INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 3)
Available online at: www.ijariit.com

Face recognition using CNN

Mrunmayee Vaidya
mrunmayeesv@gmail.com
Sinhgad Institute of Technology, Pune, Maharashtra

Jigyasa Solanki
anyashah447@gmail.com
Sinhgad Institute of Technology, Pune, Maharashtra

Sravani Wayangankar
gauri1002@gmail.com
Sinhgad Institute of Technology, Pune, Maharashtra

ABSTRACT

Face recognition has been the need of the hour for a long time. Now being implemented, we look forward to enhance its features. The topic of Face recognition using Convolutional Neural Networks is the most upcoming technology. This paper outlines the different algorithms used in the face recognition system based on Convolutional Neural Networks. It has put forth the applications and the algorithms that are used in Softmax classifier and the Deep Learning systems.

Keywords*— Artificial Intelligence, Convolutional Neural Networks, Classifier, Neural Networks, Biometrics, Deep learning*

1. INTRODUCTION

Face recognition systems have made great advancements in the fields of biometric identification. Every face has unique identification features and psychological as well as behavioral characters. Numerous algorithms and techniques have been developed for improving the performance of face recognition. Deep learning is making crucial advances in solving problems that have restricted the best attempts of the artificial intelligence community for many years.

Among one of the deep learning, approaches is Convolutional Neural Networks (CNN) which is used in developing softwares for facial recognition systems. This software uses deep learning algorithms to compare a live capture or digital image to the stored face print in order to verify an individual's identity. The general structure of the face recognition process is of three stages. It starts with pre-processing stage: color space conversion and resizing of images, continues with the extraction of facial features, and afterwards extracted feature set is classified.

2. METHOD AND MATERIAL

2.1 Data

Training and testing datasets which comprise of 48-by-48-pixel grayscale images. Images in the data sets vary considerably, some images spectacles or beard. All these images are preprocessed, grayscale and reduced to box side around the image to make identification of features easy hair like feature identification to detect darker and lighter regions of the face and reduce boundaries.

2.2 Webcam

The webcam is used to capture images. Usually, inbuild with laptops or connected webcam to computer systems.

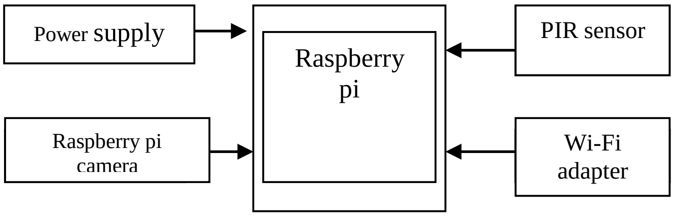

Fig. 1: Hardware diagram

3. SOFTMAX CLASSIFIER

Softmax classifiers give you probabilities for each class name. Moreover, for datasets, for example, ImageNet, we regularly take a gander at the rank-5 exactness of Convolutional Neural Networks (where we verify whether the ground-truth mark is in the main 5 anticipated names returned by a system for a given info picture). Understanding Multinomial Logistic Regression and Softmax Classifiers.

The Softmax classifier is a speculation of the double type of Logistic Regression. Much the same as in pivot misfortune or squared pivot misfortune, our mapping capacity f is characterized with the end goal that it takes an information set of information x and maps them to the yield class names through a straightforward (direct) speck result of the information x and weight framework W:

$$f(x_{i}, W) = Wx_{i}$$

Be that as it may, dissimilar to pivot misfortune, we decipher these scores as unnormalized log probabilities for each class name — this adds up to swapping out our pivot misfortune work with cross-entropy misfortune:

$$L_{i} = -log(e^{s_{y_{i}}}/\sum_{j} e^{s_{j}})$$

Things being what they are, how could I land here? How about we break the capacity separated and investigate.

