CCTV Intelligent Surveillance on Intruder Detection

Kajenthani Kanthaseelan
Dept. of Information
Technology SLIIT Malabe,
Sri Lanka

Paskaran Pirashaanthan
Dept. of Information
Technology SLIIT Malabe,
Sri Lanka

Jasmin Jelaxshana A.A.P
Dept. of Information
Technology SLIIT Malabe,
Sri Lanka

Akshaya Sivaramakrishnan
Dept. of Information Technology
SLIIT Malabe,
Sri Lanka

Kavinga Yapa Abeywardena
Dept. of Information Technology
SLIIT Malabe,
Sri Lanka

Tharika Munasinghe
Dept. of Information Technology
SLIIT Malabe,
Sri Lanka

ABSTRACT

With the technology today, the detection of intruders was introduced to decrease theft and crimes. The implementation of this system is made at home. A system with just records of theft would not work, the system should be efficient and faster to detect thefts effortlessly. The goal of this system is to give a new direction of innovation in the closed- circuit television video image processing in computer vision. Having a system to monitor and notify users of potential threats must be implemented, as well as predicting a child's movements at home. Doing this in real-time ensures a lot of human labor to be minimized while also ensuring recognition using computer vision. This document provides an outline of the intrusion detection system and danger prevention system for children outside and inside the home environment, with the help of image processing, face recognition, and training of the model through machine learning. For an instance, if a kid at home reached near the well in the garden area, system predicts whether the kid is reaching the well by deriving it from the distance calculated earlier and sends an alert to the mobile or any device where an application is installed. Same prediction applies for any intruders to reach the home or any restricted places at home.

Keywords
Intrusion, image processing, face detection, rule-based notification

1. INTRODUCTION
Surveillance is the monitoring of humans for behaviors and activities to protect them. Humans did it manually in the early days, which was a tiring task as suspicious activities were uncommon compared to usual activities. In recent years automated surveillance has been increasing with the development of technology and the identification of new techniques and concepts. Crimes and disasters that occur because of the negligence of humankind and the rate of it appearing are having a drastic increase in recent times. With intelligent CCTV (closed-circuit television) surveillance, it introduced various approaches to analyzing it. There are many approaches [1] used and below mentioned are some of them,

• Object tracking

• Background subtraction

• Suspicious activity detection – intruder detection

• SVM

• Bayes decision theorem

• Sequence models

• Multi-frame differencing method

Using surveillance cameras is on the rise [2]. It is playing a crucial role in the security sector. CCTV cameras record vast amounts of footage. Therefore, it is impossible to check them all in actual time. It is also expensive to find a human resource to monitor all footage. To overcome this, intelligent CCTV surveillance systems are being developed to recognize situations in actual time and respond [3]. According to an intelligence CCTV surveillance application, the most basic functionality is to identify people. There are many methodologies measured over time with its strength and weakness. These ancient people detect methods that have some common drawbacks. The main motivation of this research work is to provide an economical, and efficient security solution for households.

• To detect images accurately for intruder detection purposes.

• To train the model for identifying between familiar and intruders.

• To implement a rule-based notification system for predicting movements of intruders and kids at home.

• To implement an accurate facial recognition algorithm for intruder detection purposes.

With the development of the social economy and the improvement in the quality of life, family safety has become one of the most important aspects of life. Therefore, Intelligent CCTV surveillance is a vital concern nowadays. Since human actions and abnormal behaviors are highly reached a part in computer vision. The authors propose several systems to ensure a person's safety in different situations. Increasing physical security by appointing someone will not result in being efficient to all, which is why the authors proposed a proper detection system. The top priorities of this system are to maintain access control to a secure area, monitor the perimeter for intrusion and the personalities, predict a child's movements before something goes wrong, and watch for suspicious behaviors.

2. METHODOLOGY
With the increasing rate of modernization of the world, the rate of threats occurring is drastically increasing. Thus, everybody has an extreme concern regarding our security everywhere.
• Physical security around the residential area and keeping account of people visiting and identifying the intruders is a

major issue identified. Increasing physical security by appointing a security guard will not result to be effective and feasible to all. Thus, an automated detection system is suggested by the researchers as an effective solution.

• Children's safety is also another issue identified. There have been many instances which had led to the loss of children's life because parents nowadays are busy with work and spend less time with children. Thus, to provide alert and monitor children from anywhere, a system is suggested to be introduced.

