

Performance Analysis of Facial Image Feature Extraction Algorithm for Smart Home Security System Detection

Muhammad Ihsan Adly¹, Satria Mandala^{2*}

^{1,2}*School of Computing, Telkom University
Bandung, Indonesia*

*satriamandala@telkomuniversity.ac.id

Abstract

In Indonesia, alongside advancements in multi-family security technology, smart home security tools have gained popularity. These tools operate automatically in real-time without environmental limitations. However, existing tools often lack consistent accuracy and performance. To address this, we propose a smart home security system incorporating an Arduino UNO-connected camera, two relay modules, a magnetic lock, and integration with a home Internet of Things system. Our research methodology involved: 1. Reviewing existing literature on Smart Home Security, specifically focusing on facial image feature extraction algorithms; 2. Deploying Arduino UNO, 2 Relay Module, and Solenoid Lock; 3. Employing the Wavelet feature extraction algorithm, chosen for its compatibility with human behavior patterns. Our proposed method aims for an accuracy rate of 80% or higher. Experimental results demonstrate that our prototype achieved an accuracy of 85.7%, along with a precision rate of 87.94%, recall rate of 87.56%, and an f1-score rate of 87.28%.

Keywords: Internet of Things, Smart Home Security, Arduino UNO, 2 Relay Module, Solenoid Lock, Wavelet

I. INTRODUCTION

WHIT the rapid development of economy, living standards and environmental degradation, people's demands for comfortable and safe living environment are getting higher and higher[1]. Criminal crimes such as robbery and theft are serious problems for any family[2]. The victims of theft are not only the middle to upper economic communities, as even the middle to lower economic communities are the victims of theft. The cause of many house break-ins is because the security level of the building is still low, making it easy for criminals to enter. Home security systems have become an important issue in recent years. In fact, the system is designed to ensure that your property and loved ones are always safe[3].

A home security system commonly used by Indonesians is only an ordinary door lock, where the lock can be easily broken with a screwdriver. There is also the problem of human error where people forget to lock the door at night, and so criminals can easily enter the house also, Traditional door locks are inconvenient because they require a key to access[4]. Consequently, it is crucial to implement preventive measures, such as developing a face recognition smart home security.

A smart home can be seen as an environment that utilizes communication and computing technologies to enhance the residents' quality of life and enable remote control and management of household appliances. The

primary goal is to improve convenience, safety, and comfort for the occupants of the building by integrating advanced technological solutions within the home[5]. With this advance technology, burglar are not easy to enter the house.

In recent times, face recognition has emerged as a prominent subject within pattern recognition. In response, researchers have put forward diverse approaches and methodologies to tackle this area[6]. Face recognition is a widely studied subject with applications spanning across various domains, including military, commercial, and public security fields[7]. There are many smart security devices for homes being sold nowadays, but these tools still have weaknesses, such as the low accuracy of tool detection[8]. And its application is still complex. By installing this detection tool, it is considered a first attempt to improve security and reduce cases of home break-ins.

Among the many smart home security tools that use facial recognition designed by researchers are[9] and [10], Most analysts center on the accuracy of detection in supporting conditions and disregard one of the concepts of this instrument, which is infinite. Usually, this device requires good lighting, and so if the lighting conditions are not stable, this instrument does not work accurately, and this can lead to break-ins. Besides the problem of detection algorithms, it is still rare to develop smart home security tools using facial recognition that can work with poor lighting. Security is one of the main concerns in the smart home environment[11].

This research proposes a feature extraction algorithm using Wavelet and a deep learning algorithm using CNNs for the classification. Wavelet feature extraction algorithm transforming facial images using wavelets to capture important spatial details and frequency information. It is expected that with the application of the proposed feature extraction algorithm, this tool can detect accurately under difficult conditions, and the obtained accuracy can be quite high and also can maintain the best security possible.

