ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

Implementation of Face Recognition System Using Convolutional Neural Networks for Automated Door Access Using Raspberry Pi

Druva Kumar S[1], Syeda Arzoo Irshad[2], Channayya[2], Rakshith[2]

Assistant Professor, Dept. of ECE, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India[1]

UG Student, Dept. of ECE, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India[2]

ABSTRACT: Despite the existence of various biometric techniques, like fingerprints, iris scan, and hand geometry, face recognition is considered to be the most efficient. This is because it is inexpensive, non-intrusive and natural. This experiment involves the implementation of the OpenFace project consisting of a face recognition system built using deep convolutional neural networks and its real time application in the form of automated door access using an embedded system such as Raspberry Pi 3. The system uses a combination of various machine learning algorithms such as Histogram of Oriented gradients (HOG), convolutional neural networks and Support Vector Machines. When the visitor presses the push-button, the system executes a blink detection program to check if the person is genuine as opposed to a photograph of the person by detecting if the person has blinked at least once. The blink detection is implemented as a protection against malicious attempts to spoof the face recognition system. Once it's confirmed that it's a real person at the door, an image is captured by the camera. The image is then subjected to the high precision face recognition system. The captured image is compared with images in the existing database. If a match is found, the door automatically opens. If the picture is not recognised, an e-mail is sent to the owner alerting him of a possible intruder with an attachment of the captured image. If the owner recognises this person, the owner can unlock the door remotely else the door remains shut. The door is controlled by a servo mechanism that will push and pull the latch. The blink detection system calculates the changes in the 'Eye Aspect Ratio' as proposed by Tereza Soukupova and Jan ´ Cech in their paper on real time eye blink detection. Histogram of Oriented gradients(HOG) is used to detect a face, face landmark estimation is used to obtain 68 face landmarks and a convolutional neural network created by FaceNet generates 128 unique 'embeddings' for an image. These embeddings (measurements) are then compared to the 128 measurements of the known image and an appropriate result is produced. For the servo mechanism to be controlled from a remote location, a framework called Flask is used that enables us to control the door remotely.

KEYWORDS: Machine Learning, Convolutional Neural Networks, Blink detection, Raspberry Pi 3.

I.INTRODUCTION

Face recognition systems are essentially computers that can identify a person from an image or a video. With the recent developments in the field of big data analytics and artificial intelligence, there is huge scope for improvement in the field of computer vision and face recognition software. With the widespread developments in the field of AI, computers are becoming increasingly intelligent and gaining human like abilities. Giving human vision capabilities is the next step to make robots more intelligent.

Conventional Methods:

There are two types of recognition algorithms - geometric, which differentiate based on peculiar features, and photometric, which convert an image into values and compare them against template values to remove differences. Some of the commonly used face recognition algorithms are principal component analysis that uses eigenfaces, linear discriminant analysis and elastic bunch graph matching using the Fisherface algorithm.

Conventional face recognition systems are designed in two ways, through holistic features or through local features. The holistic features can be categorised into linear and nonlinear methods. Many applications have shown good results

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

for the linear appearance-based methods such as principal component analysis (PCA) [5], independent component analysis (ICA) [6], and linear discriminate analysis (LDA) [7,8]. However, under different lighting conditions or even different facial expressions, these approaches often fail to correctly identify the face. This is mainly because faces are nonlinear in nature. 'Face patterns lie on a complex nonlinear and non-convex manifold in the high-dimensional space' [1].

Modern technology methods going to be employed:

Artificial Intelligence is the field of science that deals with developing machines that possess human like intelligence. Machine Learning is an application of AI based on the idea that when machines are fed enough amount of data they can learn and deduce meaningful relationships from it. The most efficient learning tool we know is the human brain. Training computers to work like the human brain is the first step towards machine learning. Neural networks fulfil this very objective. They are mathematical expressions denoting how the human brain works.