2.1 Intruder Detection

In this phase, it analyzed the background study and research domain. Understanding current processes and similar implementation was done. Understanding the domains that the authors would cover was a challenge as the scope is huge. Authors who are involved in the respective module also collectively gather information on ideal implementations for the problem at hand. In the application's case, the authors also conducted user experience surveys to find out the most appropriate features. After completing the surveys, the authors moved forward by aggregating the results of it. Also, consider peer review authors who completed the features of the application. It gives some important steps which the researchers followed below,

• Collect related research papers.

• Background study in domains.

• Feasibility study.

• Financial viability.

• Read and analyze research papers.

• Collecting information from peers and analyzing their perspective.

• Filter out features and components.

• Complete the gathered requirements

The fundamental challenge in this proposed system was for the dataset, which the authors had to create the whole dataset from scratch. The authors used two different kinds of datasets to identify a person whether the person is known or unknown. The dataset should contain real data which is collected from daytime through the CCTV footages. Dataset is our main reference for our testing and training processes. Researchers transform face images into a vector to reshape, then split the data into training and testing (80% for training and 20% for testing).

The utmost level of implementation of the system is face detection. This research focuses on implementing a face detection system that could uniquely identify a person by analyzing patterns based on human facial textures and shapes and capable of detecting and recognizing multiple faces in live frames. It includes skin-like regional segmentation, facial feature extraction, and face image extraction of a facial candidate. Then, this will create a mathematical formula which is called a facial signature. For face detection, researchers used PyCharm implementation of the Haar Cascade classifier model and used a Haar cascade algorithm for feature extraction. It trains this face detection on the Haar cascade GUI dataset, which is an algorithm dataset where the detector can learn to generally detect faces. The model will be comparatively fast, with low positive rates and an accuracy level. This is the reason to select this algorithm for detection.

It mentions a comparison of HAAR and LBP in Table 1,

Table 1. HAAR versus LBP

Algorithm	Metrics			
	Detection Accuracy	Computation Complexity	False Positive	Robustness to different lighting conditions
HAAR	High	Complex and slow	Low	Less robust
LBP	Low	Simple and fast	High	Highly robust

Facial detection is a bio-metric artificial intelligence information-based application that can identify and verify a person from a video frame on a CCTV and is also widely used in the home environment. Making this possible in actual time ensures human security to be minimal while also confirming efficient recognition of intruder activities. This solution proposes a method to detect faces of humans and draw bounding boxes around them that will then be recognized using image processing, which helps in cropping the image from the detected face's video frame and applies enhancements in preparing for training the model.

In the last decade, there are different methods introduced that are based on Artificial neural networks and evolutionary algorithms for CCTV image processing. The authors' proposed system also focuses on CCTV image processing.

There are many basic image processing techniques to deal with. Considering all these ancient people detect methodologies, they have some common drawbacks. Some could be the identification of false classifications. Since it could be unreliable most times [4]. However, there have been new advancements in this area such as Faster (convolutional neural network) CNN, which is the most accurate single model [5]. For system design pre-processing, conversion to grayscale, background subtraction, image segmentation, image enhancement, object detection, the region of interest is to be considered. The initial step is of image processing is to pre-process the image where the image is being suppressed on unwanted distortions and focuses on the important features of the image. It helps in further processing the image. The next foremost classifications are image segmentation and image enhancement. Image processing is implemented by the CNN algorithm.

After making the image easier to process by the conversion of color, the next step is to subtract the unwanted background image related to the authors' domain area of the image. If the authors' goal is to detect the person's face for intruder checkup, then there is no need for background image to be further processed. For the image, after the face detection phase, the researchers extracted features from the face. The researchers will compare it with vectors for other detected faces. This Support Vector Machine (SVM) Classifier model gets face vectors/features as input, and it will predict the id of that face as shown in Fig. 1. SVM is a supervised machine learning algorithm to use models that divide and classify data. The accuracy of face recognition is highly dependent on the quality and variety of sample images. The variety of sample images in the sense, by capturing multiple images with multiple directions for the same face. In identification, the algorithm is presented with an image of an unknown person. The algorithm reports its best estimate of the identity of an unknown person from a database of known individuals. In a more general response, the algorithm will report a list of the most similar individuals in the database.
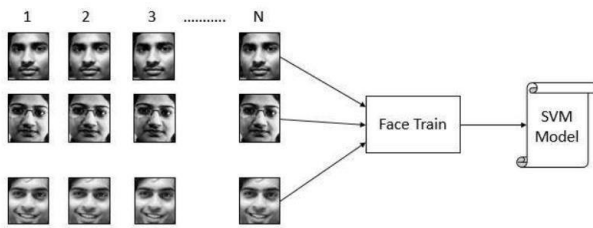
Fig. 1. Block diagram of face training. [9]

