International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (May-23)

http://dx.doi.org/10.12785/ijcds/1301108

Convolutional Neural Network (CNN) Model to Mobile Remote Surveillance System for Home Security

Rana Ayad[1] **and Farah Q. Al-Khalidi**[2]

[1,2]*Department of Computer, College of Science, Mustansiriyah University, Baghdad, Iraq*

Received 12 Aug. 2022, Revised 4 Mar. 2023, Accepted 10 Apr. 2023, Published 30 May. 2023

Abstract: Every citizen wants to eliminate any potential threat to themselves or their belongings. A proper and modern surveillance system is necessary given the enormous increase in the security needs of both persons and organizations. Remote surveillance is one of the major issues which attracted the attention of researchers recently. Surveillance cameras, commonly known as closed-circuit television (CCTV) have grown rapidly in popularity over the last few decades. Nowadays, video surveillance is quite essential. It greatly aids in reducing crime rates and can be used to keep track of the condition of buildings. In this paper, a remote surveillance system using the Convolutional Neural Networks (CNN) model was proposed. This system consists of a camera that takes more than one image of the object that passes in front of it and sends these images to the mobile phone of the intended person (the owner of the place). Images taken by the security camera may be analyzed using the CNN model depending on the region of interest for detection. The system will be delivering an alarm to the user depending on intelligent detection depending on a deep learning approach that can improve a smart home automation structure by detection people. The results of the experiments demonstrate a high degree of accuracy in detecting human beings , the accuracy of the system reached 100% in an ideal time.

Keywords: Convolutional Neural Networks, Internet of Things , Remote surveillance system, Mobile, Deep learning.

1. INTRODUCTION

Security systems are important to their personal safety. A house alarm system's purpose is to notify the homeowners of any unauthorized entry attempts into the area. Others access the home through ground-floor windows or the back entrance, which is more secluded and has less illumination than the front door [1] . Because of the advancement of new digital techniques and their integration with older conventional building methods, the smart house is rapidly becoming a reality. This is accomplished by integrating many subsystems into the surveillance system through a single control unit, including surveillance, theft control, and access control. The focus of Internet and Bluetooth-based surveillance systems is on managing and monitoring home surveillance systems from anywhere in the house [2] . Every organization, whether public or private, needs a low-cost, simple, and effective surveillance system for internal monitoring and post-event analysis [3]. Internet of Things (IoT) techniques enable the networking, interactivity, and data sharing of devices, sensors, and devices. Furthermore, overall remote control of the house is now available via the internet, Bluetooth, Wi-Fi, or other network elements via laptops, smartphones, tablets, and additional similar devices, giving users the essential comfortability in the applicable field of IoT. The IoT technology, often known as smart home automation, can monitor and regulate household activities and environmental conditions [4] When it comes to the IoT, self-driving automation systems have benefited greatly from the incorporation of Artificial Intelligence (AI). Smart home automation is essential for a relaxed way of life, and home security is a key component of it. By utilizing AI models and specific IoT technologies, homeowners can keep tabs on their home's infrastructure and evaluate any issues that may arise with it from afar [5] A Convolutional Neural Networks (CNN) model is an artificial intelligence subfield that is used to detect and analyze images. The CNN model, which works well with several convolution layers, is frequently employed to address image-based issues. Motion detection and surveillance are at the heart of a surveillance system that may be used in the house. The CNN model can analyze security camera images for detection depending on the region of interest [6] , [7] .

The goal is to create a low-cost, integrated deep learning surveillance platform to prevent crime in residential areas. The system can run on a Raspberry Pi or any inexpensive device with an internet connection and a webcam with a decent amount of pixels. Streams of live video will be recorded and analyzed. Deep neural networks will examine the video feeds. If any suspicious activity is discovered, users will be alerted instantly via a remote app. The app's users will then be able to take preventative measures to thwart criminal activity. Simply put, the suggested system employs cutting-edge technology and a well-thought-out project pipeline in an effort to improve the security of buildings.

The remains of the paper are arranged as follows: The subject of relevant studies in the field of remote surveillance systems using deep learning techniques is covered in Section 2. The suggested method is thoroughly described in Section 3. The evaluation metrics are illustrated in Section 4. The experimental outcomes of the proposed detection and classification model along with comparisons to similar research are presented in Section 5. The final conclusion of the method as well as its prospective applications in the future are provided in Section 6.

