International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 325-331, May-2024

https://doi.org/10.47001/IRJIET/2024.805043

Advancing Security Measures in Governmental Institutions: Integration of Facial Recognition and Movement Monitoring Technologies in the House of Representatives

Hala Wael AlFadhel

Department of Computer and Communications Engineering, Faculty of Engineering & Computer Science,
American University of Science & Technology, Beirut, Lebanon

***Abstract* - With the aim of improving security measures in public and government buildings, this study presents a system to protect the House of Representatives using Vicon sensors and facial recognition technology to detect threats as soon as they occur and alert accordingly to take appropriate measures. A face recognition model was realized, which took advantage of a DenseNet169-based feature extractor and a dense layer-based classifier. The machine learning models and Vicon sensors used by the physical motion capture system allowed for highly accurate analysis of real-world movements. Using Decision Tree (DT) and K-Nearest Neighbors (KNN) algorithms, the system achieved optimal accuracy using a dataset that included ten categories of behaviors performed by ten employees. Thanks to the combination of motion capture and facial recognition technology, allowing for precise threat classification, the House of Representatives will have a robust security system. This research highlights the importance of technical improvements in defending public employees and facilities.**

Keywords: Vicon Sensors, Physical Security, Machine Learning, Government Buildings, Motion Capture, Face Recognition.

I. INTRODUCTION

In today's security world, especially in sensitive government settings, there is a growing demand for sophisticated surveillance systems that can proactively detect and address possible dangers. Conventional security procedures, however essential, frequently fail to proactively identify less visible risks such as subtle deviations in behavior or symptoms of an impending attack [1]. The security of public and government facilities has been a pressing concern in recent years, since there has been a rise in threats and incidents targeting the disruption of governmental services. The House of Representatives, as a crucial institution for democratic administration, is particularly susceptible to both internal and external threats that can take on different forms, ranging from physical assaults to disruptions caused by individuals displaying hostile conduct. Existing security procedures frequently fall short in proactively detecting and addressing these threats before they evolve into real incidents [2].

Existing security systems lack the ability to effectively integrate real-time data processing with precise threat detection, a critical requirement for proactively preventing incidents. Hence, there is a requirement for a robust security system that accomplishes prompt identification and swift reaction to potential security risks by leveraging fortified and strategically positioned protection equipment and sensors within the premises, such as surveillance cameras. The aim of this study is to create and apply a state-of-the-art physical security system that integrates advanced Vicon sensors and facial recognition technology. The purpose of this system is to improve the effectiveness of security measures by using a proactive and adaptable strategy to identifying and handling threats to guarantee the safety of House of Representatives.

II. RELATED WORK

Some studies have developed security systems to protect the system from actual attacks, relying on cameras, surveillance devices, and advanced Internet of Things devices. The physical environment usually offers diverse information that comes from many sources. Physical Access Control Systems (PACSs) govern the physical admission to different zones, both from outside and within the monitored area. The cornerstone of Personal Identity Verification (PIV) for persons identification is formed by three factors: knowledge-based authentication (such as a password), possession-based authentication (such as a smart card), and biometric-based authentication (such as a fingerprint or other biometric information) [3]. Video surveillance systems are increasingly relying on communication systems that utilize wired or wireless technology to automatically evaluate recorded movies using artificial vision algorithms [4]. Numerous environmental sensor systems rely on information and communication technology solutions to convey data, and they integrate a wide variety of possible sources of information. Some examples of sensors include those that gauge voltage for battery or uninterruptible power supply applications, those that assess

ISSN (online): 2581-3048
Volume 8, Issue 5, pp 325-331, May-2024
https://doi.org/10.47001/IRJIET/2024.805043

humidity levels to prevent rapid deterioration of equipment, those that measure temperature to detect air conditioning failures that could be catastrophic for devices such as servers, those that detect fluid leaks, those that measure airflow to ensure sufficient cooling in specific areas, those that detect motion to grant access to restricted areas, and those that detect audio signals, such as alarms or the sound of shattered glass.

