Research Design II

Matthew Micallef
Institute Of ICT
Malta College of Arts Science and Technology
Malta
matthew.micallef.e21498@mcast.edu.mt

***Abstract*—The increase of effectiveness of both Artificial intelligence (AI) and Machine learning (ML) have have changed the field of image processing and image recognition. This paper goes over the implementation and the testing of Local Binary Patterns Histograms (LBPH) algorithm in a face recognition system. With the aim to use this technology to eliminate the need for the use of a physical key to enter a house.**

I. Introduction

The proposed system uses a Haar cascade classifier to detect faces, a Local Binary Patterns Histograms (LBPH) algorithm for face recognition, and a Raspberry Pi camera for real-time video capturing. The ultimate goal for the research carried out is to evaluate the feasibility of using the above mentioned technologies as a cost-effective and secure home entry solution in replacement of a conventional lock and key. The rational behind researching this topic is due to flaws with traditional security measures which include, the need for a key, the chance of the key being lost or stolen or even forgotten in the premise, these flaws all lead to the same scenario, the user being unable to enter the home.

Facial recognition has been emerging more and more as a technology and should be further investigated, in the above use case due to its non-transferability which climates the loss of the key in the traditional security system and also the convenience, the user can never be locked out of the house cause his face, now the replacement key is always available. This paper goes through fundamental aspects such as the number of training images required for reliable facial recognition, more over it goes through the effect of using gray scale images.

The paper positions itself within two major areas of studies, it basis itself on computer vision and also practical home security solutions, it is set apart from the vast majority of studies that research facial recognition by the use of accessible hardware in this case the raspberry pi four and the raspberry pi camera, thus extending to the more cost effective smart home solution rather than a purely academic or high-end applications of facial recognition.

The hypothesis for this research is the integration of a Haar cascade classifier, an LBPH algorithm, and a Raspberry Pi four and Pi camera to potentially create a superior form of entry to unlocking a lock without using a key in the aim of make it being more secure, affordable and user convenient.

A. Research Approach

The research being carried out will follow a systematic approach based on a Research Onion depicted in the appendices [Fig 2]. The Layers to the research onion are explained below.

1) Research Philosophy: The study concentrates on observable phenomena which utilizes a structured and organised methodology and thus follows a Positivism Philosophy.
2) Research Approaches: Due to the study using established theories and research carried out in the domain of facial recognition and the LBPH algorithm, the paper will follow a Deductive approach.
3) Research Strategy: The strategy could be classified as an experiment due to having a devised system and are adjusting a factor which is the quantity of training images to examine the influence on the system's effectiveness.
4) Research Choices: The research adopts a mono-method, by the use of quantitative data gathered from testing the system.
5) Time Horizons: The research is conducted over a specific point in time rather than longitudinal and thus being cross-sectional.
6) Techniques and Procedures: The data being collected involves testing the facial recognition system under different conditions and collecting statistical data. Data that will investigate the correlation between the number of training images and the systems performance.

II. Literature Review

The methodology undertaken in this paper is based on the integration and the implementation of different technologies to identify an alternative home entry solution.

A Haar cascade classifier is a machine learning object detection method which is used to identify objects in an image or video. This classifier is trained based on both positive and negative images of the object to be recognized. Calculable rectangular features are used to detect the presence of the object being recognized in the image, these are also known as Haar features[1]. The cascading aspect for this classifier refers to a number of increasingly complex classifiers that reject negative samples to primarily focus the processing resources on more promising areas of the image.

The Local Binary Patterns Histograms (LBPH) algorithm is used in image processing and pattern recognition due to it being a powerful feature extractor, this algorithm is mainly used for facial recognition applications. This algorithm works by comparing each pixel in the image with its neighboring pixels it encodes the resulting relationship into a binary pattern theses patterns can then be used to create histograms that can be used for image comparison and other recognition tasks.A representation of how the LBPH algorithm works can be seen in the below figure (Fig 1).