II. LITERATURE REVIEW

A. Existing Research

In this study[12], researchers designed an Internet of Things-based Face Recognition tool that uses physical location monitoring. The accuracy of this tool reaches 87.77% using the hold out method and reaches 89.81% using the stratified 10-fold cross validation method, for accuracy with other classifications is also still above 75% of all accuracy so that it can be said that this tool is competitive. In this study[8], researchers designed a smart home security system that can detect faces and movements using motion sensors and IoT cameras. The accuracy of this tool reaches 80%. This system uses KNN algorithm and Complex decision tree. In this study[13], researchers designed a system to detect facial expressions using Dual tree multi-band wavelet transform and Gaussian mixture model. The accuracy obtained from implementing the DTMBWT system and GMM classifier reached 98.14%. This system gets very high accuracy, so there are no shortcomings. In this study[2], researchers designed a face detection and voice detection system, which is able to perform detection when the user uses a face mask or medical mask. The advantage of this system is its high detection accuracy reaching 87.5% for face detection and 84.62% for voice detection. Even if the user uses a face mask or medical mask, the accuracy does not decrease much.

B. Theoretical Foundation

1) *Wavelet Feature Extraction*: Feature extraction is the process of extracting information from the original image. Efficient feature extraction is a very difficult problem for machine approaches because of human visual perception[13]. Feature extraction algorithms in facial recognition are used to identify relevant facial features, reduce dimensionality, normalize pose and lighting, handle noise and variation, and enable matching and classification. These algorithms extract characteristic facial features, create compact representations, account for variations in facial appearance, ensure robustness to noise, and accurately match faces for identification and authentication. One of the most potent and useful methods for face recognition is the wavelet transformation[14]. Wavelet feature extraction is a valuable face detection technique that can effectively analyze facial features in images. We apply a wavelet transform to the image and decompose it into its various

frequency components to capture global and local facial details. Multi-level decomposition allows extraction of finer features such as eyes, nose, and mouth. Reducing feature dimensionality improves efficiency. These extracted features are used to recognize faces using classification algorithms. This approach is more robust and can respond to changes in facial appearance, lighting conditions, and poses, making it useful for facial recognition systems, surveillance, and human-computer interaction. The extracted feature representations are used for subsequent representation classification[15].

2) *CNN Classification*: Classification is the final stage of the recognition process. Mainly used to categorize emotional states about a subject[16]. Perception is the human face. Despite the impressive success of traditional face recognition methods through extracted hand-crafted feature, researchers have turned to deep learning over the past decade for its high automatic recognition capabilities[17]. Deep learning is an artificial intelligence application with the ability to automatically learn from experience to improve performance[18]. A CNNs or Convolutional Neural Networks is a type of deep learning model used to analyze visual data such as images, which will automatically learn and extract important features from different levels of data. It captures progressively more complex details, enabling accurate detection and classification of facial image data. CNNs is successful in tasks such as image classification and face recognition due to its ability to learn and extract features from visual data. CNNs is used for classification tasks where it learns how to assign labels to input data such as facial images. This is achieved by automatically extracting relevant features from the input during the training process. CNNs adjusts weights to improve predictions, minimize errors, and achieve accurate classification. A feature of CNNs is its ability to capture and represent the key features needed to distinguish different objects and classes without the need for manual feature engineering. This property makes CNNs very effective in tasks where a feature extraction Wavelet plays an important role in face recognition.

3) *Test Parameters*: After constructing a model that consisted of a Wavelet feature extraction algorithm and CNNs (Convolutional Neural Network) Classification, the authors evaluated the performance of the model. The importance of performance evaluation of the model is to show that the model proposed by the authors is fulfilling the purpose of this experiment and can achieve the measurement goals of performance analysis in the model. The performance evaluation used in this performance analysis of facial image feature extraction algorithm experiment consisted of Accuracy, Precision, Recall, and F1-Score.

- Accuracy

Accuracy is a metric that measures the accuracy of the predictions made by the model. Quantify the percentage of correctly classified samples of the facial image data relative to the total number of instances in the facial image dataset. The formula for calculating accuracy is given below.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative}$$

- Precision

Precision is calculated by dividing the number of true positive predictions by the sum of true positive and false positive predictions. This provides insight into the model's ability to minimize false positives and accurately identify positive cases. The formula for calculating precision is given below.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

- Recall

Recall is a metric that measures the model's ability to correctly identify positive instances from the total number of actual positive instances. It quantifies the proportion of true positive predictions made by the model to the total number of positive instances in the dataset. The formula for calculating recall is given below.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

- F1-Score

F1-Score is a metric that combines precision and recall into a single measurement. F1-Score provides a balanced assessment of a model's performance by taking into account both false positives and false negatives. The formula for calculating F1-score is given below :

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)}$$

III. RESEARCH METHOD

A. Research Stages

This research begins with designing the system, then gathering the dataset, then create the proposed method using the proposed algorithm, and evaluate the outcome. The details of the process can be seen in Fig. 1 below.