The Face Net model is the best model for face classification and identification purposes. It will generate embedding vectors for an image of faces. Researchers can predict face class, which means what is the name of that person and the probability of that prediction. As per the analysis conducted, authors identify that such instances mainly occur because of manual errors. Thus, to provide security to protect the needy by keeping track and monitoring the movements of people to identify intruders builds the system. Some Face recognition methods differentiate by facial features from one person to another person. In terms of identification, it gives the image compared with stored images in the database, and if it is positive return the Id or name of that class. To address the possibility that a face belongs to an unknown identity [6]. Authors trained many images of strangers or intruders (these do not belong to that housing member). After the training phase, an unknown face or the face not belonging to it classifies any other classes as strangers. All the datasets are preprocessed. This dataset was gathered in real-time faces using CCTV cameras. The verification approach uses the following steps:

• Get pre-processed dataset/images from the stored database.

• Randomly split the extracted images, 80% for training and 20% for testing.

• Using the training dataset to get the best accuracy model.

• The result is compared with the actual labels.

To evaluate the models, the second approach uses the same dataset (with trained dataset) to comparing predicted values with the actual one. This process is using a support vector machine, the test dataset is predicted.

The model can differentiate the child between the adult faces more accurately through the face recognition feature comparison. This recognition model will categorize two age limits as under 15 or above 15 that will detect that through the human faces which appear in CCTV or video frames. This feature helps the child safety alert notification system to provide safety for children. Such a solution reduces human effort significantly, as authors can filter down to segments of the footage.

This child's facial recognition can differentiate between the child and the adult faces and recognize the child faces more accurately through the face recognition feature comparison. This feature helps the child's safety alert notification system to provide safety for children. Hence this prediction model is designed to capture the moving trend of the child with the help of the Scented Kalman Filter algorithm, which shall run to predict the movement of the child. This model will detect the movement of the children, who were in the video frame, and no other human faces. This recognition model will categorize two age limits like under 15 or above 15 and, it will detect that through the human faces which appear in CCTV or video frame. The below cartoon Fig. 2, will show the structure of the model,
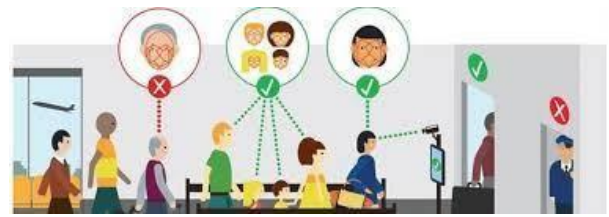

Fig. 2. Child recognition

Users of this system can view and operate the CCTV via a mobile application. The mobile app includes live streaming video, intruder alert notifications, accepting or avoiding options for intruder-relatives, and adding relative face options. Using this system provides an added security measure in which the user can decide the detections of the face recognition model to be an intruder. If the user accepts the intruder, the user can give an alert to the police or someone. If the user avoids the intruder, the user can add this intruder to the database as a known individual. It builds this system in a way that it would detect the movements [7] and capture the face of the person to conduct classification to identify the intruder.

2.2 Child Movement Prediction and Rule-based Notification

The goal is to design a model that can accurately predict several probable future movements of a child, by calculating, as in (1), the multimodal distribution

$$P(Lh+1{:}T \mid L1{:}h, V1{:}h) \tag{1}$$

Where, Lh+1: T, is the order of the past video frames. V1:h and agent locations, L1:h, where, h is the history length, and T–h is the prediction length. The ground truth location history and the set of video frames that are preprocessed are the inputs to this model. History Encoder is used to encode the input and the output is fed to the RNN decoder to predict the location which outputs the heatmap. A combination of those generates a multimodal distribution for predicted locations.

The observed trajectory from time 1 to h and the future child trajectories from time h + 1 to T are trained in this model.

There have been many systems that detect movements of children however no system to predict the movement prior and notify the relevant parties. This proposed system investigates all aspects where the movements of the child can be predicted by using relevant algorithms by calculating distances to hazardous places inside and outside the home environment. The rules are being declared already to predict the movement of kids at home, thus ensuring the safety of the child inside the home premises. Different dataset of video frames has been considered and the most optimal algorithm has been chosen.