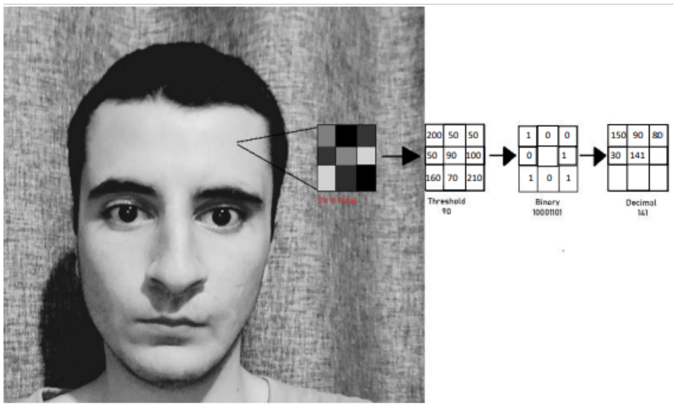

Fig. 1. Conversion of a Gray-scale image to decimal

A raspberry pi is a small and affordable single board computer. It is based on a Quad core 64-bit ARM-Cortex A72 which runs at 1.5GHz and the model which was used in the research carried out was equipped with 4Gb of RAM. This not only classifies the board as affordable but only has good performing hardware.

Their are similar studies that use a raspberry pi for facial recognition, one paper uses a raspberry pi connected to a web cam, passive infrared (PIR) sensor and uses the OpenCV library for image processing, the aim of this study was to use the PIR sensor to identify movement in the room and to trigger the web camera to start transmitting images to a cascade classifier to look for facial features to detect a person[2].

Another research paper went into more depth, two models were evaluated the use of the Haar cascade classifier and also the Histogram of Oriented Gradients (HOG), the evaluation for the Haar model used Adaboost training which selects the best features to form a better classifier. An HOG model uses histograms to represent the distribution of pixel intensities in an image. This algorithm divides the given image into 8 * 8 cells and a histogram for each cell is generated, normalization takes place over a larger 16*16 cell to make the descriptor insensitive to lighting variations. All vectors from each block are concatenated into one vector to form the final HOG feature vector. The latter HOG algorithm explained above was was found to be too computationally intensive for the Raspberry Pi 3 Model B+ and was implemented using a system with an i5 CPU which also deemed to be too computationally intense for [3].

A different research paper uses a similar approach to the one that was undertaken it uses a raspberry pi zero a more affordable model by the Pi foundation and a web cam which although the model is not stated, the resolution is said to be more or equal to 720p. A data set of four individual was created with each individual having 10 images in a different environment, the process of creating a trainer file was done through a python script which loaded the images into gray scale and subsequently into a numpy array. The Haar Cascade frontal face classifier was used to detect a face in the images provided and each image was appended with a respective id. A model was trained using the trainer file generated and the LBPH algorithm being executed on the raspberry pi uses this model for facial recognition. The study also examined the size and loading time for the trainer file by varying the number of images for each of the four users[4].

In both the above studies they use Haar classifier although the research carried out by Singh et al, goes a step further to integrate Adaboost which should help to to construct a stronger classifier than that of Negpal et al, the use of the better Raspberry pi 3 should also make training and facial recognition faster due to the better hardware.

in the study carried out by Mladenova et al, a similar approach was taken with the aim to open locker doors in an airport by means of a raspberry pi 3 and the use of Principal Component Analysis (PCA) as a dimensionality reduction algorithm, this was done in order to simplify the the process of comparing facial images. The algorithm treats each image inputted as a one dimensional feature vector created from the pixel values of the image. PCA reduces the non-informative parts while preserving the most informative aspects of the image this is done to reduce the dimension of the feature vectors. In facial recognition, the use of PCA is known as the Eigenfaces method this involves representing the faces as a smaller collection of the key features, these are said to be the principal components of the collection of face images, this makes the comparison between the faces slightly more computationally manageable[5].

When comparing the three main approaches above two of the papers follow a similar flow by the use of a combination of a Convolution Neural Networks (CNN) and Recurrent Neural Networks (RNN) for face recognition, this approach basis itself on deep learning models which are better able to learn complex patterns in the input data. CNN helps to adaptive and automatically learn spatial hierarchies of features, while RNN helps in learning temporal dependencies in the data. Although this method of facial recognition gives high accuracy it requires a substantial amount of computational resources and a lot of training data. On the other hand the research carried out by Mladenova et al, uses PCA for feature selection and dimensional reduction. Although this approach is less resource intensive in terms of computational resources, this approach could lead to lower accuracy when compared to the above two research papers discussed above more particularly when it comes to complex patterns in the data set.

The following Conference paper written by Nikita et al, uses a different hardware approach that the four other studies discussed above it uses an Arduino based hardware solution rather than a raspberry pi but ultimately they are both affordable single board computers. The implementation captures a video input stream and uses a motion detection module that is written in MATLAB. if motion is detected the frame that motion was detected on is passed through to a facial detection module also written in MATLAB. if this module identifies facial features, the coordinates of the face and also the frame is sent to a facial recognition module, which uses the coordinates to extract the face from the frame. The following toolboxes are used in terms of software imported from MATLAB, the Computer Vision Toolbox, the Statistical Toolbox and the image acquisition tool box[6].