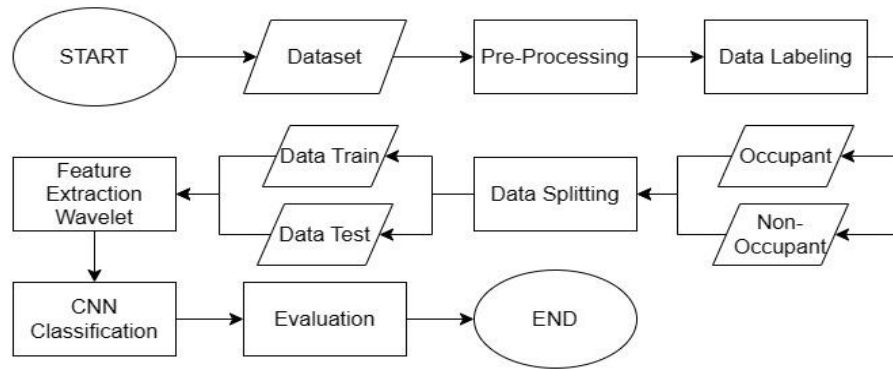


Fig. 1. Flowchart of the Processing Flow

Fig. 1 shows the process of performance analysis of feature extraction on facial image data. This experiment begins with a base design of the system, and then prepares the dataset. The dataset will undergo pre-processing development, where the pre-processing consists of labeling the facial image data ‘Occupant’ [Pnghuni] and ‘Non-Occupant’ [Bkn_Pnghuni]. After the facial image data has been labeled, the datasets are divided into two sets, which are Data Train and Data Test. After splitting, the feature of facial image data is extracted using the feature of an extraction algorithm Wavelet. The extracted dataset is then classified using a CNNs algorithm. The end result of the classification of the facial image data will be the same as the labeled data such as ‘Occupant’ and ‘Non-Occupant’.

B. Dataset for Experiments

The data required for this experiment consisted of facial images of family members and non-family members. The images were captured differently, with most of the facial image data taken from the internet website that provides datasets to be used for author experiments, and part of the facial image data was taken with author's smartphone Redmi Note 9 Pro. After gathering the facial image data, the author put a label on the data. For family members the label was ‘Pnghuni’ [Occupant] and for the non-family the data was labeled as ‘Bkn Pnghuni’ [Non-Occupant]. The labeled data from the internet and author's smartphone were split evenly for the Data Test and Data Train. The total amount of facial image data that was used in this experiment was four hundred and ten (407) images.

IV. RESULTS AND DISCUSSION

A. Dataset Input and Image Training using Proposed Model

The dataset that the author gathered consisted of privately collected data using the author's smartphone and data that the author collected from internet websites, specifically human facial data. The total combined data from the author's smartphone and internet website were 407 facial images. The data that the author privately collected using the author's phone were captured using various different angles while taking the facial image photo, and the data that the author collected from the internet were chosen variously to show different angles as well. The detailed facial image dataset that author used in this experiment is shown in TABLE I and TABLE II. After gathering and selection, the facial image data were trained using the proposed model that consists of a Wavelet feature extraction algorithm and CNNs (convolutional neural networks) classification algorithm. The model created will be used in this experiment.

TABLE I
 OCCUPANT FACIAL IMAGE DATASET SAMPLE.








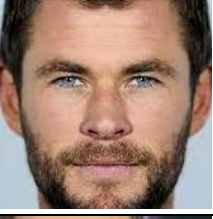



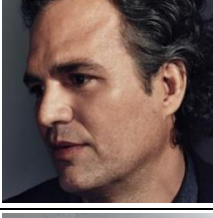



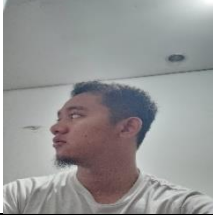
Amount	Source	Example of Facial Image Data
61	Writer Smartphone	
51	Internet Website	
54	Internet Website	
51	Writer Smartphone	

TABLE II
NON-OCCUPANT FACIAL IMAGE DATASET SAMPLE

Amount	Source	Example of Facial Image Data		
50	Internet Website			
53	Internet Website			
66	Internet Website			
21	Writer Smartphone			

B. Experiment

In this experiment, the model that was already created before using a feature extraction algorithm Wavelet and a classification algorithm CNNs will be the one that author uses to create a facial recognition program. The performance measurement of this model is evaluated using k-fold cross validation with the number of the fold is 5 (5-Fold). The author chose a K-Fold of 5 specifically because of the small size of the dataset which is less than 1000 data points in total. The performance measurement results of the model are shown in TABLE III below.