2. RELATED WORK

Many Convolutional Neural Network topologies have been presented in the literature to increase the accuracy of classification in the remote surveillance system [8]. The effectiveness of image categorization led to further research in object detection using neural networks. In this section, a set of relevant studies will be presented in chronological order from the oldest to the most recent. In 2017, Intriago-Pazmiño et al. [9] presents a comprehensive evaluation of works published in the last decade and identify, implement, and test the most-used and best-rated algorithms. A feature extraction approach similar to Histograms of Oriented Gradients (HOG) as well as two classification algorithms, AdaBoost and Support Vector Machine(SVM), have been discovered by us. 50 photos from Penn-Fudan were tested. Both the SVM-HOG and Adaboost-HOG detection rates are below the rates seen in similar works, at 0.66 and 0.72, respectively. While the accuracy of 0.96 in the SVM-HOG combination is on par with that of comparable works.

Xiao and Liu [10] examined visual saliency's potential for use in pedestrian picture detection. In order to train a Convolutional Neural Network, our method first generates a visual saliency map out from input image, and then feeds the result into the network. Faster R-CNN uses DSS to generate visual saliency map and detects pedestrians. Faster R-CNN was trained using Penn-Fudan, INRIA, and Daimler datasets. The experimental results show that the suggested method outperforms the state-of-the-art methods by a wide margin, with a 91% detection accuracy on the Penn-Fudan Dataset, a 15% miss rate on the INRIA Dataset, and a 28% miss rate on the Daimler Dataset. For the purposes of indoor location, activity tracking, and person recognition in smart building/home environments, Shi et al. [11] developed An intelligent floor monitoring system. Each step's location data activates lights at the appropriate locations while the entire walking signal is analyzed by a CNN model to predict if the person is a legitimate room user and automatically regulate door access. The smart floor monitoring system utilizes different gait-induced output signals for video-privacy protected, convenient, and secure individual recognition. A CNN model with 10 persons and 1000 data samples may reach 96.00% prediction accuracy based on their walking gaits, offering excellent accuracy in real-world applications.

In 2020, Irjanto and Surantha [12] employed the CNN AlexNet facial recognition technology, which is embedded in a door locking mechanism. Data is collected by utilizing a device to record 1048 facial data points on the homeowner's face, which is then utilized to train machine learning, and the outcomes are fairly accurate, with an accuracy of 97.5% to an earlier study. As a consequence, the CNN AlexNet technique can do highly accurate facial recognition and could be installed on an IoT device. A computer vision system based on CNN is developed by Navea [13] can evaluate if a person who comes into a location is permitted or not by employing face, height, and build recognition with gender issues. The developed method achieved more than 98% with balanced prediction accuracy. This is a supplementary technology that may be used in conjunction with automatic locks or security systems. Benito-Picazo et al. [14] also demonstrate a video surveillance system for detecting things in motion with anomalous behavior using a panoramic 360° surveillance camera. Its objective is to track and detect things in an environment, whether static or moving, that are regarded as abnormal or possibly harmful. This system detects and characterizes abnormal items in the scene using the CNN-based module designed for microcontroller architectures. While Khudhair and Ghani [15] focused on developing CNN, and the cloud to use in a video surveillance system. The system is made up of multiple nodes, each of which contains a central processing unit (CPU), The Raspberry Pi and camera nodes connect via client-server architecture. The captured footage is sent to the central node, where it is processed by a person-recognition algorithm trained on the Common Objects in Context (COCO) dataset. From there, it is uploaded to the cloud. Furthermore, the major node will alert security personnel of the finding of persons through SMS. According to the experimental results, the accuracy rate reached 96.1%.

In 2021, a mobile application is developed by Kumar et al. [16] to control house appliances. The project is built around the ARM11 Raspberry Pi microcontroller board. Python is the programming language used to create the software, which is an integrated development environment. A fan and a light bulb are used as loads to illustrate the functional prototype. The suggested model is additionally utilized for remote monitoring of the house or business by informing the user when there is any unauthorized entrance into the premises. To achieve the intended surveillance, a quadcopter-based system is offered by Iqbal et al. [17] to keep an eye out for suspicious behavior in a designated area, including the movement of armed individuals and facial detection. The suggested system produces an alert for security staff after detecting any suspicious behavior. In the initial feature extraction stage, a renowned Faster RCNN algorithm is tuned to learn quickly with feature reduction. The capacity of five distinct types of CNNs to detect items of interest in surveillance images was tested. The system's average accuracy is 79% across all categories. Malhotra and Chhabra [18] demonstrated a computer vision-based automated approach for detecting aberrant activity during examinations. The goal of this project is to use closedcircuit television (CCTV) cameras to watch for suspicious conduct in students during physical exams. The proposed system combines You Only Look Once (YOLOv3) with residual networks as the backbone architecture to detect suspicious activity in the test. The findings show that the suggested strategy is reliable and effective. The experimental findings show that the students' examination invigilation is effective. The detection rate of any dubious activity in the classroom was 88.03%. Gajjar et al. [19] demonstrate a visual saliency based approach for human detection. Saliency maps of people in a picture are extracted using a Deep Multi-Layer Network, then multiplied by the input picture and given to a CNN. We use the Penn-Fudan Dataset and the TudBrussels Benchmark to prepare DetectNet for its detection task. The network learns mid- and high-level body traits after training. We exhibit the efficiency of our strategy on human detection by achieving 91.4% accuracy on the Penn-Fudan Dataset and 53% on the TudBrussels Benchmark.

