

Door Camera Facial Recognition with Raspberry PI

¹Mohit Tiwari, ²Tripti Tiwari, ³Anuj S, ⁴Amal V

¹Bharati Vidyapeeth's College of Engineering New Delhi

²BVIMR, New Delhi

³Galgotias University, Noida

⁴REVA University, Bangalore

Abstract

The most significant of highlight of any home security framework is to identify the individuals who go into or go out. Rather than checking that through passwords or pins remarkable appearances can be utilized as they are one's biometric characteristic. These are intrinsic and can't be adjusted or taken without any problem. The degree of security can be raised by utilizing face location. The proposed face acknowledgment entryway lock security framework has been created to forestall theft in exceptionally secure territories like home condition with lesser force utilization and increasingly solid independent security gadget for both Intruder identification and for entryway security. This framework is controlled by raspberry pi circuit. Raspberry Pi electronic board is worked on Battery power gracefully, remote web availability by utilizing USB modem, it incorporates camera, PIR motion sensor and a door. At whatever point the individual comes before the entryway, it perceives the face and in the event that it is enlisted, at that point it opens the entryway, if the face isn't enrolled it will raise a caution and snaps an image and send it on the enlisted number.

Keywords: Raspberry Pi, USB modem, PIR movement sensor

INTRODUCTION

Raspberry Pi is a minimal effort and minuscule estimated PC that connects to a PC screen and uses a console and mouse. It empowers individuals of any age to pick up registering in a simple and proficient manner with programming dialects like Python [1], Windows Programming and Android. Nowadays it has a major role in security systems like CCTV, Door lock and unlocking system, automatic gate system etc. Convolutional Neural Network (CNN) is the most commonly wide algorithm for image processing, object detection and the like. In CNN image classification, images are captured and processed or classify under certain categories. This study was conducted to utilize an image processing technique, implement Convolutional Neural Network (CNN) and measure the effectiveness for classifying and detecting the faces [2].

DETECTION USING CNN

CNN stands for convolution neural networks. This is different from other neural networks in the way it process image and the mathematical expression it uses for the classification [3]. General technique involved in CNN is, when the image is passed to the network, there will be several layers in the network. Each layer has its own work. The basic work done by specified layer is they contain filters [4]. Since the retina image is made of many pixels. The basic building block of filter is a pixel of an exudate or an MA or hemorrhages. This filter is passed on to each and every pixel of the image, the pixels which match with the filters and feature mapped and send to the next layer. The same process gets continued until each pixel

is checked and result is generated as per label given. The training set and testing set should consist of at least 12000 retinal images of all classes so that the training of the classifier gets easy.

Convolution layer: This is the primary layer which extracts the features of input. This contains several filters(kernel). It learns the input image and each filter is convolved against it, so that the feature extraction is done by forming feature maps [5]. Every layer represents a different feature. It is a differential operation that considers image matrix and a filter as $ip \cdot A$ Matrix that represents an Image comprises of 3 dimensions($H*W*D$)

A filter ($fh*fw*d$)

• Output is $= (H-Fh+1)*(W-F+1)*1$

Pooling layer: Pooling layer is a discretization process and is used to grab the key parameters when the input images are large to learn and provides basic reduction of them to internal representation, thus it reduces the complexity [6]. Spatial pooling can reduce the dimensionality but retains the important information. It consists of several types of spatial Pooling namely

- Max Pooling
- Average pooling
- Sum Pooling

Rectification layer: Rectification layer adds a specific non-linear functionality to each layer to avoid the linear passage in the neural net, Compared to Sigmoid and regressive functions. ReLU trains Deep neural network with appropriate functionality. Let Kth layer is the rectification layer and it considers activation volume $X(k-1)$ from a Non-Linear layer (k-1) and generates the modified activation volume $X(k)$: $X(k) = |X(k-1)|$ These operations will not change the input parameters and input constraints and their merger lead to a single layer The output is $f(X) = \text{Max}(0, X)$ [7].

Fully Connected Layer: The resultant matrix from the above layers is converted into vector and fed into a fully connected neural net. This is used to perform high level reasoning after various iterations of convolution and Max pooling . It connects to each and every neuron. We have an activation function to classify the outputs [8]. Classification Layer: After processing through multiple layers, it is the final layer which classifies the fundus image.

METHODOLOGY

The steps involved for this face detection are

- I) Feature extraction
 - Convolution
 - Activation
 - Pooling
- II) Classification
 - Flattering
 - Fully connectedness
 - Prediction

Every layer in a convolution network isn't connected to each and every neuron in the next layer. A layer is nothing but a set or a series of convolution operations involved. Receptive field is the part of the image that is focused on which the convolution operation is applied [9]. This receptive field slides across the

image. Sliding is nothing but applying a dot product between the weight matrix at a layer and every part of the image iteratively. [10] A series of dot product is performed between the weight matrices and input matrix. Figure 1 shows the flow diagram of this work.