TABLE III
PERFORMANCE MEASUREMENT RESULTS.

Fold	Accuracy	Precision	Recall	F1-Score
1	84.1%	87.6%	86.6%	86.3%
2	86.4%	89.4%	89.3%	89.1%
3	82.5%	86.1%	85.5%	84.8%
4	84.8%	86.1%	87.7%	86.7%
5	87.1%	90.5%	88.7%	89.5%

With the result of all five folds, the average result of performance measurement of the model that author proposed was 85.7% in accuracy, 87.94% in precision, 87.56% in recall, and 87.28% in F1-Score. The highest performance measurement results were achieved in Fold 5, which are 87.1% in accuracy, 90.5% in precision, 88.7% in recall, and 89.5% in f1-score. The lowest performance measurement results were achieved in Fold 3, which are 82.5% in accuracy, 86.1% in precision, 85.5% in recall, and 84.8% in f1-score.

C. Hardware and Software

In this experiment, the author used the available hardware and software that was owned by the author personally. The details of the hardware and software specifications are shown in the next point.

1) Software Specification

- Windows 10 Pro
- Python
- OpenCV
- Jupyter Notebook
- Visual Studio Code
- Arduino IDE

2) Hardware Specification

- Processor : Intel Core i5-7200U
- Memory : 8192MB RAM
- Arduino UNO
- 2 Relay Module
- Solenoid Lock
- 12v Power Supply
- Wires male/female

D. Prototype Testing

In this prototype testing, the author used the already trained model and implemented the trained model into the prototype device that the author proposed. There are two conditions for the end results. The details of the results will be shown in the next point below.

1) *Accepted Condition*: The accepted condition is where the facial image data that has been captured was matched according to the description of the trained model. This condition will allow the solenoid lock to open because it detects that the person is classified as 'Penghuni' [Occupant]. The example of the accepted condition is shown in Fig. 2 below.

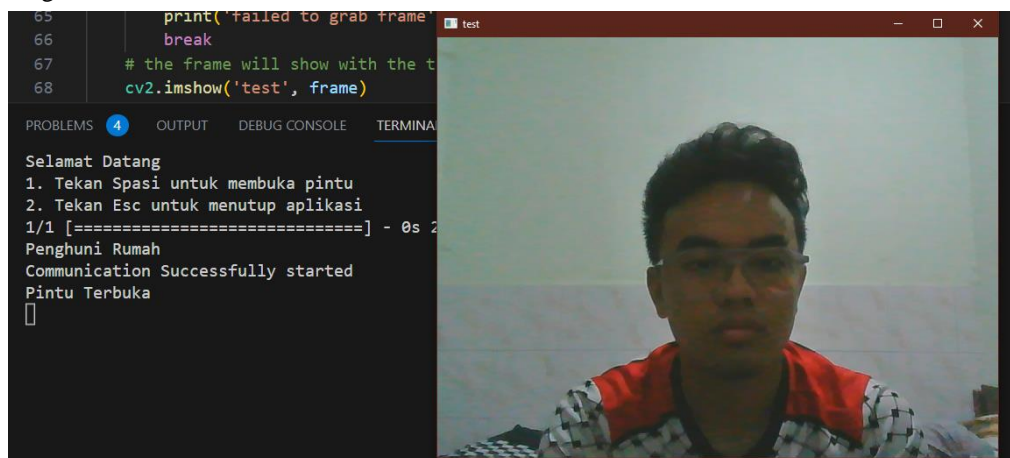


Fig. 2. Accepted Condition Example

2) *Declined Condition*: The declined condition is where the facial image data that has been captured did not match according to the description of the trained model. This condition will not allow the solenoid lock to open, because it detects that the person is classified as 'Bkn_Penghuni' [Non-Occupant]. The example of the declined condition is shown in Fig. 3 below.