The security system of a property serves as its primary means of protection, as it promptly alerts owners to the presence of any unauthorized individuals. Contemporary security systems possess the capability to detect motion through a diverse range of sensors and promptly notify the owner in the event of an intrusion. Many of these systems lack features like as zone barriers, facial recognition, remote video monitoring, and power failure detection. Additional significant attributes are user-friendliness, financial viability, and energy conservation. The main goal of the proposed architecture in [5] is to tackle these difficulties by developing an intelligent security and monitoring system. This system would utilize a range of ultrasonic sensors to detect any instances of property trespassing and notify the owner about the presence of an unauthorized individual. Regardless of whether someone intentionally broke in or not, the system will promptly notify them, enabling them to revert to an earlier time point without triggering any alerts. The owner can remotely watch his property using the remote video surveillance feature. In the event of a power outage, this system also sends a notification to the owner. This system utilizes face recognition technology as an authentication method to ensure that only authorized individuals are permitted access to their property.

Advanced gadgets make it easier to convey identified emotions to the user by processing raw visual data directly from each camera using efficient and tailored algorithms. In order to address this need, a lightweight distributed system was developed in [6] that relies on edge computing and Raspberry Pi. The system optimizes data transfer and component deployment, lowers round-trip delay, and completes complex computing tasks, as well as provides large-scale communication services with high reliability. A secure and environmentally friendly smart building is presented in [7] by use of modern Internet of Things (IoT) technology. The Constrained Application Protocol is one of the most important web transport protocols, and every device has its own unique address. It is a protocol at the application layer that avoids sending data over encrypted channels. Datagram Transport Layer Security (DTLS) safeguards data with automatic key management, confidentiality, authentication, and data integrity. Enhancements from the Certificate Authority (CA) are used to combine the DTLS protocol with the Secure Hash Algorithm (SHA-256), which further improves security.

Privacy and security are two of the most important areas where technology helps us. Since smartphones are the most popular smart device, they can also serve as a security alarm system. Also, there has been a recent uptick in the usage of smart IoT devices that incorporate AI. A system for smart home security was created in [8]. The Raspberry Pi's No Infrared (NoIR) camera module records movies and takes images, making it a security system in a nutshell. Additionally, a PIR motion sensor is employed for this purpose. With the help of the built algorithm and data gathered from motion sensors and the NoIR Pi camera module, it is suggested to use facial recognition classification technology to forecast potential security threats. An intelligent home automation system is introduced in [9] with the purpose of managing household appliances, monitoring environmental conditions, and detecting activity within the home and its environs. The suggested approach involves the use of a sophisticated deep learning model that can accurately recognize and classify motion based on detected patterns. This model is specifically designed to detect intruders and effectively prevent false alarms. An individual captured by a surveillance camera is categorized as either an invader or a resident of the residence, depending on their distinctive walking style. Convolutional Neural Networks (CNN) model was utilized to conduct an experimental examination of human mobility patterns for evaluation purposes.

Various technologies exist to differentiate between different suspicious actions by actively monitoring live video data. Human behavior is often unpredictable, making it challenging to determine whether it is suspicious or typical. A deep learning technique is employed in [10] to identify abnormal or typical behavior in the academic setting. If a suspect action is anticipated, an alarm message is sent to the appropriate authorities. Surveillance is commonly conducted by extracting successive frames from video footage. The entire frame is bisected into two sections. The initial phase involves the computation of features from the video frames, while the subsequent phase utilizes the generated feature classifier to forecast if the class is suspicious or normal. The growing cloud computing paradigm offers a chance to manage the immense volume of monitoring data that is continuously generated on-site across IoT systems. Nevertheless, the detection performance remains unsatisfactory due to the intricate monitoring environment. The emphasis in [11] is on multi-target detection for the purpose of real-time monitoring in smart IoT systems. A novel deep neural network model named A-YONet is utilized in a cutting-edge cloud monitoring system. This model is created by merging the strengths of YOLO and MTCNN. Its purpose is to enable efficient training and feature learning with limited computational resources. A sophisticated detection algorithm is subsequently created using the installation box preset system and a technique for

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 325-331, May-2024

https://doi.org/10.47001/IRJIET/2024.805043

integrating features at many levels. The method's efficiency in improving detection accuracy, particularly for multi-target detection, is demonstrated through experiments and evaluations utilizing two data sets. These data sets include a public data set and a custom data set obtained from a genuine surveillance system.