The facial recognition process is based on PCA using the Eigen faces method similar to the approach taken by Mladenova et al, where the system creates a training set of face images, then separates the frames down to individual vectors, A covariance matrix is formed, from which the eigenfaces better known as eigenvectors are derived. The system focuses on the key facial attributes rather than the entire facial data set to be able to best determine the weights of each frame that it is processing. it than uses the weights of the new face image and compares it the weights of the images that it has stored in a database.

The Euclidean Distance (ED) method is used for classification, with the maximum threshold set to 4.00×10^{15}. If the ED difference is greater than the above set threshold the classifier will fail to recognize the individuals identity. The above classifier is given eight images of the user to be trained on compared to the ten images that Negpai et al used.

All the implementations given above use a different method and algorithm for facial recognition the first three implementations relay on various deep learning techniques with different loss functions and different data preparation stages. The last two implantation's use a simpler yet still effective method of facial recognition by means of PCA. The dependent factor on choosing any of these implementations relays on the specific computational resources that the system has to offer that the margin of acceptable error rate.

A Literature Map is attached in the appendices section [Fig 3] to gather insight of how research was mapped out.

III. Methodology

To address the hypothesis these objectives and research questions where identified.

Objectives

1) To dynamically create a dataset based on facial features.
2) To implement an efficient face detection mechanism using a Haar cascade classifier.
3) To apply an effective facial recognition algorithm using LBPH.
4) To determine the optimal number of training images for high accuracy.
5) To evaluate the feasibility and safety of the technology for practical use.

Research questions

1) What kind of data set is required for detecting faces?
2) What algorithm can be used to best distinguish between faces?
3) Is there any difference in the confidence levels when increasing the number of images taken for each new individual?
4) Is it safe to adopt this technology for the above mentioned use case?

The following subsequent stages will be used to carry out the execution of the research paper.

1) The first stage of the Development process would be to gather the initial user data, the data that will be used to train a model with. A mechanism for the user to enter their name is created and a sequence of images is subsequently taken.
2) The images are then collected converted to gray scale and saved in a folder structure using a new given id
3) A Haar cascade classifier will then be used to identify weather their is a face in the image or not and the image will than be cropped and the original image taken will be overwritten by the cropped image of the users face.
4) An LBPH algorithm will be used to train a model based on the images that were gathered and processed.
5) The model will then be automatically used to start recognizing users that are in front of the raspberry pi camera.
6) The model will then be tested with different amounts of training images to understand the correlation between the number of images taken for each new individual.
7) A comprehensive review of the results should be undertaken to identify weather or not this implementation can be deployed as a product, due to the known potential for high false positives/negatives that the LBPH model can generate.

A. Data Set

In the context of this implementation, their is no need for a pre-existing data set, the reason being is that the model training is being based on the individuals who have been granted entry to the home. As mentioned above whenever a new user is allowed entry to the home, a number of images are taken and stored the latter being that the images contains a face, upon the images being saved in the particular structure the model will get retrained and the new user will now be recognised by the model in conjunction with the existing users that the model has already been trained on. The above described operation can be executed effortlessly, in the scenario that a user is no longer granted entry to the home, the images of that particular user can be purged and the model can be retrained and as a result the model will no longer recognise the user.

B. Face Trainer

To asses the presence of a face in the provided image, the Frontal face Haar classifier was chosen for it being fast in execution, having a high accuracy, and minimal false positives. This classifier is trained using a number of different images which includes images with a face and without, different skin tones, identifying eyes and facial structure.

C. Facial recognition

The Open CV library was chosen for it being open-source and has three built in facial recognition algorithms which include the Eigenfaces, Fisherfaces and the LBPH. The LBPH algorithm when compared to the above mentioned algorithms can not only recognise the front of the faces but also the sides of the face which makes this algorithm a better option in this particular use case.

D. Number of images

The number of images used in the initial stage of this implementation will be tested to see the affect of the number of images that the LBPH algorithm is trained on, this would be done to identify a number of images which offer a high accuracy without sacrificing storage space, also with reducing the number of images that the model is trained on the retraining of the model will be faster given the less images that the model needs to train on, both for existing users and the user that is being added.