One method calculates the statistics of a random variable that undergoes a non-linear transformation. This algorithm also addresses many of the approximation issues of the Extended Kalman Filter and constantly achieves a better level of performance at a comparable level of complexity. For an instance, if a kid at home reached near the well in the garden area, the system predicts whether the kid is reaching the well by deriving it from the distance calculated earlier and sends an alert to the mobile or any device where an application is installed. The same prediction applies to any intruders to reach the home or any restricted places at home. The below-

mentioned pre-requisites should be filled to make this notification system function. It should implement a prediction algorithm and a detection classification algorithm.

Rule 1

Using the detection and classification algorithm, the faces will be captured and classified into "known" and "unknown" categories. Once any of the faces captured fall into the "unknown" category, an intruder detection [8] notification would trigger the mobile application. Fig. 3 below explains how the hardware equipment is connected and functions.
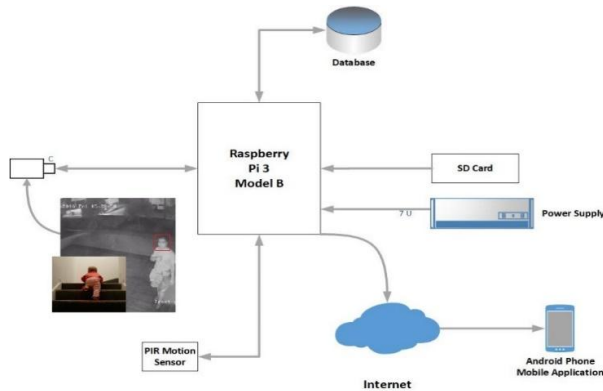
Fig. 3. System diagram of the rule-based notification system

Rule 2

Analyzing experiments through the identification and verification show that the researchers achieved excellent results. This paper ensures a secure environment with intelligent CCTV systems. Using the prediction model, if it detects any movement, this notification system will start functioning. Once it recognizes any movement, a snap of the moving face will be taken and compared to classify it into either "known" or "unknown" face. If it detects an unknown face, it will discontinue the prediction process. Else, if the captured face falls into the known face category, then it will classify this into "above 15" or "below 15" age categories. If the category is only "below the age of 15" then the Scented Kalman Filter algorithm shall run and find the trend of movement and alert the relevant parties, respectively. Below Fig. 4 elaborates on the child movement prediction process using a flow diagram, explaining all the conditions and optimal algorithm used.
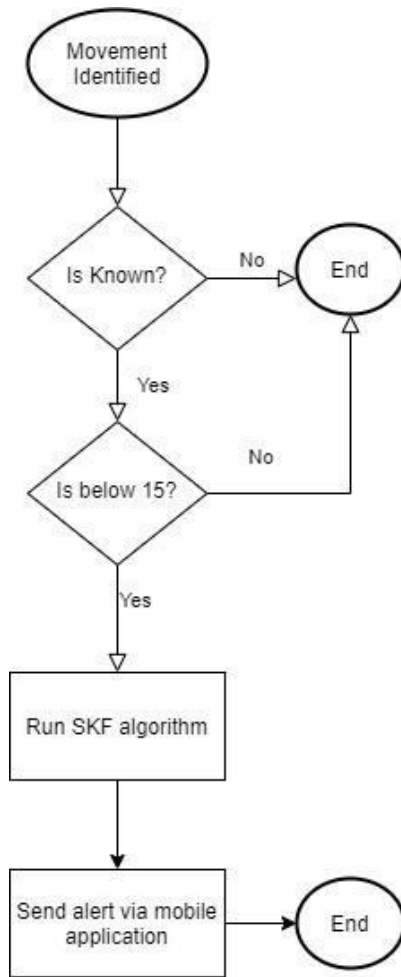

Fig. 4. Child movement prediction notifications' process flow diagram

Applying both these defined rules, the proposed system works as expected. Physical security around the home environment as well as the child's security inside and outside the home environment is maintained by applying the methods proposed. It creates this system intending to derive the utmost protection to children who are not under the direct surveillance of their parents. Hence, this system is also built to predict the movements of a child and send an immediate notification to the relevant parties.

3. RESULTS AND DISCUSSIONS
Various approaches were demonstrated and tested. Researchers came up with the optimal algorithms and research findings in proposing the desired system. Image processing and classification were done by the CNN algorithm and reached the optimal results, enhancing the cropped face-detected images as shown in Fig.5 and classifying them.