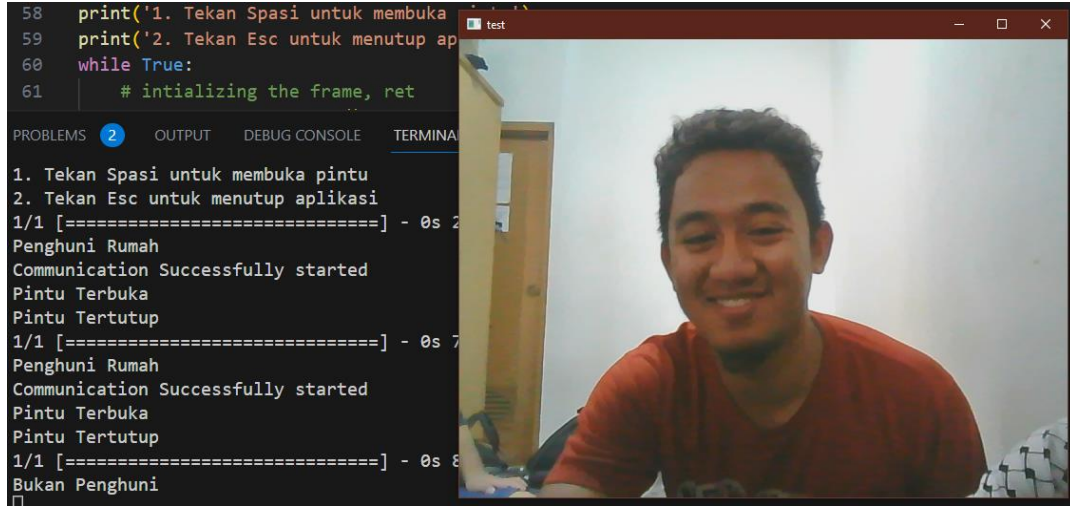


Fig. 3. Declined Condition Example.

E. Discussion

Recent research experiments into facial recognition have already demonstrated a significant increase in terms of accuracy through the application of a feature extraction Wavelet algorithm. This experiment also highlighted the effectiveness of a feature extraction algorithm Wavelet in facial image detection. With the use of a feature extraction algorithm Wavelet, the result of accuracy has been improved and the result is stable. The advantages of the Wavelet algorithm compared to other feature extraction algorithm, include its ability to analyze data at different scales, accurately localize features in time and frequency, detect edges, compress data effectively, adapt to different data parts, and succeed in tasks like object recognition and biomedical signal analysis. Nevertheless, the choice of the appropriate method depends on the unique characteristics of the data and the specific requirements of the task at hand. TABLE iv below is a comparison of accuracy from other experiments using different methods also using the feature extraction algorithm Wavelet.

TABLE IV
PERFORMANCE COMPARISON WITH OTHER RESEARCH ON FACE DETECTION.

Method	Accuracy
Wavelet with Convolutional Neural Networks	85.7%
HOG with Feedforward Neural Networks[19]	92.34%
Gabor-Based[20]	76.67%
Quanvolutional Neural Networks with a Single Quanvolutional Filter[21]	95.8%

The objective of this comparison in TABLE iv is to show the advantage of a Wavelet feature extraction algorithm from other methods. The accuracy performance outcome of using Wavelet feature extraction algorithm that writer achieved was 85.7%, another experiment from[19] that used HOG (Histogram of Oriented Gradients) achieved 92.34% accuracy, where the experiment from[20] achieved 76.67% using the Gabor-Based method. The experiment from[21] using Quanvolutional Neural Networks achieved 95.8% QNNs are simply

an extension of classical CNNs. The various classification methods mentioned earlier are part of the artificial neural network category, which relies on shared mathematical modeling principles in the field of machine learning. These methods incorporate neuron activation functions, such as ReLU, sigmoid, or tanh, within their layers to introduce non-linearity into their models. Despite their unique architectural variances, they have practical applications across a wide range of machine learning tasks. The difference between results are the dataset that been used in every research is not the same, therefore there is a quite significance differences.