In 2022, Keerthana et al. [20] suggested CNNs as a method for detecting the object and spotting odd behavior close to the door. An electric door lock solenoid is required to unlock the door. If an ultrasonic sensor detects an approaching person over a certain threshold value, it only attempts to pick up a human image if there is no match for it in the database. An alarm message is delivered to the registered mobile phone number when a stranger attempt to open the door, allowing the owner to manage the door locking mechanism and check the mailer's image. The experimental results demonstrate that detection accuracy is more than 90%. Taiwo et al.[21] built a deep learning system for motion detection and classifications, allowing for smart home automation that can control appliances, check the weather, and detect movements inside and outside the house. A deep learning schema is utilized to create an algorithm that improves intruder detection and prevents false alarms in the smart home automation system. Depending on movement style, a human observed by the security camera is classed as an invader or a home inhabitant. The CNN classification model was 99.8% accurate.

3. METHODOLOGY

A deep learning CNN algorithm is employed to detect person movement as a security measure for motion detection [22] . The suggested system is divided into many phases, which will be discussed in detail later.

A. *Dataset Description*

Penn-Fudan dataset for pedestrian detection and segmentation is an image database that includes images

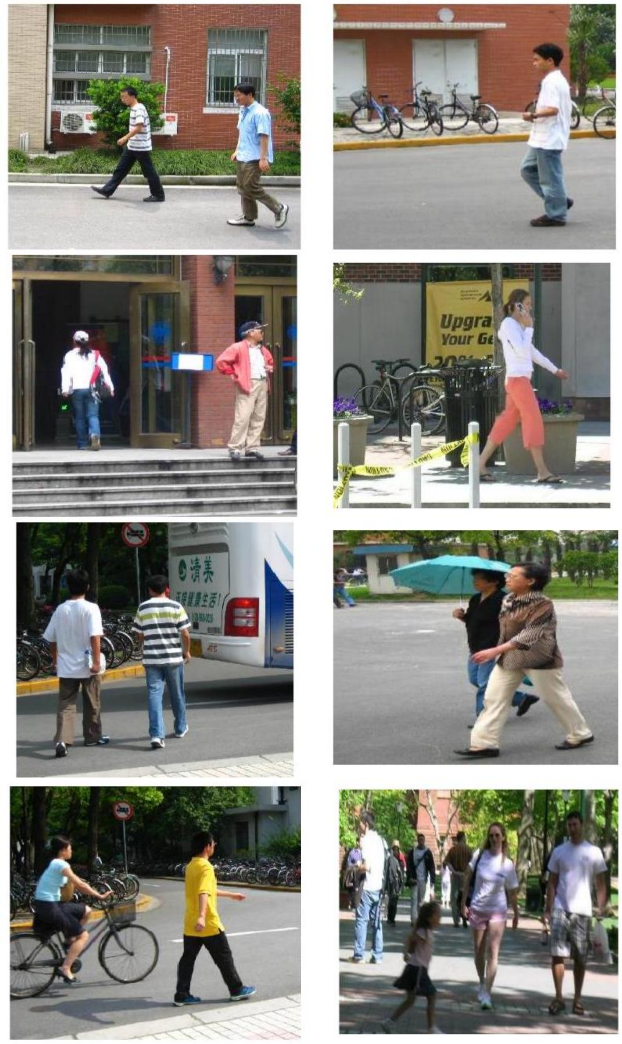

Figure 1. Penn-Fudan dataset Samples.

used in the published tests for pedestrian detection. The images were taken on campus and on city streets. The pedestrians in these images are of particular interest to us. At least one pedestrian will appear in each image. In this dataset, the heights of tagged pedestrians lie between [180,390] pixels. All of the pedestrians with labels are standing. There are 170 images with 345 tagged pedestrians, 96 from the University of Pennsylvania region and the others 74 from Fudan University. Figure 1 show some samples of the Penn-Fudan dataset downloaded from [23] . Figure 2 depicts the general diagram of the proposed surveillance system.

