Journal of Engineering and Applied Sciences, Vol:12, issue:3, pp: 685-689, ISSN: 18196608
11. Maddala, Teja Kiran Kumar, P. V. V. Kishore, Kiran Kumar, and Anil Kumar. "YogaNet: 3D Yoga Asana Recognition Using Joint Angular Displacement Maps with ConvNets." *IEEE Transactions on Multimedia* (2019).
12. Kumar, E.K., Kishore, P.V.V., Kumar, M.T.K., Kumar, D.A. and Sastry, A.S.C.S., 2018. Three-Dimensional Sign Language Recognition With Angular Velocity Maps and Connived Feature ResNet. *IEEE Signal Processing Letters*, *25*(12), pp.1860-1864.
13. Prasad, Ch Raghava, and P. V. V. Kishore. "Performance of active contour models in train rolling stock part segmentation on high-speed video data." *Cogent engineering* 4, no. 1 (2017): 1279367.
14. Kumar, E.K., Kishore, P.V.V., Sastry, A.S.C.S., Kumar, M.T.K. and Kumar, D.A., 2018. Training CNNs for 3-d sign language recognition with color texture coded joint angular displacement maps. *IEEE Signal Processing Letters*, *25*(5), pp.645-649.
15. Zhao, R., Ouyang, W. and Wang, X., 2013. Person re-identification by salience matching. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2528-2535).
16. Raghava Prasad C., Kishore P.V.V., Morphological differential gradient active contours for rolling stock segmentation in train bogies ,2016, ARPN Journal of Engineering and Applied Sciences, Vol: 11, Issue: 5, pp: 2799 - 2804, ISSN 18196608

AUTHORS PROFILE

Ch. Raghava Prasad
Assoc.Professor (Electronics and Communications Engineering)
K L E F
Guntur, India
chrp@kluniversity.in

Mallavelli Jaya SriNandan
B.Tech Student (Electronics and Communications Engineering)
K L E F
Guntur, India
nandanmallavelli99@gmail.com

Veerubhotla Surya Atchyuth
B.Tech Student (Electronics and Communications Engineering)
K L E F
Guntur, India
suryaav123@gmail.com

Siddana Phaneendra
B.Tech Student(Electronics and Communications Engineering
K L E F
Guntur, India
siddanaphaneendra@gmail.com