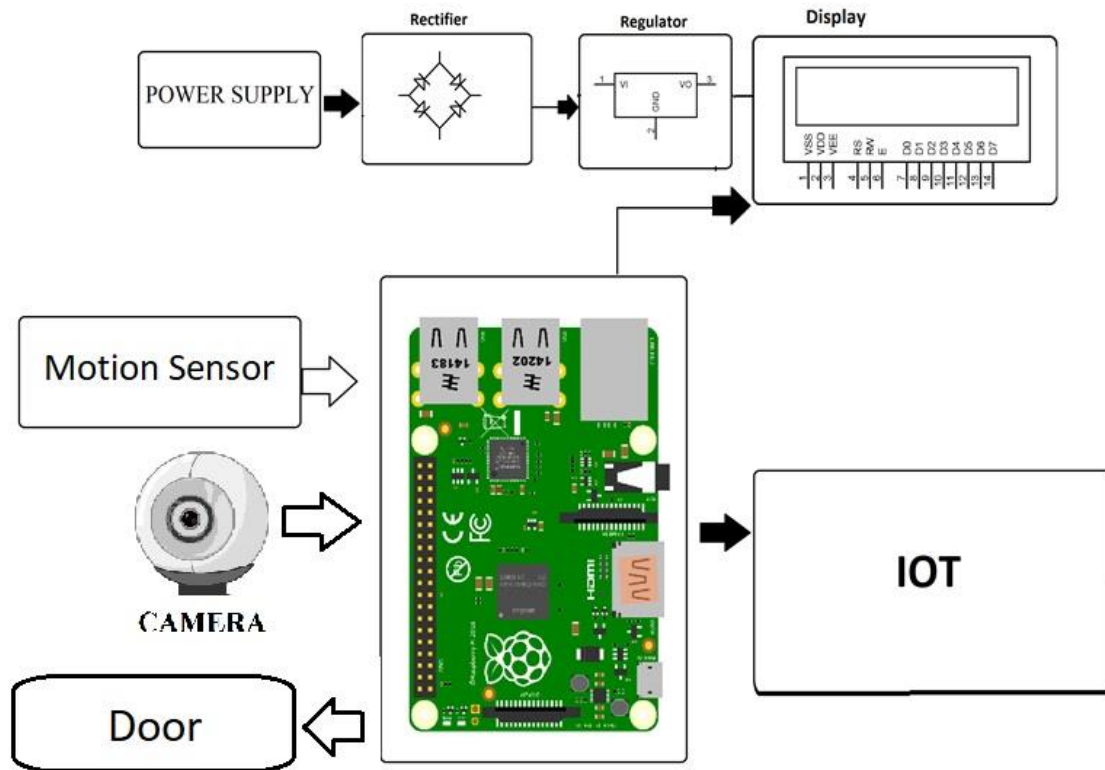


Fig.1: Flow Diagram of Face detection

EXPERIMENTAL RESULTS

The following results provide performance of the neural networks for the face detection. Figure 2 shows the layer information and parameters of the model which we developed.

Model:

Layér(type)	Output.Shapé	Parám.#
input5(InputLayér)	(None, 224, 224, 3)	0
block1conv1(Conv.2D)	(None, 224, 224, 64)	1792
block1conv2(Conv.2D)	(None, 224, 224, 64)	36928
block1pool(MaxPooling.2D)	(None, 112, 112, 64)	0

block2conv1(Conv.2D)	(None, 112, 112, 128)	73856
block2conv2(Conv.2D)	(None, 112, 112, 128)	147584
block2pool(MaxPooling.2D)	(None, 56, 56, 128)	0
block3conv1(Conv.2D)	(None, 56, 56, 256)	295168
block3conv2(Conv.2D)	(None, 56, 56, 256)	590080
block3conv3(Conv.2D)	(None, 56, 56, 256)	590080
block3conv4(Conv.2D)	(None, 56, 56, 256)	590080
block3pool(MaxPooling.2D)	(None, 28, 28, 256)	0
block4conv1(Conv.2D)	(None, 28, 28, 512)	1180160
block4conv2(Conv.2D)	(None, 28, 28, 512)	2359808
block4conv3(Conv.2D)	(None, 28, 28, 512)	2359808
block4conv4(Conv.2D)	(None, 28, 28, 512)	2359808
block4pool(MaxPooling.2D)	(None, 14, 14, 512)	0
block5conv1(Conv.2D)	(None, 14, 14, 512)	2359808
block5conv2(Conv.2D)	(None, 14, 14, 512)	2359808
block5conv3(Conv.2D)	(None, 14, 14, 512)	2359808
block5conv4(Conv.2D)	(None, 14, 14, 512)	2359808
block5pool(MaxPooling2D)	(None, 7, 7, 512)	0
flatten1(Flatten)	(None, 25088)	0
dense1(Dense)	(None, 5)	125445
=====		
Totalparams: 20,149,829		
Trainableparams: 125,445		
Non-trainableparams: 20,024,384		

Fig 2: Layer information and parameters of the model

Figure 3 describes the accuracy of the developed model. It explains the data loss and accuracy of the model.