B. *Image Preprocessing Phase*

The primary goal of this stage is to improve the image data in addition to reducing undesired distortions or enhancing certain image attributes. The preprocessing process includes the following:

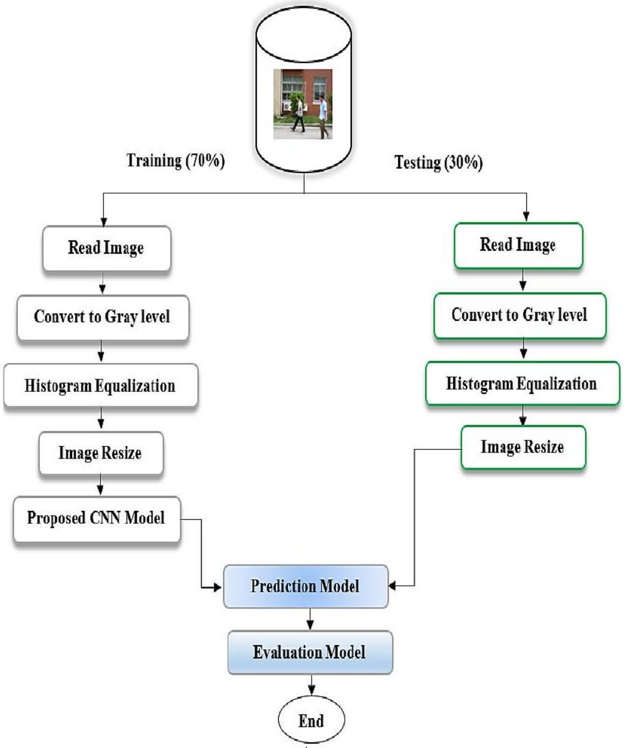

Figure 2. The proposed surveillance system's overall diagram.

Figure 3. RGB to grayscale image transformation

1) Convert images from RGB to gray levels

Grayscale images are a combination of two colors, white and black, with countless gray shades in between. The luminosity approach refers to the enhanced form of the average method. It also averages values, but it adds a weight value to accommodate for human comprehension, where the Green value weight has been seen as the greatest value of the weight [24], as shown in Eq. (1) .

$$\text{Gray Image} = 0.21 * R + 0.72 * G + 0.07 * B \qquad (1)$$

Figure 3 shows the application of this process to a sample of the dataset used in the paper.

2) Histogram Equalization Method

The histogram equalization (HE) method has long been regarded as a well-known method for improving image

Figure 4. image histogram equalization

Figure 5. Image resizing

contrast. The image histogram is a sort of histogram that serves as a graphic depiction of tonal distribution in a digital image. In the HE method, the cumulative distribution function (CDF) is used as the transformation curve for the image gray values. Let I and L denote an image and its gray levels, respectively. I(i, j) represents the gray value at the position with coordinates (i, j), N represents the total number of pixels in the image, and nk represents the number of pixels of gray level k. Then, the gray-level probability density function of image I is defined as p(k) = (k = 0, 1, 2, . . . , L-1), and the CDF of the gray levels of image I is $c(k) = \sum_{r=0}^{k} p(r)$.

Where k from 0 to (L - 1) . The standard HE algorithm maps the original image to an enhanced image with an approximately uniform gray-level distribution based on the CDF. The Histogram Equalization for all dataset images is determined in this stage [25] . The histogram was done by applying Eq. (2) as follows:

$$f(k) = (L - 1) * c(k) \qquad (2)$$

3) Resize the image

As an alternative to the typically utilized bilinear interpolation, bicubic interpolation can be used in realtime rendering to make textures look better when scaled. All images were resized and cropped to 200×200 pixels using bicubic interpolation as shown in Figure 5.

C. Image Classification and Object Detection with CNN

CNN is the most well-known and commonly utilized deep learning algorithm. CNN has a fundamental benefit over its predecessors in that it detects significant traits without the necessity for human interaction [26], [27]. The CNNs have been widely employed in a wide range of applications, consisting of computer vision, speech recognition, and person detection. The CNNs, like traditional Neural Networks (NNs), are informed by neurons found in human and animal brains. CNN has three major advantages: comparable presentations, sparse interactions, and parameter sharing [28]. Because object detection systems, video surveillance, vehicle monitoring, and autonomous vehicle driving are becoming increasingly popular, rapid and precise object detection systems are in high demand [29]. Object detection typically generates a bounding box around an object with the confidence value specified. Single-class object detection occurs when just one object is detected in a given image. In multi-class object detection, more than one item from different classes must be found [30] , [11] . This section will go through the classification and detection of objects using the suggested CNN model in depth. The localization procedure is then based on the Intersection Over Union (IOU) approach.