Fig. 5. Image Enhancements, before and after

The face detector's performance was evaluated using the test data as the researchers stated before. This detection model's training accuracy was 90.1% as well as the testing accuracy was 88.2% as per the split of data used to train the model. Detecting a face in a video frame of home surveillance footage is sufficient that the researchers focus on the classification part of the face detector rather than the localization of the detected face (bounding box).

To fine-tune this model to the researchers' specific task, the researchers trained it on the researchers' accumulated dataset of faces. Here, researchers used 400 positive images and 400 negative images to train the model. This model will detect faces in any situation in day mode. Some features are below,

• Single/Multi face detection.

• Side view face detection.

• Detect the face within 20 meters distance.

• It will detect human faces even; they wear masks, helmets, cooling glasses, spectacles.

• This model will only work in daytime
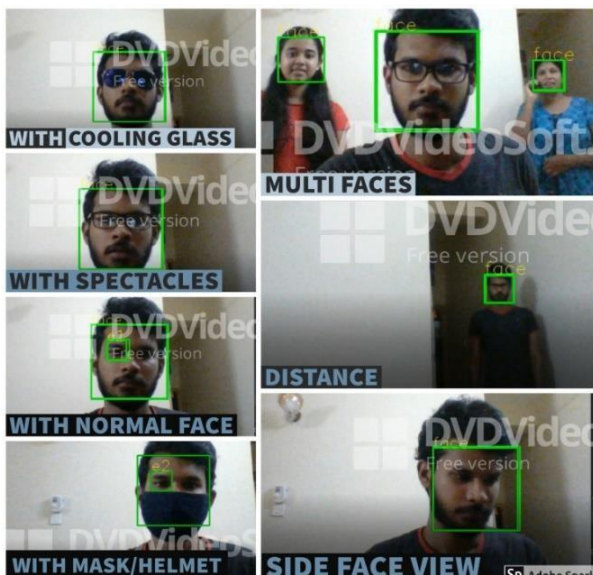
These features are shown in Fig. 6 below,

Fig. 6. Face detection sample results testing all features.

The model gives the authors' best accuracy as it should eliminate over-fitting scenarios from its performance. Based on the evaluation's methods, select an exemplary model to work within, to identify such a model [10].

Training the model was tested using the SVM algorithm, and

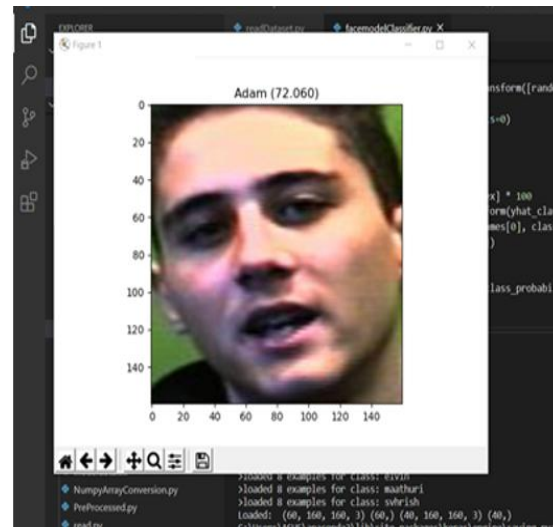the best results were achieved resulting as shown in Fig. 7 below,

Fig. 7. Face recognition sample results

Face vectors are used in recognizing the face of people and the classified dataset of faces from image classification is used to train the model successfully and finds the intrusion. After running the proposed solution, with a dataset of 400 images, there was a classification accuracy of 96%, resulting in some wrong images. But, the epochs, that define how many times the training will repeat, were reduced from 50 to 10 to get better results in a short period of time. And the results of the testing are shown in Fig. 8 below, referring to validation accuracy and dependence of accuracy on the epochs. This defines that the model can be able to generalize any new data being added.
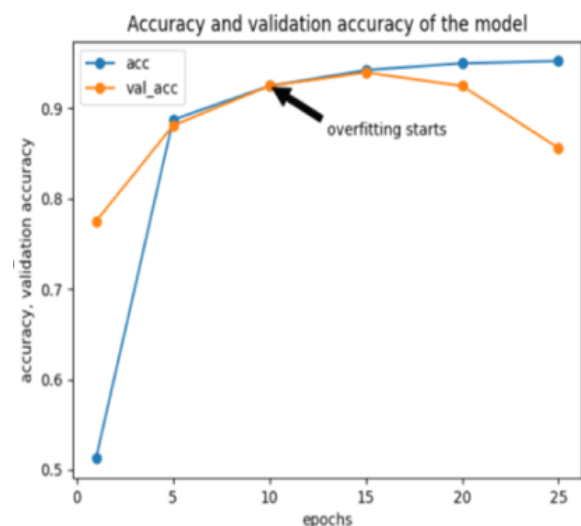

Fig. 8. Accuracy and validation accuracy of the model.