III. METHODOLOGY

Government buildings must be extremely secure as the number of threats and security breaches has skyrocketed in the past several years. Because of its central role in the legislative process, the House of Representatives need stringent security measures to protect its members and employees. To meet this need, an advanced physical security system will be introduced, utilizing the latest technology in the form of Vicon sensors and facial recognition cameras.

3.1 Components of the proposed security system

The system was designed with two basic components, the first to see and analyze facial expressions and the second to analyze the physical movements that occur inside the House of Representatives building. The first system is a facial recognition system that uses surveillance cameras to record complex facial characteristics of individuals. The placement of these cameras throughout the facility has been carefully planned to ensure complete coverage. The main purpose of this system is to use advanced algorithms to scan facial expressions and identify indicators of violent behavior, which may indicate potential security risks.

The second part of the security system incorporates Vicon sensors, which are widely recognized for their exceptional accuracy in capturing and analyzing bodily movements. The sensors are strategically placed in important locations across the facility to monitor the movements of individuals, enhancing the security measures alongside the facial recognition technology. The system can detect abnormal or unpredictable behaviors by tracking individuals' movements inside the area, which may require additional examination.

Both systems employ sophisticated machine learning algorithms to analyze the gathered data. Machine learning algorithms are trained on extensive datasets of facial expressions and movement patterns to accurately differentiate between normal and hostile actions. This allows the security system to offer immediate evaluations and notifications, enabling prompt reactions to any security risks. The incorporation of facial recognition and movement monitoring technology into the security system of the House of Representatives signifies a substantial advancement in safeguarding public officials and upholding the authenticity of governmental activities. Continual improvement and fine-tuning of these systems are essential for adjusting to changing security issues, guaranteeing that safety measures are both proactive and reactive.

3.2 Facial recognition system

The development of facial recognition technology has made substantial strides in recent years, particularly due to improvements in deep learning and neural network architectures. These technological breakthroughs enable the immediate analysis of facial expressions to determine emotional states that may indicate possible dangers. The suggested method employs strategically positioned surveillance cameras within the building to capture facial characteristics and transmit the images to the face recognition system. The objective is to examine the facial expressions and ascertain whether an individual is exhibiting aggressive conduct or not.

The FER2013 dataset contains images and categories describing the intensity of emotions felt by the person appearing in each image. The dataset contains grayscale images with dimensions of 48 x 48 pixels. The pictures depict seven separate emotions, two categories neutral and angry, which are relevant to the concept of safety. Figure 1 shows examples from the dataset.

Figure 1: Samples from FER2013 dataset

Image data generators are designed for training, validation and testing purpose. These generators automatically import batches of images and quickly modify their dimensions and color mode (RGB). Table 1 indicates the number of images in each subset.

Table 1: Number of samples in each subset

	Number of images
Training Set	7168
Validation Set	1792
Testing Set	2191

Figure 2 shows the distribution of classes in the training set, where there are 3196 images of an angry face and 3972 images of a neutral face.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 325-331, May-2024