To begin, our misfortune capacity should limit the negative log probability of the right class:

$$L_{i} = -log\ P(Y = y_{i}|X = x_{i})$$

This likelihood of articulation can be translated as:

$$P(Y = k|X = x_{i}) = e^{s_{y_{i}}}/\sum_{j} e^{s_{j}}$$

Where we utilize our standard scoring capacity structure:

$$s = f(x_{i}, W)$$

All in all, this yields our last misfortune work for a solitary information point, much the same as above:

$$L_{i} = -log(e^{s_{y_{i}}}/\sum_{j} e^{s_{j}})$$

Note: Your logarithm here is really base e (regular logarithm) since we are taking the backwards of the exponentiation over e prior.

The genuine exponentiation and standardization through the total of types is our real Softmax work. The negative log yields our genuine cross-entropy misfortune.

Similarly as in pivot misfortune or squared pivot misfortune, registering the traverse a whole dataset is finished by taking the normal:

$$L = \frac{1}{N} \sum^{N}_{i = 1} L_{i}$$

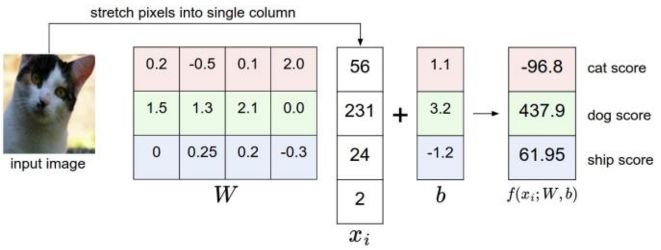

Fig. 2: Probabilities in softmax

4. DEEP LEARNING

Denoted mathematically, if a kernel i, at position $(x; y)$ is used to

Compute the activity of a neuron denoted by $a^{i}_{x,y}$ and the ReLU nonlinearity is then applied, the response-normalized activity $b^{i}_{x,y}$ can be expressed as

$$b^{i}_{x,y} = a^{i}_{x,y} \Bigg/ \left(k + \alpha \sum_{j=max(0,i-n/2)}^{min(N-1,i+n/2)} \left(a^{j}_{x,y}\right)^{2}\right)^{\beta},$$

Where N the total number of kernels in the layer and the sum is runs over n “adjacent” kernel maps at the same spatial position. This scheme aided generalization and reduced its network classification error rates. They further reduced the classification error by overlapping the network's max-pooling layers. It consisted of five convolutional layers, three of which were followed by max-pooling layers, and three fully connected layers. The various layer parts in the top half of the figure ran on one GPU, while the layer parts at the bottom ran on the second GPU. The GPUs interacted with each other only at specific layers.

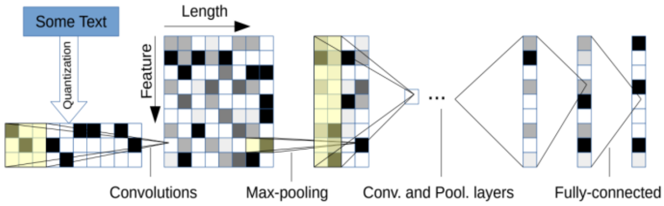

Fig. 3: Classification with deep learning

5. CONCLUSION

In this paper, we have devised methods and techniques to successfully build a facial recognition system. Along with a basic 5 layered CNN, we have also involved techniques like max pooling and fine tuning to get into deeper layers of CNN and get the highest accuracy rates. The proposed paper gives an idea to build a system that would detect motion using a PIR sensor, switch the camera to capture the image and provide the identity of the person.

6. REFERENCES

[1] Abdulkader, A. (2006). A two-tier Arabic offline handwriting recognition based on conditional joining rules. In Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition (pp. 1–6). Los Alamitos, CA: IEEE Computer Society

[2] Basu, S., Karki, M., DiBiano, R., Mukhopadhyay, S., Gangu ly, S., Nemani, R., & Gayaka, S.(2016). A theoretical analysis of deep neural networks for texture classification. arXiv 1605.02699

[3] Cheng, Z., Soudry, D., Mao, Z., & Lan, Z. (2015). Training binary multilayer neural networks for image classification using expectation backpropagation. arXiv 1503.03562.