E. Research Philosophy and Approach

The research follows a positive philosophical approach given that Facial recognition is in essence a quantitative problem, involving precise measurements of facial features and characteristics computed through mathematical formulae. Based on this approach the research utilized the deductive approach to establish a hypothesis and test it through the methodology described.

The experimental research strategy was used to focus on the correlation between the number of images used for training and the resulting model's accuracy. The mono -method will be used, relying mainly on quantitative data to potentially provide precise, numerical data which is beneficial for statistical analysis in the field of facial recognition.

The above research philosophies were used to fulfil the research objectives outlined above, but to also provide substantiated knowledge to implement facial recognition technology in the given context.

F. Method of Analysis

The performance of the facial recognition model will be assessed based on the confidence level it exhibits when identifying individuals, this method was chosen to mimic the intended task in correctly identifying individuals. By using the confidence level, the model can be assessed based on the correct identification but also on the certainty that the model has in the identification.

G. Ethical Considerations

Due to the nature of facial recognition and the need for collecting and analysing potentially sensitive data these ethical considerations are being taken into consideration

1) Informed Consent: Before any data is collected the user should be informed and given the option to accept or deny the capturing of their images for the use case.
2) Data Protection and Privacy: To ensure data anonymity the user data will be stored based on an id and security measures like data encryption are used.
3) Compliance with Regulations: The research will comply with all the relevant data protection regulations such as the General Data Protection Regulation (GDPR) in the European Union. Meaning that the data collected will be processed only for the process of facial recognition and will not be retained for more than necessary.

IV. Results

The implementation followed the methodology as described above, an observation was made that when the image being processed through to the Haar Classifier and the LBPH algorithm the algorithms performed better when converted to gray scale, by means of reduction in the number of false positives, this was also identified by the four studies that used a similar pipeline which, are mentioned above in the literature review section.

The conversion of the image to gray scale was also carried over to when the live video stream feed was being captured at 24 frames per second through the Pi camera, the facial recognizer was given gray scale images rather than the full color images.

With regards to the research questions identified in the methodology stage, a data set was not required given that the users that need to be recognized are being added and trained on on initial set up. The LBPH algorithm was chosen as the facial recognition algorithm this choice was based on research and analysis carried out from previous studies, by other researchers, respectively cited in the research methodology phase.

Testing the number of images to identify weather increasing the number of images used to train the model returns a higher confidence level or not was carried out, as follows in Table 1, where subject refers to the number of users the model was trained on, the number of images per user and the average confidence level for 10 frames. From the above table we can see the Haar cascade classifier performed perfectly in identifying faces in test case one, with a confidence level that

Test case	Subject	Number of images per subject	Average confidence level for 10 frames
1	1	1	68.28
2	1	10	81.973
3	1	20	87.40
4	1	40	88.89
5	5	1	84.13
6	5	10	67.24
7	5	20	75.70
8	5	40	68.32

TABLE I
RESULTS

improved by an average of 13 when the number of images increased from one to ten. Although there was a marginal increase in confidence from test case three to four, it was not significant, showing that 20 images are sufficient for a face identification confidence of around 87. Further testing involving four additional people (five users in total), revealed varied results. The highest confidence rate was in test case one despite some false positives. Test case six revealed a drop in confidence and an increase in false positives when each user was trained with 10 images. Test cases seven and eight showed inconsistent results, with test case seven having a higher confidence rate and lower false positives than test case eight. Repeated trials confirmed these outcomes. In conclusion, using 20 images per user is optimal for confident face identification.

A comparison into a similar study carried out by Delbiaggio [7] revealed divergent outcomes, in his research, four facial recognition algorithms were tested, and it was found that using 10 images per subject led to a higher false positive rate than 20 or 40 images. Contrarily, this study recorded false positive, even with 20 or 40 images. The difference may be attributed to Delbiaggio's use of a high-quality Nikon D3100 camera, while this study used a Pi camera, which offers comparatively lower image quality. The computational platform could also play a role in the imbalance as this project used a System-on-Chip (SOC) device, which is less powerful. Furthermore, this study used a video stream for recognition, while Delbiaggio used still images.

A severe security concern arose when A test case revealed that an image of a person, displayed on a mobile device, was not only recognized as a person by the Raspberry Pi but also displayed high confidence rates.

V. Conclusion

Further investigation is essential to devise a method for distinguishing between a live face and an image, either on display or printed. Moreover, comprehensive testing in varying lighting scenarios - such as in the dark or under intense sunlight - must be carried out prior to implementing this system.