1) The Proposed CNN Model Layers

The CNN architecture is made up of numerous tiers. The proposed CNN architecture's layers are depicted in Figure 6 below:

a) Convolutional Layer: The convolutional layer is still the core part of CNN design. It is consisting of several convolutional filters (so-called kernels). Some of these filters are convolved with the input image to generate an output feature map described as N-dimensional metrics [31].

b) LeakyReLU: Rather than downscaling negative inputs as ReLU does, this activation function enables them to be never ignored. It is used to address the issue of Dying ReLU. Leaky ReLU may be formally defined as follows in Eq. (3):

$$f(x)_{LeakyReLU} = \left\{ \begin{array}{cc} x & if X > 0 \\ mx & if x \leq 0 \end{array} \right\} \tag{3}$$

The leak factor is indicated by the symbol m. It is often set to an extremely low number, such as 0.001 [32].

c) Pooling Layer: The pooling layer's primary function is to subsample the feature maps. These maps are created using convolutional methods. In other words, this method replaces large-scale feature maps with smaller feature maps. During the pooling phase, it keeps the bulk of the dominant data (or features) [33].

d) Fully Connected Layer: This layer is usually seen at the end of any CNN structure. Every neuron inside this layer is connected to every neuron in the previous layers by utilizing the Fully Connected (FC) method. It serves as the CNN classifier. As a feed-forward ANN, it adheres to the fundamental methodology of the typical multiple-layer perceptron neural network. The FC layer gets input from the preceding pooling or convolutional layer. This is a vector created after fattening the feature maps. The FC layer output represents the final CNN output [34].

e) Output Layer: The output layer, which is the final layer of the CNN structure, provides the final classification. Certain loss functions are employed in the CNN model's output layer to predict the anticipated error across the training data. This error demonstrates the difference between the actual and expected output. The CNN learning approach will then be utilized to optimize it [35]. Algorithm 1 explains the overall proposed system in detail.

2) Object localization

To localize an object in the proposed remote surveillance system, the Intersection Over IOU method was used [36], [37]. In IOU, both real and predicted bounding box variables are used, and also the IOU for both is determined using the equations in Eq. (4), as shown below:

$$IOU = \frac{IntersectionArea}{UnionArea} \tag{4}$$

4. Mobile Application

Mobile app frameworks now have easier access to home automation thanks to the proliferation of smartphones, Wi-Fi, and the Internet of Things. Users may check in on what's happening at home in real time and manage their Internet of Things gadgets from anywhere using the remote surveillance system app. A remote notice will be sent to the app whenever a person is detected, allowing users to take rapid action with any Internet of Things devices they have enabled. Those who have downloaded the mobile app for a remote surveillance system will be able to subscribe to and get notification when a person is detected.

5. EVALUATION METRICS

Two metrics have been utilized to evaluate the system. In this section, we explain these metrics as follows:

A. *Accuracy rate*

The probability that any pixel in that category has been correctly classified is defined as accuracy [38], [39]. It is determined in Eq. (5) as follows:

$$Accuracy = \frac{Numberoftestsiscorrect}{Totalnumberoftests} \tag{5}$$

B. *Loss rate*

This value is calculated as shown in Eq. (6) as follows:

$$LoSS = \frac{Numberoftestsiswrong}{Totalnumberoftests} \tag{6}$$

6. RESULTS AND DISCUSSION

Graphs will demonstrate the various results that were produced after the proposed technique was used to detect a person ; these results will be compared to reveal the best outcome among the techniques that were utilized in comparison. The results of the proposed system are compared

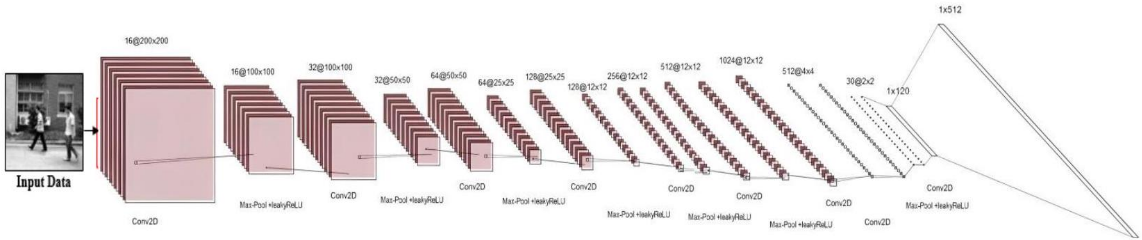

Figure 6. The Proposed CNN Model.