https://doi.org/10.47001/IRJIET/2024.805043

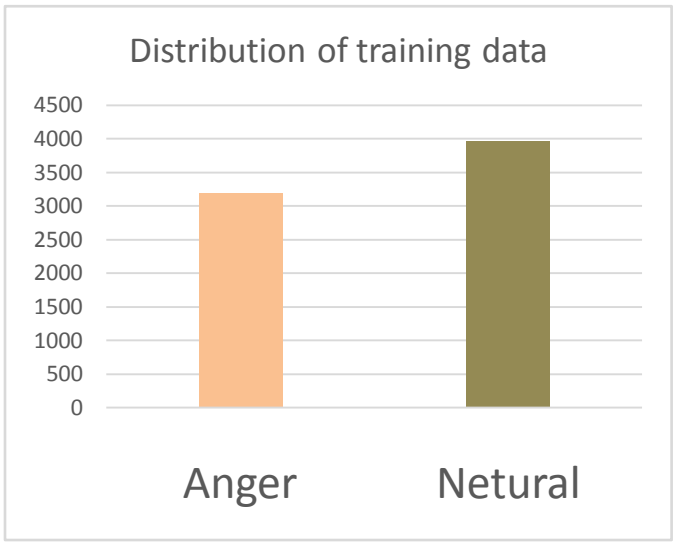

Figure 2: Distribution of the training data

The approach employs a transfer learning framework, in which a pre-trained model on a vast dataset is utilized to transfer the acquired knowledge to a particular task using a smaller dataset. The classification challenge involved constructing a neural network model, as depicted in Figure 3, with dense layers that incorporate dropout regularization. The model utilized DenseNet169 as a pre-trained feature extractor and excluded the completely connected layers at the top.

Early Stopping callback was employed during model training to address overfitting. It halts the training process when the model's performance on the validation set ceases to improve. The model underwent multiple iterations of refinement by utilizing training and validation data generators to exploit existing knowledge from the feature extraction layers.

```
Model: "model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, 50, 50, 3)]       0

 densenet169 (Functional)    (None, 1, 1, 1664)        12642880

 global_average_pooling2d (  (None, 1664)              0
 GlobalAveragePooling2D)

 dense (Dense)               (None, 256)               426240

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 1024)              263168

 dropout_1 (Dropout)         (None, 1024)              0

 dense_2 (Dense)             (None, 512)               524800

 dropout_2 (Dropout)         (None, 512)               0

 classification (Dense)      (None, 2)                 1026

=================================================================
Total params: 13858114 (52.86 MB)
Trainable params: 1215234 (4.64 MB)
Non-trainable params: 12642880 (48.23 MB)
_________________________________________________________________
```

Figure 3: Deep neural network structure for face recognition system

Figure 4 illustrates the correlation between the number of epochs and the increase in accuracy, while Figure 5 depicts the relationship between the number of epochs and the decrease in loss for both the training and validation sets.

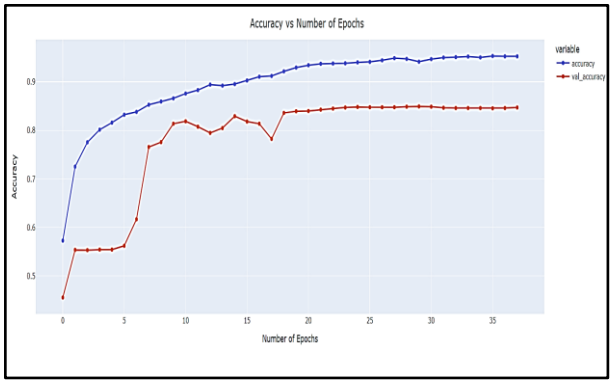

Figure 4: Correlation between the number of epochs and accuracy

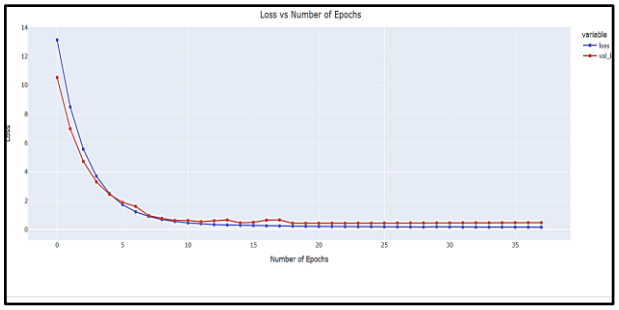

Figure 5: Correlation between the number of epochs and loss

Evaluation metrics were obtained by making predictions on the test set. The rating report displayed in Figure 6 provides a comprehensive summary of the model's performance in each category, including overall metrics. It shows that the model performs well on the test set, with balanced precision and recall for both classes.

	precision	recall	f1-score	support
0	0.83	0.81	0.82	958
1	0.85	0.87	0.86	1233
accuracy			0.84	2191
macro avg	0.84	0.84	0.84	2191
weighted avg	0.84	0.84	0.84	2191

Figure 6: Classification report of facial recognition model

Predicting a neutral expression is easier for the model than an angry one, according to the improved accuracy and F1 score for the "neutral" category. These all measures demonstrate that the model is successful in making accurate forecasts.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 325-331, May-2024

https://doi.org/10.47001/IRJIET/2024.805043

3.3 Body movement analysis system

Vicon sensors, used in this approach, efficiently capture exact physical motions across the House of Representatives building from different angles and spots. Machine learning is then used to examine the data collected by the sensors, providing a first assessment of how hostile or normal the person's behavior is.