A better camera could be used to have more resolution in turn detail for the LBPH algorithm to train on. Research would be carried out to identify whether a deep convolutional neural network would have been better suited for this use case.

While the current implementation may not be viable as it stands, there is potential for future enhancements. Given its inherent potential, it's possible to develop a cost-effective product with this software that could compete favorably in the market, despite other more expensive alternatives.

Appendices

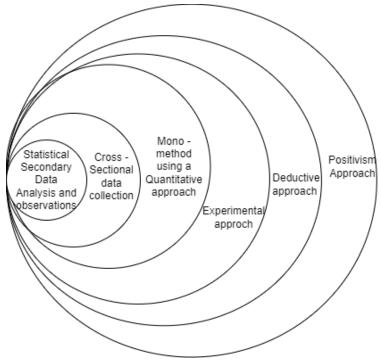
Statistical Secondary Data Analysis and observations
Cross - Sectional data collection
Mono - method using a Quantitative approach
Experimental approch
Deductive approach
Positivism Approach

Fig. 2. Research Onion

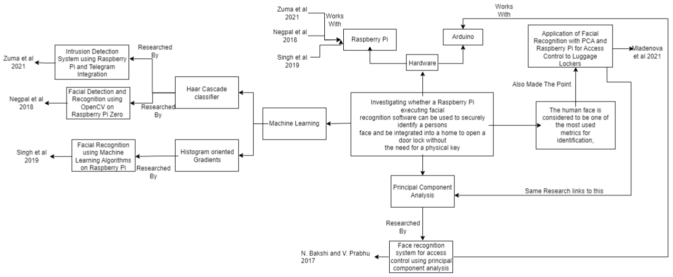
Zuma et al 2021
Works With
Negpal et al 2018
Singh et al 2019
Raspberry Pi
Hardware
Arduino
Works With
Application of Facial Recognition with PCA and Raspberry Pi for Access Control to Luggage Lockers
Mladenova et al 2021
Also Made The Point
Zuma et al 2021
Intrusion Detection System using Raspberry Pi and Telegram Integration
Researched By
Negpal et al 2018
Facial Detection and Recognition using OpenCV on Raspberry Pi Zero
Researched By
Haar Cascade classifier
Machine Learning
Investigating whether a Raspberry Pi executing facial recognition software can be used to securely identify a persons face and be integrated into a home to open a door lock without the need for a physical key
The human face is considered to be one of the most used metrics for identification,
Singh et al 2019
Facial Recognition using Machine Learning Algorithms on Raspberry Pi
Researched By
Histogram oriented Gradients
Principal Component Analysis
Same Research links to this
Researched By
N. Bakshi and V. Prabhu 2017
Face recognition system for access control using principal component analysis

Fig. 3. Literature Map

References

[1] López, Laura Sánchez. "Local Binary Patterns applied to Face Detection and Recognition." (2010).

[2] Mfundo Zuma, Pius A. Owolawi, Vusi Malele, Kehinde Odeyemi, Gbolahan Aiyetoro, and Joseph S. Ojo. 2021. Intrusion Detection System using Raspberry Pi and Telegram Integration. In Proceedings of the International Conference on Artificial Intelligence and its Applications (icARTi '21). Association for Computing Machinery, New York, NY, USA, Article 5, 1–7. https://doi.org/10.1145/3487923.3487928

[3] IS. Singh, R. Ramya, V. Sushma, S. R. Roshini and R. Pavithra, "Facial Recognition using Machine Learning Algorithms on Raspberry Pi," 2019 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, 2019, pp. 197-202, doi: 10.1109/ICEECCOT46775.2019.9114716.

[4] G. S. Nagpal, G. Singh, J. Singh and N. Yadav, "Facial Detection and Recognition using OpenCV on Raspberry Pi Zero," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2018, pp. 945-950, doi: 10.1109/ICACCCN.2018.8748389.

[5] T. Mladenova, I. Valova and N. Valov, "Application of Facial Recognition with PCA and Raspberry Pi for Access Control to Luggage Lockers," 2021 International Conference Automatics and Informatics (ICAI), Varna, Bulgaria, 2021, pp. 141-145, doi: 10.1109/ICAI52893.2021.9639574.

[6] N. Bakshi and V. Prabhu, "Face recognition system for access control using principal component analysis," 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2017, pp. 145-150, doi: 10.1109/INTELCCT.2017.8324035.

[7] Delbiaggio, N., 2022. A comparison of facial recognition's algorithms. [online] Theseus.fi. Available at: ¡https://www.theseus.fi/handle/10024/132808¿