Algorithm 1. *The Proposed Remote Surveillance System*

Input: Dataset
Output: Evaluate the Proposed CNN Model
Begin
// Input//
1: Load dataset.
2: Apply Holdout splitting into (training set (70%), testing set (30%))
// Training Phase //
3: Pre-Processing phase
- Convert RGB image to grayscale image.
- Apply Histogram Equalization using Eq. (2)
- Resize Image.

4: Classify and detect an object by the proposed CNN model.
//Testing Phase//
// Classification Phase //
5: Classify Samples
- Results (1) = CNN (Features Set, Targets)

// Output//
6: Test the remainder of the data that have been entered for training.
7: Calculate the location of the person (object) using "Eq. 4".
8: Real time footage capture by camera analyse based on step (4) when person detected.
9: Send an alarm and person's pictures to the owner of the home, its content is that there is a person in the house.
10: End Procedure.

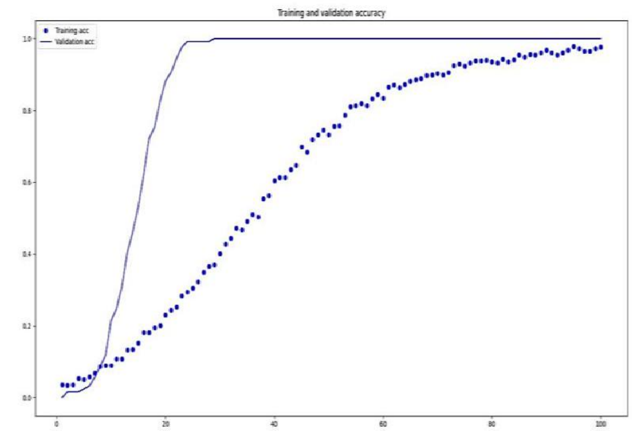

Figure 7. The accuracy rate of the proposed system.

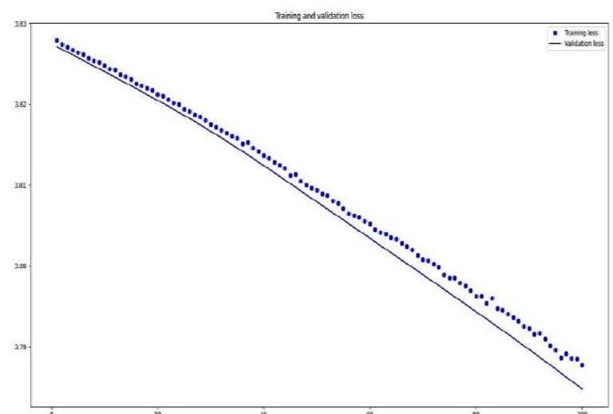

Figure 8. The loss rate of the proposed system.

in this section with the findings of earlier research, which were discussed in Section 2 and are displayed in Figure 7 and 8.

Figures 7 and 8 shows how well the CNN architecture performs during training. With a graph of the 100 training epochs' worth of loss values, derived from both the training and validation samples. the CNN model's training dropped progressively throughout the first 20 epochs, but the validation samples displayed a highly unstable pattern. the results obtained between epochs 40 and 100 are particularly intriguing since they show that the CNN model successfully learns the classification issue and that the loss values for training and evaluation are in sync. in a similar vein, when the loss values increased, the training and evaluating phases' accuracy also improved. around 60-100 epochs into the training procedure, the accuracy values for both phases stabilized.

From Table 1 , by comparing the results we obtained with the results of the relevant studies, we find that the proposed model is superior in the accuracy of detecting persons by 100%.

TABLE I. Accuracy Comparison with related works.

Approach	Accuracy
SVM-HOG [9]	96.00
Adaboost-HOG [9]	72.00
Xiao and Liu [10]	91.00
CNN [11]	96.00
Gajjar et al.[19]	91.4
CNN-MufHAS [21]	99.80
Proposed	100.00

7. Conclusions and Future Work

This study is concerned with developing a security and surveillance system based on deep learning for a firm, university, or other critical location in real-time, the surveillance system synchronizes real-time data with the mobile notification. By notifying the proper agencies of any concern, the surveillance system assures individual and community protection. This research proposes utilizing CNN to detect person motion as a security measure for the detection of persons. In the context of a smart house, the suggested detection of movement is meant to detect someone passing in front of the camera, capture an image of him/her, and then transmit an alarm to the homeowner. The experimental results on the Penn-Fudan dataset give a detection accuracy rate reached 100%.