When we utilise machine learning with AI in face recognition systems, we can improve their accuracy significantly even under varied lighting conditions or partial occlusion. Face recognition is a non-linear concept; hence it is best to use machine learning techniques such as neural networks that can handle nonlinear inputs effectively.

A Convolutional Neural Network is a type of neural network that consists of convolutional layers, a subsampling layer and fully connected layers. CNNs are useful in identifying patterns that have large variance and are robust to distortions and simple geometric transformations like translation, rotation and scaling [6], [7].

II.LITERATURE SURVEY

The field of Biometrics has seen vast improvements in forensics, secured access and prison security. Current techniques include face recognition, handwriting, fingerprints, hand geometry and iris scan. However, it has been deduced that Facial recognition systems are more inexpensive, non-intrusive and extremely accurate. Presently Facial recognition technology is the most reliable out of all recognition technologies.

Conventional approach to location access often involves live streaming of data and hence a need for continuous monitoring of data [2]. Facial recognition algorithms can automate this task and make the process smoother. Face Recognition has also been used as an attendance system. It has been noted that Eigenface method yields better results than Fisherface method. 'Eigenface implemented inside the Attendance System returned between 70% to 90% similarity for genuine face images' [3].

There is scope for development of more robust algorithms that can increase the accuracy of face recognition subject to various conditions. Face recognition often fails when under conditions such as poor lighting, glasses, partial occlusion, different facial expressions and low quality images [4]. This could be improved through advanced techniques coupled with deep learning and augmentation of skin textures and features peculiar to a person (moles, scars, etc.).

In recent times Machine Learning algorithms have yielded better results in fields of object recognition. This experiment aims to utilise a new and robust face recognition system called OpenFace [9] that increases the accuracy of facial recognition systems by employing various machine learning techniques in subsequent stages of face recognition and to implement it in an embedded system environment such as Raspberry Pi 3. It involves four crucial steps [10] – face detection, face landmark estimation, generation of face embeddings (generated via a deep convolutional neural network) and a final classification step to find a possible match.

The face detection step is achieved through the state of art 'Histogram of Oriented Gradients(HOG)' method [5] developed by Dalal and Triggs. It gives far better results than Haar cascade classifiers executed in OpenCV (i.e. the Viola-Jones detectors)

Face Landmark Estimation [6] algorithm is implemented next to center the face in the captured image. The algorithm generates 68 face landmarks that define a person's face. By making sure that the landmark points corresponding to the nose and lips are always situated at the same coordinate location, we ensure that the face is centered. Doing so increases the accuracy in further stages.

Deep Learning is a method that tries to achieve artificial intelligence. Deep learning takes in an input, modifies it through the different layers that exists in the neural network and then provides an output. The network is then modified to provide us with desired output through multiple rounds of training. The third step involves using a trained deep convolutional neural network to generate 128 unique measurements for each face called as 'embeddings' [7]. These embeddings are used to classify images using a Support Vector Machine (SVM) classifier [8].

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

Face recognition is a fast developing field that faces many challenges but also has a wide scope in real time applications. Our experiment aims to utilise a facial recognition system that has been built using new machine learning techniques and implement additional features such as eye blink detection, intruder alert mailing system and its implementation in real time embedded system.

III.PROPOSED METHODOLOGY AND DISCUSSION

The main objective of the experiment work is to implement a Face Recognition System based on Convolutional Neural Networks (OpenFace) in a real-time embedded system such as Raspberry Pi and demonstrate a practical application in the form of Automated Door Access. Further improvements to the above system include an anti-spoofing technique in the form of eye-blink detection program and an intruder alert mail sent to the owner containing a picture of the visitor at the door.