Vicon dataset consists of 10 physical movements and 10 aggressive physical movements, representing human activity. A data set of 10 individuals (7 males and 3 females) was acquired using a Vicon capture system. Analytical tests were conducted and took only 10 seconds per participant, with a sampling frequency of 200 Hz. Data were classified into 20 distinct categories, as well as a binary classification of aggressive or normal behavior. All activities were used, shown in Figure 7, which shows the distribution of samples in each category.

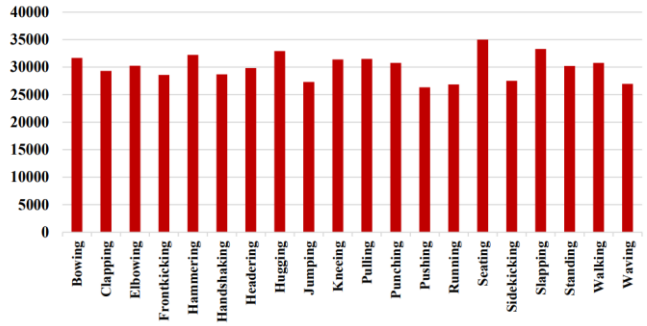

Figure 7: Distribution of the activities on Vicon physical action dataset

The data was processed using Label Encoder to translate categorical labels into numeric format and MinMaxScaler to rescale numeric features within a specified range, typically between 0 and 1. The dataset was partitioned into two subsets: a training set and a testing set. The training set was assigned 80%, while the testing set was assigned 20%. Multiple machine learning classifiers, including Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), Multi-Layer Perceptron (MLP), and Logistic Regression (LR), were employed, trained, and assessed to determine the optimal one. The duration of training for each classifier was measured using a timer.

Figure 8 shows the accuracy results for the classifiers. Both the DT and KNN algorithms obtain an excellent accuracy score of close to 100%. NB shows the lowest level of accuracy.

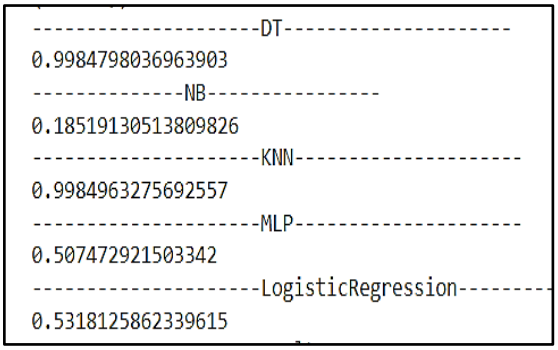

Figure 8: Test accuracy of capture physical movements models

Figure 9 shows the precision results of the classifiers where both DT and KNN maintain their flawless scores of 1.0.

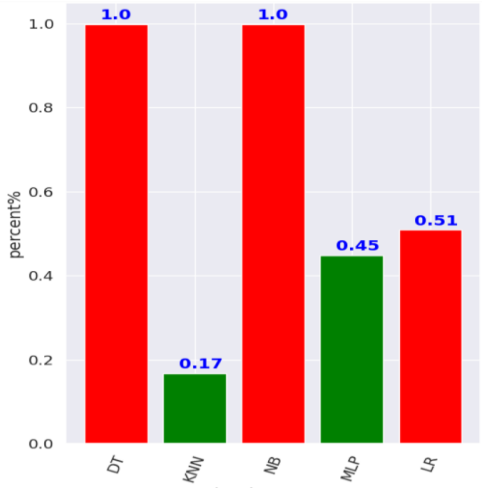

Figure 9: Precision results

Figure 10 shows the recall results for the classifiers. Once again, the DT and KNN algorithms achieved perfect recall rankings.