V. CONCLUSION

In this section, the writer will present the conclusion and analyze the result of the experiment in order to show if the proposed research objective has been achieved. After careful analysis, the experiment showed a successful outcome of the proposed accuracy percentage. The proposed method of facial image recognition using a Wavelet feature extraction algorithm and CNNs (Convolutional Neural Networks) classification algorithm achieved a percentage 85.7% in accuracy, 87.94% in precision, 87.56% in recall, and 87.28% in F1-score. The detection of faces using a combination of Wavelet feature extraction and CNNs (Convolutional Neural Networks) classification provides significant benefits over a Gabor-Based method. Both wavelet-based feature extraction algorithms and convolutional neural networks (CNNs) offer distinct advantages in various aspects of facial recognition. Wavelet-based methods are well suited for tasks requiring multiresolution analysis, feature localization, edge detection, texture analysis, and energy compression. These are especially useful for tasks with varying scales and noise reduction requirements. Wavelets can efficiently retrieve important information from images and are suitable for various feature fusion techniques. CNNs, on the other hand, have revolutionized the field of computer vision with their ability to automatically learn hierarchical and abstract features from raw pixel data. They excel at tasks such as object recognition, image classification, and semantic segmentation. CNNs can effectively handle complex and large datasets, offering high accuracy and generalization ability. They can learn hierarchical representations, making it suitable for tasks where the underlying patterns are not explicitly known. It is important to consider the specific requirements of the task at hand when deciding on the appropriate approach. Wavelet-based techniques may be suitable for tasks involving multiresolution analysis, localization, and denoising, whereas CNNs are better suited for tasks requiring advanced pattern recognition and large-scale image analysis. In some cases, a combination of both methods can take advantage of each other's strengths and provide even better performance. Combining wavelet-based functions with convolutional neural networks (CNNs) offers several advantages over Gabor-based functions. CNNs can automatically learn hierarchical and abstract features from raw data, making them more adaptive and capable of capturing complex patterns. This data-driven approach improves performance across a wide variety of tasks and datasets. Additionally, end-to-end training with CNNs optimizes both feature extraction and classification, increasing model effectiveness. This combination also makes it more resistant to changes in pose, lighting, and scale. Additionally, wavelet-based functions help reduce dimensionality and make processing of large image datasets more efficient. However, when choosing between Gabor functions and wavelet-based functions, you should consider your specific task requirements, as either method may work well for your particular application. Experimentation and evaluation are essential to determine the most appropriate feature extraction approach for a particular problem. However the effectiveness of this research method's accuracy percentage is also determined by the dataset used, the complicatedness of the condition, and the implementation of the method on different problems. In real-world try-out, the result may produce a different outcome of detection rate.

DATA AND COMPUTER PROGRAM AVAILABILITY

Data and program used in this paper can be accessed in the following site <https://github.com/GNite12/Tugas-Akhir.git>, and <https://s.id/1RZCN>

ACKNOWLEDGMENT

The authors would like to thank everyone that has support the author for this entire research, especially author mother and father, also author lecturer for the guidance during the research, and to all of author friends that support author along the way.