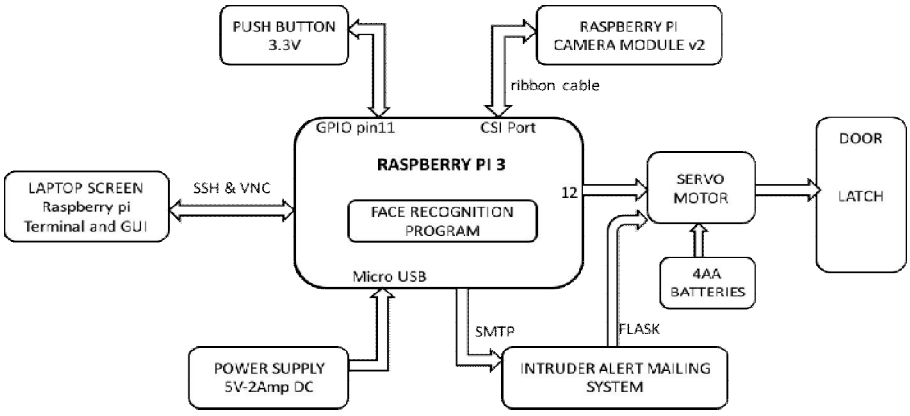


Figure 1: Block-diagram of the proposed methodology

The experiment execution begins when the push button is pressed, triggering the face recognition system. The raspberry pi camera activates to run the blink detection program.

The eye-blink is detected based on the value of 'Eye Aspect Ratio (EAR)'. From the set of 68 facial landmarks, 6 landmarks represent the coordinates of each eye. [11], [12]. The Eye Aspect Ratio is calculated as follows –

$$EAR = ||p2-p6||+||p3-p5||/2||p1-p4||$$

The EAR threshold is usually 0.3 but after trial and errors we derived at an appropriate EAR threshold- 0.28 that suited our experiment needs. The EAR is constantly monitored; when the value of EAR falls below the 'set threshold' a blink is registered and the program execution moves forward. The captured image is then subjected to the face recognition system [10].

The face recognition system called OpenFace [9] includes the following stages - Histograms of Oriented Gradients, Face Landmark Estimation, Convolutional Neural Network from FaceNet and Support Vector Machines. Histograms of Oriented Gradients [5] is used for face detection.The picture initially is converted as a black and white picture. Every pixel in the image is replaced by an arrow called as gradient that points to the direction of the darker pixel. This is to ensure that the same picture differing in brightness does not get identified as a different picture altogether. The whole image containing all the gradients will then look like an outline of a face if face exists in picture.

The HOG algorithm applies a 'sliding window' across the entire image. For every stage of the sliding window, an HOG descriptor for that region is calculated. The descriptor is then compared to a template using a trained SVM to determine if a person's face is located in that region or not. The detection window is 64 pixels wide by 128 pixels tall. The 64 x 128-pixel window is divided into 7 rows and 15 columns, a total of 105 blocks. Each block is comprised of 4 cells, with each cell having a 9 bin histogram. thus the total vector size will be 7x15x4x9 = 3780 values.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

 On HOG -- 

Face Landmark Estimation [6]: This is to account for different poses of the face. It starts with identifying 68 points that are there in almost any face. We then want to identify the positions of these points. These points include the position of edge of eyes and chin etc. After these positions are identified, these points are rotated, scaled and sheared until they are all aligned to the desired and fixed position of these 68 features.
The facial landmark detector based on the One Millisecond Face Alignment with an Ensemble of Regression Trees paper by Kazemi and Sullivan (2014) is implemented in the Dlib library. This method includes:

1. A training set that has the coordinates of each feature and these features are manually labelled.
2. *Priors*, or more specifically, the *probability on distance* between pairs of input pixels.

Such a face landmark detector is made available in the Dlib library. It is used to generate the 68 x-y coordinates that represent the location of specific facial structures on the face.