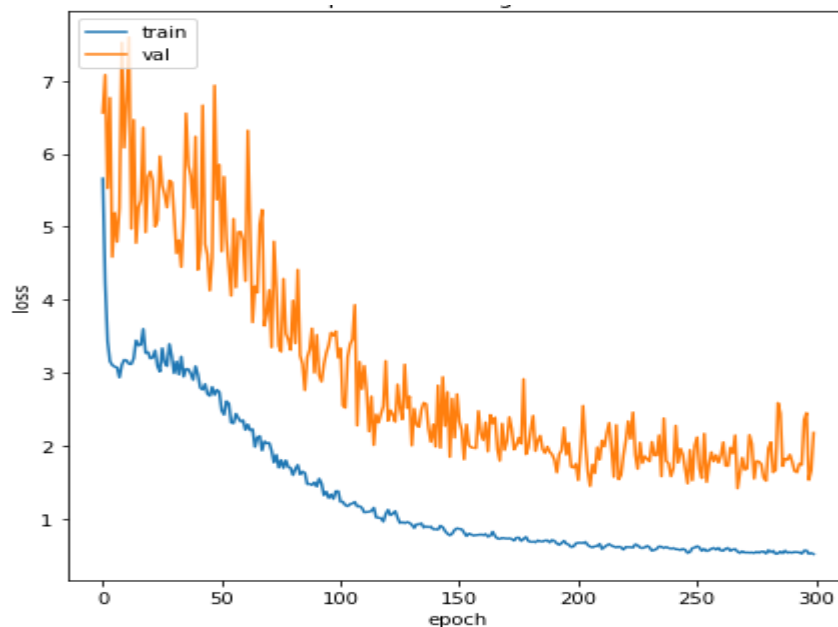


Fig 3: Train loss and validation loss of the model

In the above figure, the X-axis represents epochs and Y-axis represents the loss value. Here we used 300 epochs for this face detection project. If the epoch number is increasing then the validation and train loss getting reduced. It clearly shows the accuracy of the model. If the model has less validation loss and train loss then we can assume the accuracy is good and the model is also good.

CONCLUSION

Door camera technique is widely used for security purpose based on the face detection nowadays. We got the accuracy of 97 percent. It is correctly detecting the faces of the people who is coming to home and analysis the faces correctly and taking action on time. There is a future scope for this we can implement in the CCTV camera and identify the objects easily.

REFERENCES

1. Ji Shunping & Chi, Zhang & Xu, Anjian & Shi, Yun & Duan, Yulin.(2018). —3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images" 10. 75. 10.3390/rs10010075.
2. Florent Marie, Lisa Corbat, Yann Chaussy, Thibault Delavelle, Julien Henriet, Jean-Christophe Lapayre, Segmentation of deformed kidneys and nephroblastoma using Case-Based Reasoning and Convolutional Neural Network, Expert Systems with Applications, Volume 127, 2019, Pages 282-294, ISSN 0957-4174
3. Bullock, Joseph & Cuesta, Carolina & Quera-Bofarull, Arnau. (2018). XNet: A convolutional neural network (CNN) implementation for medical X-Ray image segmentation suitable for small datasets.

4. Yan Song, Bo He, Peng Liu, Tianhong Yan, Side scan sonar image segmentation and synthesis based on extreme learning machine, *Applied Acoustics*, Volume 146, 2019, Pages 56-65, ISSN 0003-682X
5. Shan E Ahmed Raza, Linda Cheung, Muhammad Shaban, Simon Graham, David Epstein, Stella Pelengaris, Michael Khan, Nasir M. Rajpoot, Micro-Net: A unified model for segmentation of various objects in microscopy images, *Medical Image Analysis*, Volume 52, 2019, Pages 160-173, ISSN 1361-8415,
6. Dan C Cirean, Alessandro Giusti, Luca M Gambardella and Schmidhuber, —Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images||, *Advances in neural information processing systems* 25 · January 2012
7. Wang, Shengke& Liu, Lu & Qu, Liang & Yu, Changyin& Sun, Yajuan& Gao, Feng & Dong, Junyu. (2018). —Accurate Ulva Prolifera Regions Extraction of UAV Images with Superpixel and CNNs for Ocean Environment Monitoring|| *Neurocomputing*, 10.1016/j.neucom.2018.06.088.
8. Jie Chang, Luming Zhang, Naijie Gu and Xiaoci Zhang, —A Mix-pooling CNN Architecture with FCRF for Brain Tumor Segmentation||, *Journal of Visual Communication and Image Representation* 58 · December 2018, DOI: 10.1016/j.jvcir.2018.11.
9. AyseBetulOktay, Anil Gurses, Automatic detection, localization and segmentation of nano-particles with deep learning in microscopy images, *Micron*, Volume 120, 2019, Pages 113-119, ISSN 0968-4328,
10. Guotai Wang, Wenqi Li, Michael Aertsen, Jan Deprest, SébastienOurselin, Tom Vercauteren, Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks, *Neurocomputing*, Volume 338, 2019, Pages 34-45.
11. Gardner, Daniel, and David Nichols. "Multi-label Classification of Satellite Images with Deep Learning." (2017): 2017.