REFERENCES

- [1] A. Yang, C. Zhang, Y. Chen, Y. Zhuansun, and H. Liu, "Security and Privacy of Smart Home Systems Based on the Internet of Things and Stereo Matching Algorithms," *IEEE Internet Things J*, vol. 7, no. 4, pp. 2521–2530, Apr. 2020, doi: 10.1109/JIOT.2019.2946214.
- [2] N. Saxena and D. Varshney, "Smart Home Security Solutions using Facial Authentication and Speaker Recognition through Artificial Neural Networks," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 154–164, Jun. 2021, doi: 10.1016/j.ijcce.2021.10.001.
- [3] A. R. Syafeeza, M. K. Mohd Fitri Alif, Y. Nursyifaa Athirah, A. S. Jaafar, A. H. Norihan, and M. S. Saleha, "IoT based facial recognition door access control home security system using raspberry pi," *International Journal of Power Electronics and Drive Systems*, vol. 11, no. 1, pp. 417–424, Mar. 2020, doi: 10.11591/ijped.v11.i1.pp417-424.
- [4] S. Phawinee, J. F. Cai, Z. Y. Guo, H. Z. Zheng, and G. C. Chen, "Face recognition in an intelligent door lock with ResNet model based on deep learning," in *Journal of Intelligent and Fuzzy Systems*, IOS Press BV, 2021, pp. 8021–8031. doi: 10.3233/JIFS-189624.
- [5] P. K. Sharma, J. H. Park, Y. S. Jeong, and J. H. Park, "SHSec: SDN based Secure Smart Home Network Architecture for Internet of Things," *Mobile Networks and Applications*, vol. 24, no. 3, pp. 913–924, Jun. 2019, doi: 10.1007/s11036-018-1147-3.
- [6] X. Ning, W. Li, B. Tang, and H. He, "BULDP: Biomimetic Uncorrelated Locality Discriminant Projection for Feature Extraction in Face Recognition," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2575–2586, May 2018, doi: 10.1109/TIP.2018.2806229.
- [7] G. Y. Chen, T. D. Bui, and A. Krzyzak, "Illumination invariant face recognition using dual-tree complex wavelet transform in logarithm domain," *Journal of Electrical Engineering*, vol. 70, no. 2, pp. 113–121, Apr. 2019, doi: 10.2478/jee-2019-0017.
- [8] O. Taiwo and A. E. Ezugwu, "Internet of Things-Based Intelligent Smart Home Control System," *Security and Communication Networks*, vol. 2021, 2021, doi: 10.1155/2021/9928254.
- [9] J. Pierce, "Smart home security cameras and shifting lines of creepiness a design-led inquiry," in *Conference on Human Factors in Computing Systems - Proceedings*, Association for Computing Machinery, May 2019. doi: 10.1145/3290605.3300275.
- [10] M. Sajjad *et al.*, "Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Generation Computer Systems*, vol. 108, pp. 995–1007, Jul. 2020, doi: 10.1016/j.future.2017.11.013.
- [11] S. Nagarkar and V. Prasad, "Evaluating Privacy and Security Threats in IoT-based Smart Home Environment," Maharashtra, 2019. [Online]. Available: <http://www.ripublication.com>
- [12] A. S. Shitole and M. H. Devare, "TPR, PPV and ROC based performance measurement and optimization of human face recognition of IoT enabled physical location monitoring," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 3582–3590, Jul. 2019, doi: 10.35940/ijrte.B3186.078219.
- [13] J. Kommineni, S. Mandala, M. S. Sunar, and P. M. Chakravarthy, "Advances in computer–human interaction for detecting facial expression using dual tree multi band wavelet transform and Gaussian mixture model," *Neural Comput Appl*, vol. 34, no. 18, pp. 15397–15408, Sep. 2022, doi: 10.1007/s00521-020-05037-9.

- [14] T. V, Vellore Institute of Technology, Institute of Electrical and Electronics Engineers. Madras Section, IEEE Communications Society., and Institute of Electrical and Electronics Engineers, *Conference proceedings, International Conference on Vision Towards Emerging Trends In Communication and Networking (ViTECoN 2019) : 30-31, March 2019, Vellore, Tamilnadu, India.* 2019.
- [15] J. Kommineni, S. Mandala, M. S. Sunar, and P. M. Chakravarthy, "Accurate computing of facial expression recognition using a hybrid feature extraction technique," *Journal of Supercomputing*, vol. 77, no. 5, pp. 5019–5044, May 2021, doi: 10.1007/s11227-020-03468-8.
- [16] Z. Balogh, M. Magdin, and G. Molnár, "Motion Detection and Face Recognition using Raspberry Pi, as a Part of, the Internet of Things," 2019.
- [17] W. Mellouk and W. Handouzi, "Facial emotion recognition using deep learning: Review and insights," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 689–694. doi: 10.1016/j.procs.2020.07.101.
- [18] G. Katare, G. Padihar, and Z. Qureshi, "Challenges in the Integration of Artificial Intelligence and Internet of Things 11," Indore, Dec. 2018.
- [19] M. Awais *et al.*, "Real-time surveillance through face recognition using HOG and feedforward neural networks," *IEEE Access*, vol. 7, pp. 121236–121244, 2019, doi: 10.1109/ACCESS.2019.2937810.
- [20] L. Li, H. Ge, Y. Tong, and Y. Zhang, "Face Recognition Using Gabor-Based Feature Extraction and Feature Space Transformation Fusion Method for Single Image per Person Problem," *Neural Process Lett*, vol. 47, no. 3, pp. 1197–1217, Jun. 2018, doi: 10.1007/s11063-017-9693-4.
- [21] M. Henderson, S. Shakya, S. Pradhan, and T. Cook, "Quantum convolutional neural networks: powering image recognition with quantum circuits," *Quantum Mach Intell*, vol. 2, no. 1, Jun. 2020, doi: 10.1007/s42484-020-00012-y.