Convolutional Neural Networks (CNNs) are similar to artificial neural Networks. Many neurons are interconnected and each will receive an input, perform some function on it and then pass on its output to some other neuron.
There are four main operations in the ConvNet [15]

1. Convolution
2. Non Linearity (ReLU)
3. Pooling or Sub Sampling
4. Classification

The main steps in the process of convolutional neural networks are as follows [16]:
Step1: Initialize every filter and weight with some value
Step2: An image from the training database will be passed through all of the layers to provide us with some output probabilities which will tend to be random in nature because of the random assignment initially.
Step3: Calculate the total error at the output layer
Total Error = $\sum$ ½ (target probability – output probability) ^{2}

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

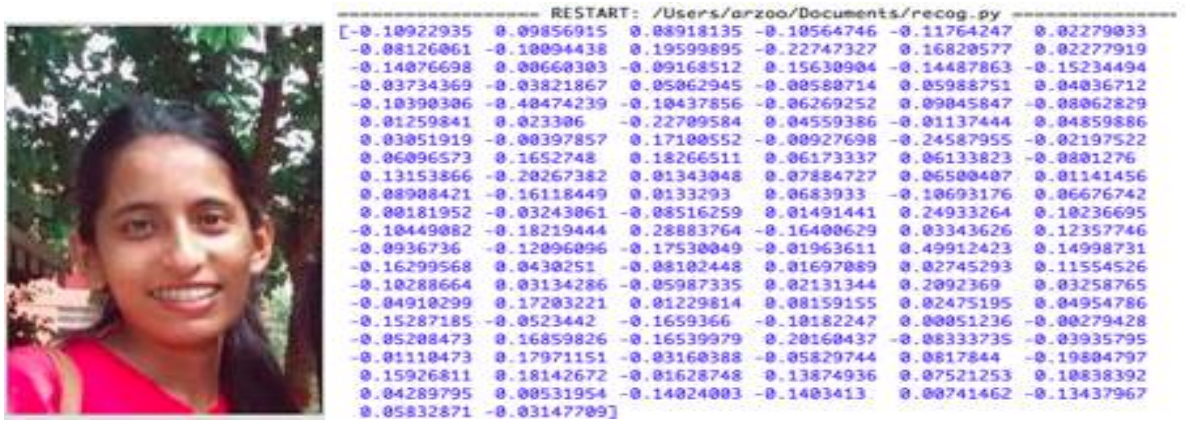
(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

Step4: Here backpropagation is used to calculate gradient of error with respect to the weight values and then update all of the values in such a way so that it reduces error.
Step5: Repeat steps 2-4 with all images in the training set.
Google's FaceNet records an accuracy of 99.63% on the Labelled faces in the wild (LFW) database while using state of the art techniques.

The training process works by looking at 2 pictures of one person and comparing that with another person**.** The algorithm makes changes to the network in such a way so as to make sure that measurements from 2 different pictures of the same person are closer in distance as compared to the measurements of the third person. The above step is repeated several times till the neural network learns to generate the 128 embeddings for captured input image.
'The method is based on learning a Euclidean embedding per image using a deep convolutional network. The network is trained such that the squared L2 distances in the embedding space directly correspond to face similarity: faces of the same person have small distances and faces of distinct people have large distances.' [7].



The final step in the face recognition system involves utilisation of a linear svm classifier. The Support Vector Machine takes in the 128 measurements from input test image and compares it with the 128 measurements of trained test images stored in the database. If a match is found, the svm classifier will return the name of the person who has the closest measurements to that of input image. A support vector machine is a popular machine learning algorithm used for classification and regression tasks. It constructs an N-1 dimensional hyperplane in an N dimensional space. The hyperplane will divide the input data into two categories. The distinguishing feature of SVM is a kernel function that enables it to distinguish nonlinear data. SVMs are known to achieve significantly higher accuracy than traditional methods just after few rounds of relevance feedback.
If the above steps give a positive result and the captured image is recognised by the system, the door will unlock via a servo mechanism. A servo motor connected to the door latch will turn $90^{o}$to unlock the door and after a few seconds will lock the door again. The angle of rotation of the servo motor is controlled by pulse width modulation technique.
A mail is sent to the user in case the face isn't recognised by the system. The native library smtplib is used for sending the mail. Additionally, the email module is also used which makes adding attachments easier. The file is converted into Base64 format before sending it. [14]
Flask is a web application framework. Armin Ronacher is the one who developed it. It is related to the Werkzeug WSGI toolkit and Jinja2 template engine. A micro development framework called Flask is used to control the servo remotely [13].