8. ACKNOWLEDGMENT

The authors would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq), Baghdad-Iraq for its support in the present work.

References

[1] L. Ajao, J. Kolo, E. Adedokun, O. Olaniyi, O. Inalegwu, and S. Abolade, “A smart door security-based home automation system: an internet of things,” *SciFed Journal of Telecommunication*, vol. 2, no. 2, 2018.

[2] Z. Peng, X. Li, and F. Yan, “An adaptive deep learning model for smart home autonomous system,” in *2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*. IEEE, 2020, pp. 707–710.

[3] F. Mehmood, I. Ullah, S. Ahmad, and D. Kim, “Object detection mechanism based on deep learning algorithm using embedded iot devices for smart home appliances control in cot,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–17, 2019.

[4] A. K. Gupta and R. Johari, “Iot based electrical device surveillance and control system,” in *2019 4th international conference on internet of things: Smart innovation and usages (IoT-SIU)*. IEEE, 2019, pp. 1–5.

[5] L. W. Yang and C. Y. Su, “Low-cost cnn design for intelligent surveillance system,” in *2018 International Conference on System Science and Engineering (ICSSE)*. IEEE, 2018, pp. 1–4.

[6] Y.-H. Chang, P.-L. Chung, and H.-W. Lin, “Deep learning for object identification in ros-based mobile robots,” in *2018 IEEE International Conference on Applied System Invention (ICASI)*. IEEE, 2018, pp. 66–69.

[7] M. Waqas, S. Tu, Z. Halim, S. U. Rehman, G. Abbas, and Z. H. Abbas, “The role of artificial intelligence and machine learning in wireless networks security: principle, practice and challenges,” *Artificial Intelligence Review*, pp. 1–47, 2022.

[8] Z. Liu, M. Waqas, J. Yang, A. Rashid, and Z. Han, “A multi-task cnn for maritime target detection,” *IEEE Signal Processing Letters*, vol. 28, pp. 434–438, 2021.

[9] M. Intriago-Pazmiño, V. Vargas-Sandoval, J. Moreno-Díaz, E. Salazar-Jácome, and M. Salazar-Grandes, “Algorithms for people recognition in digital images: A systematic review and testing,” in *World Conference on Information Systems and Technologies*. Springer, 2017, pp. 436–446.

[10] F. Xiao and B. Liu, “Pedestrian detection using visual saliency and deep learning.” *Acta Microscopica*, vol. 27, no. 4, 2018.

[11] Q. Shi, Z. Zhang, T. He, Z. Sun, B. Wang, Y. Feng, X. Shan, B. Salam, and C. Lee, “Deep learning enabled smart mats as a scalable floor monitoring system,” *Nature communications*, vol. 11, no. 1, pp. 1–11, 2020.

[12] N. S. Irjanto and N. Surantha, “Home security system with face recognition based on convolutional neural network,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, 2020.

[13] R. F. Navea, “Room surveillance using convolutional neural networks based computer vision system,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, p. 6700, 2020.

[14] J. Benito-Picazo, E. Dominguez, E. J. Palomo, and E. Lopez-Rubio, “Deep learning-based video surveillance system managed by low cost hardware and panoramic cameras,” *Integrated Computer-Aided Engineering*, vol. 27, no. 4, pp. 373–387, 2020.

[15] A. B. Khudhair and R. F. Ghani, “Iot based smart video surveillance system using convolutional neural network,” in *2020 6th International Engineering Conference “Sustainable Technology and Development”(IEC)*. IEEE, 2020, pp. 163–168.

[16] V. Nandalal, M. Kousalya, P. Madhumitha, R. Kamaleshwari, N. K. Selvi *et al.*, “Implementation of smart home assistance and surveillance,” in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1. IEEE, 2021, pp. 1117–1121.

[17] M. J. Iqbal, M. M. Iqbal, I. Ahmad, M. O. Alassafi, A. S. Alfakeeh, and A. Alhomoud, “Real-time surveillance using deep learning,” *Security and Communication Networks*, vol. 2021, 2021.

[18] M. Malhotra and I. Chhabra, “Automatic invigilation using computer vision,” in *3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021)*. Atlantis Press, 2021, pp. 130–136.

[19] V. Gajjar, Y. Khandhediya, A. Gurnani, V. Mavani, and M. S. Raval, “Vis-hud: using visual saliency to improve human detection with convolutional neural networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1908–1916.

[20] T. Keerthana, K. Kaviya, S. D. Priya, and A. S. Kumar, “Ai enabled smart surveillance system,” in *Journal of Physics: Conference Series*, vol. 1916, no. 1. IOP Publishing, 2021, p. 012034.