As well as the messages would be received at IoT Hub from the connected Raspberry Pi [11] sensor device as shown in Fig. 3. Then this IoT will be integrated into the mobile application for notifications. Moreover, the predictive model is shown in Fig. 9 is the two sets of outputs received. The object information is displayed on the frames.
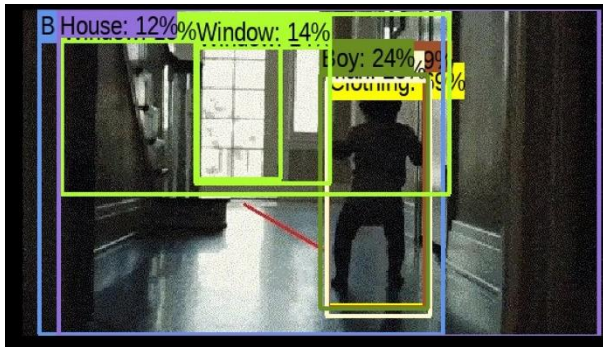
Fig. 9. Object information for rule-based system

4. CONCLUSION

In this paper, researchers have proposed the intelligent system for recognizing people's faces including children and adults, also getting notification from the system when intruders have been identified or by predicting the child's movements and notifying when they reach hazardous or restricted places at home. Researchers have discussed image processing techniques and facial detection algorithms and achieved with the optimal solution to extract facial images from the CCTV live streaming videos, enhance the recognized images, classify the enhanced images, and train the model effortlessly. Additionally, child prediction and the notification-based system is too built with the optimal use of algorithms and hardware equipment. The system uses data from a real-time environment.

Implementing this system reduces work for humans and provides an effective and faster intruder detection system as well as guiding children at home for busy parents. Research areas in image enhancements, child's movement prediction, and rule-notifying system bring a concluded proposed system to use at ease for the users.

5. REFERENCES

[1] Garima Mathur and Dr. Mahesh Bundele, "Research on Intelligent Video Surveillance Techniques for Suspicious Activity Detection - Critical Review", International Conference on Recent Advances and Innovations in Engineering 2016Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] K. Dailey, "The rise of CCTV surveillance in the US," 29, April 2013.

[3] J. Park, J. Choi, M. Park, S. Hong, and H. Kim, "A Study on Intelligent Video Security Surveillance System with Active Tracking Technology in Multiple Objects Environment," International Journal of Security and Its Applications, vol. 6, pp. 211-216, 2012.

[4] M. Vidanapathirana, "Real-time Human Detection in Computer Vision — Part 1," 24 Mar 2018. [Online]. Available: https://medium.com/@madhawavidanapathirana/https-medium-com-madhawavidanapathirana-real-time-human-detection-in- computer-vision-part-1-2acb851f4e55. [Accessed 07 Jun 2020].

[5] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks," Microsoft Research.

[6] Phillips, P. Jonathon. "Support vector machines applied to face recognition." Advances in Neural Information Processing Systems.

[7] D. S. Vasant, B. A. Arun and M. P. Shivaji, "Raspberry-pi-based anti- theft security system with image feedback," Journal of Information, Knowledge, and Research in Electronics and Communication Engineering, vol. 04, no. 02, 2017.

[8] GV Balakrishna, B. Santhosh Kumar, "Smart intruder detection using video surveillance," International Journal of Science Technology and Management, vol no.5, Issue no.2, February 2016.

[9] Rajeev Thaware, "Real-Time Face Detection and Recognition with SVM and HOG Features," Monday, May28,2018.[Online].Available:https://www.eeweb.com /profile/rajeevthawa re/articles/real-time-face-detection-and-recognition-with-svm-and-hog- features.

[10] F. Emmert-Streib and M. Dehmer, "Evaluation of Regression Models: Model Assessment, Model Selection, and Generalization Error," Machine Learning and Knowledge Extraction, vol. 1, no. 1, pp. 521–551, Mar. 2019.

[11] A.N. Ansari, M.Sedky, N.Sharma and A.Tyagi, "An Internet of Things Approach for Motion Detection using Raspberry Pi, "2015.