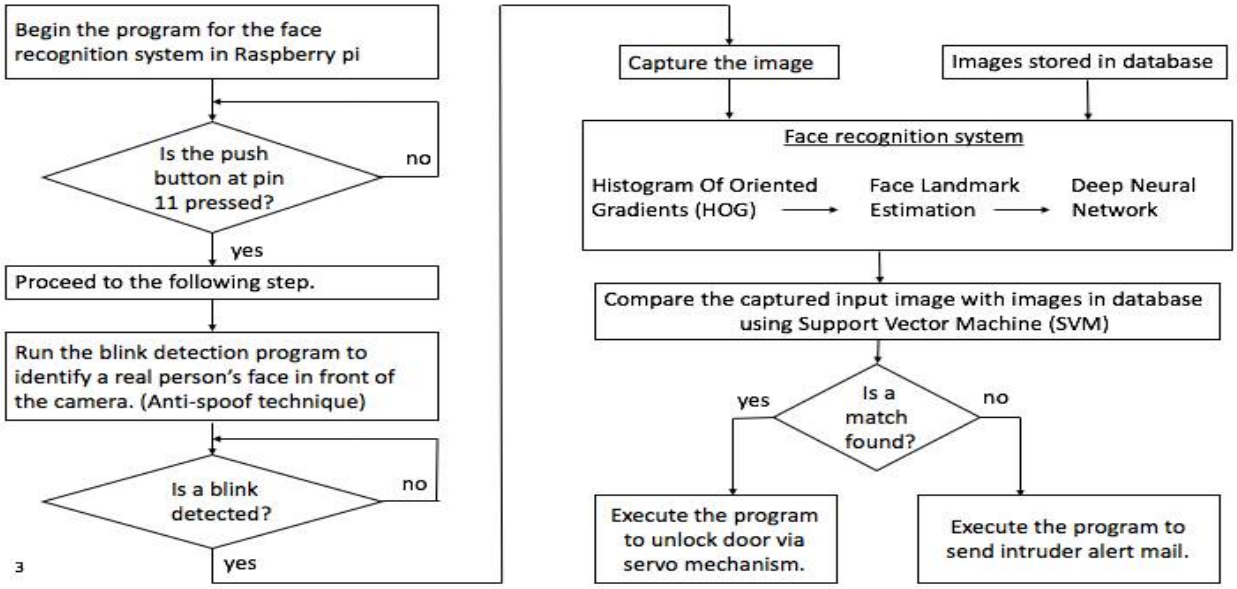
ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

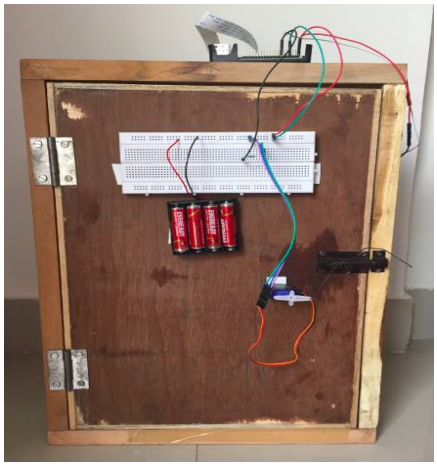
Vol. 6, Issue 6, June 2017



IV.EXPERIMENTAL RESULTS

The results of the experiment are as follows:

- When the push button is pressed the blink detection program runs.
- The blink detection program then shows a frame that displays the EAR. When the EAR falls below the threshold, blink is detected, the face recognition program runs
- Face recognition program takes around 15 seconds to calculate the 128 measurements of the image in the database.
- A picture is then captured with the Rpi camera.
- The face recognition program is then executed and searches for a match from the database
- The door is opened if the captured image is a match from the database. The door is opened with the help of latch that is controlled by servo as shown in figure below.
- Else a mail is sent to the owner with a picture of the unknown person and owner can then give/deny access.



ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

V. CONCLUSION

The implementation of door access via face recognition system using convolutional neural networks in raspberry pi has been successfully achieved. The system gives expected results and accuracy higher than Viola Jones method. Our experiment has incorporated features into the system that make it more robust. The blink detection program serves as an excellent anti-spoofing technique that prevents fraudulent visitors from accessing the system. The intruder alert mailing system further strengthens the security and responsiveness of the system.The face recognition system based on convolutional neural networks was successfully run on an embedded system in the form of Raspberry pi. However, a slight performance drop was noted. The system took 15 seconds to display the results as opposed to 4 seconds in the processor used in general purpose computer systems. This could well be overcome by utilizing the GPU of the raspberry pi to greater extents. Convolutions essentially involve matrix multiplications which require lot of GPU power. Thus, increasing the GPU consumption can increase speed of execution in raspberry pi.

REFERENCES

[1]. Waldemar Wójcik, Konrad Gromaszek and Muhtar Junisbekov (2016). Face Recognition: Issues, Methods and Alternative Applications, Face Recognition - Semisupervised Classification, Subspace experimention and Evaluation Methods, Dr. S Ramakrishnan (Ed.), InTech, DOI: 10.5772/62950. Available from: https://www.intechopen.com/books/face-recognition-semisupervised-classification-subspace-experimention-and-evaluation-methods/face-recognition-issues-methods-and-alternative-applications.

[2]. Md. Nasimuzzaman Chowdhury, Md. Shiblee Nooman, Srijon Sarker, "Access Control of Door and Home Security by Raspberry Pi Through Internet", International Journal of Scientific & Engineering Research, India, ISSN 2229-5518, Volume 4, Issue 1, page nos.(1-9), October-2013.

[3]. Adrian Rhesa Septian Siswanto ; Anto Satriyo Nugroho ; Maulahikmah Galinium, "Implementation of face recognition algorithm for biometrics based time attendance system", ICT For Smart Society (ICISS), 24-25 Sept. 2014, INSPEC Accession Number**:** 14856000, DOI: 10.1109/ICTSS.2014.7013165

[4]. Kavita Devi, Dimple. "Scope of face recognition techniques",International Journal of Computer Science and Information Technology Research, India, ISSN 2348-120X,Vol.3, Issue 2, page nos.(1-5), April-June 2015

[5]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," inProc.IEEE Int. Conference on Computer Vision Pattern Recognition, Jun.2005, pp. 886–893.

[6] V.Kazemi and S. Josephine, "One millisecond face alignment with an ensemble of regression trees", In CVPR, Columbus, OH, USA , ISSN:1063-6919., 10.1109/CVPR.2014.241., Conf. Paper id No., pp. 1867-1874, 2014.

[7] F. Schroff, D. Kalenichenko, J. Philbin, "Facenet: A unified embedding for face recognition and clustering", *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*,ISBN: 978-1-4673-6965-7, doi: 10.1109/CVPR.2015.7298682, pp. 815-823, Jun. 2015.

[8] Cortes,C. & Vapnik, "V. Mach Learn", Kluwer Academic Publishers, ISSN: 1573-0565,doi:10.1007/BF00994018, Volume 20, Issue 3, pp 273–297, September 1995

[9] B. Amos, B. Ludwiczuk, M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016.

[10] https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

[11] Tereza Soukupova and Jan ´ Cech, "Real-Time Eye Blink Detection using Facial Landmarks",21st Computer Vision Winter Workshop Luka Cehovin, Rok Mandeljc, Vitomir ˇ Struc (eds.) ˇ Rimske Toplice, Slovenia, February 3–5, 2016

[12] http://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/

[13] https://www.element14.com/community/community/design-challenges/pi-iot/blog/2016/08/15/pi-control-hubspoke3-key-less-door-entry-testing-the-servo

[14] http://naelshiab.com/tutorial-send-email-python/

[15] http://www.kdnuggets.com/2016/11/intuitive-explanation-convolutional-neural-networks.html/3

[16] https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/