[21] 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*, vol. 2022, 2022.

[22] T. Zhang, M. Waqas, Z. Liu, S. Tu, Z. Halim, S. U. Rehman, Y. Li, and Z. Han, "A fusing framework of shortcut convolutional neural networks," *Information Sciences*, vol. 579, pp. 685–699, 2021.

[23] [Online]. Available: https://www.kaggle.com/datasets/psvishnu/pennfudan-database-for-pedestrian-detection-zip

[24] J. Coady, A. O'Riordan, G. Dooly, T. Newe, and D. Toal, "An overview of popular digital image processing filtering operations," in *2019 13th International conference on sensing technology (ICST)*. IEEE, 2019, pp. 1–5.

[25] W. Wang, X. Wu, X. Yuan, and Z. Gao, "An experiment-based review of low-light image enhancement methods," *Ieee Access*, vol. 8, pp. 87 884–87 917, 2020.

[26] T. Wang, Z. Liu, T. Zhang, S. F. Hussain, M. Waqas, and Y. Li, "Adaptive feature fusion for time series classification," *Knowledge-Based Systems*, vol. 243, p. 108459, 2022.

[27] D. Liciotti, M. Bernardini, L. Romeo, and E. Frontoni, "A sequential deep learning application for recognising human activities in smart homes," *Neurocomputing*, vol. 396, pp. 501–513, 2020.

[28] D. Popa, F. Pop, C. Serbanescu, and A. Castiglione, "Deep learning model for home automation and energy reduction in a smart home environment platform," *Neural Computing and Applications*, vol. 31, no. 5, pp. 1317–1337, 2019.

[29] A. Brenon, F. Portet, and M. Vacher, "Arcades: A deep model for adaptive decision making in voice controlled smart-home," *Pervasive and Mobile Computing*, vol. 49, pp. 92–110, 2018.

[30] A. Reghunath, S. V. Nair, and J. Shah, "Deep learning based customized model for features extraction," in *2019 International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2019, pp. 1406–1411.

[31] M. A. Saleem, N. Senan, F. Wahid, M. Aamir, A. Samad, and M. Khan, "Comparative analysis of recent architecture of convolutional neural network," *Mathematical Problems in Engineering*, vol. 2022, 2022.

[32] X. Zhang, Y. Zou, and W. Shi, "Dilated convolution neural network with leakyrelu for environmental sound classification," in *2017 22nd international conference on digital signal processing (DSP)*. IEEE, 2017, pp. 1–5.

[33] M. Sahu and R. Dash, "A survey on deep learning: convolution neural network (cnn)," in *Intelligent and Cloud Computing*. Springer, 2021, pp. 317–325.

[34] M. Khan, J. Seo, and D. Kim, "Towards energy efficient home automation: a deep learning approach," *Sensors*, vol. 20, no. 24, p. 7187, 2020.

[35] E. N. Kajabad and S. V. Ivanov, "People detection and finding attractive areas by the use of movement detection analysis and deep learning approach," *Procedia Computer Science*, vol. 156, pp. 327–337, 2019.

[36] Y. Qing, W. Liu, L. Feng, and W. Gao, "Improved yolo network for free-angle remote sensing target detection," *Remote Sensing*, vol. 13, no. 11, p. 2171, 2021.

[37] S. Hussain and F. Al-Khalidi, "Eyes detection in the human face," *International Journal of Civil Engineering and Technology (IJCIET)*, vol. 9, no. 10, pp. 1001–1007.

[38] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," *arXiv preprint arXiv:2008.05756*, 2020.

[39] H. M. Fadhil, M. N. Abdullah, and M. I. Younis, "A framework for predicting airfare prices using machine learning," *The Iraqi Journal of Computers, Communications, Control, and Systems Engineering (IJCCCE)*, vol. 22, 2022.

Rana Ayad received the B.Sc. degree in Computer Science from AL-Rafidain University College in 2001, For years 2011 to 2020 , she teaches in private school, she is currently Master student to obtain a degree in computer science.

Dr. Farah Qais Al-Khalidi is an Assistant Professor, who joined the Department of computer science at Mustansiriyah University in 2010. she has a B.Sc. degree in computer science (image processing), she received an M.Sc. degree from Technology University in 2003 and received Ph.D. degree in information technology from Sheffield Hallam University (UK) in 2012., she is interested in different fields such as Computer science, image processing, information technology, Graphics, thermal imaging, Computer Sciences Department, Section Editor at AlMustansiriyah Journal of Science. faculty of Information Technology at Mustansiriyah University