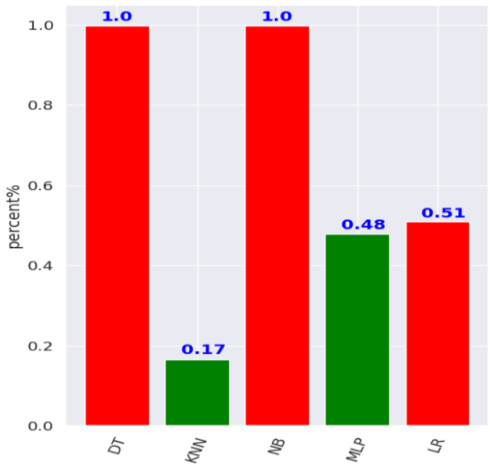

Figure 10: Recall results

Figure 11 shows the recall results for the classifiers. Once again, the DT and KNN algorithms achieved perfect recall rankings.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 325-331, May-2024

https://doi.org/10.47001/IRJIET/2024.805043

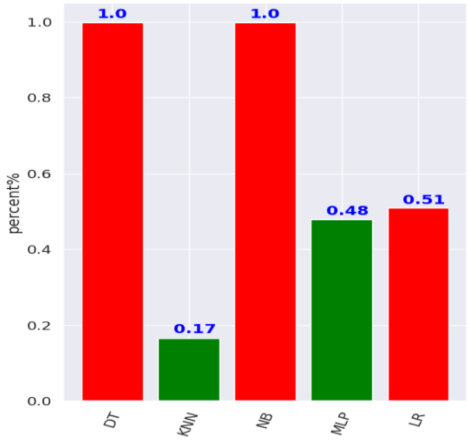

Figure 11: F1-score results

These results indicate that the DT and KNN models have exceptional performance, delivering perfect results in all criteria. Conversely, the NB model exhibits a notable decline in performance in comparison to the other models. This suggests that the underlying assumptions of the Naive Bayes method may not be applicable to this particular data set, or may necessitate a different way of representing the data.

IV. RESULTS DISCUSSION

The face recognition model utilizes a feature extractor based on DenseNet169, along with a classifier built on dense layers. This model has robust performance, attaining an overall accuracy rate of 84%. The algorithm may have an easier time identifying neutral expressions because they are likely to have less volatility compared to anger expressions. The Vicon sensors and motion models utilized in the physical motion capture system offer a very precise method for capturing and analyzing physical movements. Utilizing a comprehensive dataset that encompasses 10 different types of physical and aggressive actions offers a robust foundation for training these models. Both the DT and KNN algorithms achieve optimal performance across all criteria.

Table 2 shows that our models outperformed related studies using the same datasets in various. In [12], a real-time CNN approach with the OpenCV library predicts face emotion classification and detection by feature extraction. The Raspberry Pi study technique includes face identification, feature extraction, and emotion categorization. CNN was used to recognize facial emotions using the Facial Emotion Recognition (FER-2013) dataset. Forecast accuracy was 65.97%. In [13], a human activity identification model was trained and verified using the Vicon physical actions dataset and the convolutional neural network.

Table 2: Comparing our models with related work

	Facial recognition model	**Capture physical movements models**
Our Methods	DenseNet169	DT, NB, KNN, MLP, LR
Our Results	Accuracy of 84%	Accuracy of 100%
Comparative study	2020 – Zahara et al.[12]	2018 – Olatunji et al.[13]
Results of the comparison study	Accuracy of 65.97%.	Accuracy of 95%
Methods of the comparison study	CNN	CNN

V. CONCLUSION

The deployment of a sophisticated physical security system in the House of Representatives has demonstrated significant promise in bolstering safety protocols for this vital infrastructure. This study introduced a two-component system consisting of facial recognition technology and Vicon sensors, each specifically designed to target different aspects of security threat detection. The facial recognition component utilized an advanced DenseNet169 model, which accurately differentiates between neutral and aggressive facial expressions with a high level of precision. This skill is crucial in detecting possible threats before they escalate and equipping security personnel with the required knowledge to intervene proactively. The Vicon sensor-based motion capture system accurately assessed bodily movements, categorizing them as usual or aggressive behaviors with a 100% accuracy rate. The Vicon sensors, known for their accuracy and dependability, were used in conjunction with the powerful Decision Tree and K-Nearest Neighbors algorithms to create a very effective technique for monitoring and evaluating the physical movements of individuals on the premises. These systems provide complete security using facial and behavioral analytics. These technologies' integration into the House of Representatives' security architecture improves legislative process and staff protection. The findings stimulate security technology development and enhancement to stay up with increasing threats and defend public institutions. The field of research can be expanded in the future by developing a motion recognition system to anticipate the next movement that a person will make, providing proactive attack detection systems, and investigating advanced deep learning methods to

ISSN (online): 2581-3048
Volume 8, Issue 5, pp 325-331, May-2024
https://doi.org/10.47001/IRJIET/2024.805043

deal with differences in lighting and improve the accuracy of the face recognition system.

REFERENCES

[1] E. James, and F. Rabbi, "Fortifying the IoT Landscape: Strategies to Counter Security Risks in Connected Systems," *Tensorgate Journal of Sustainable Technology and Infrastructure for Developing Countries*, vol. 6, no. 1, pp. 32-46, 2023.

[2] S. Krishnan, M. S. Anjana, and S. N. Rao, "Security considerations for IoT in smart buildings," *In 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1-4, 2017.

[3] H. Ferraiolo, K. L. Mehta, N. Ghadiali, J. Mohler, V. Johnson, and S. Brady, "Guidelines for the use of PIV credentials in facility access," *NIST Spec. Publ*, *800*, 2018.

[4] G. S. D. M. A. Sreenu, and S. Durai, "Intelligent video surveillance: a review through deep learning techniques for crowd analysis," *Journal of Big Data*, vol. 6, no. 1, pp. 1-27, 2019.

[5] G. Lulla, A. Kumar, G. Pole, and G. Deshmukh, "IoT based smart security and surveillance system," *In 2021 international conference on emerging smart computing and informatics (ESCI)*, pp. 385-390, 2021.

[6] J. Yang, T. Qian, F. Zhang, and S. U. Khan, "Real-time facial expression recognition based on edge computing," *IEEE Access*, vol. 9, pp. 76178-76190, 2021.

[7] A. Kumar, S. Sharma, N. Goyal, A. Singh, X. Cheng, and P. Singh, "Secure and energy-efficient smart building architecture with emerging technology IoT," *Computer Communications*, vol. 176, pp. 207-217, 2021.

[8] A. J. Majumder, and J. A. Izaguirre, "A smart IoT security system for smart-home using motion detection and facial recognition," *In 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*,pp. 1065-1071, 2020.

[9] O. Taiwo, A. E. Ezugwu, O. N. Oyelade, and M. S. Almutairi, "Enhanced intelligent smart home control and security system based on deep learning model," *Wireless communications and mobile computing*, *2022*,pp. 1-22, 2022.

[10] C. V. Amrutha, C. Jyotsna, and J. Amudha, "Deep learning approach for suspicious activity detection from surveillance video," *In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*,pp. 335-339, 2020.

[11] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end–edge–cloud surveillance in smart IoT," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12588-12596, 2021.

[12] L. Zahara, P. Musa, E. P. Wibowo, I. Karim, and S. B. Musa, "The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm-based Raspberry Pi," *In 2020 Fifth international conference on informatics and computing (ICIC)*, pp. 1-9, 2020.

[13] I. E. Olatunji, "Human activity recognition for mobile robot," *In Journal of Physics: Conference Series*, vol. 1069, no. 1, p. 012148, 2018.

Citation of this Article:

Hala Wael AlFadhel, "Advancing Security Measures in Governmental Institutions: Integration of Facial Recognition and Movement Monitoring Technologies in the House of Representatives", Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 8, Issue 5, pp 325-331, May 2024. Article DOI https://doi.org/10.47001/IRJIET/2024.805043

****

