Date of publication JUN-30, 2023, date of current version FEB-20, 2023.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.22.2.3095

Deep Learning Algorithm for Detecting and Analyzing Criminal Activity

RADDAM SAMI MEHSEN
Middle Technical University, Technical Institute of Baqubah, Baghdad, Iraq
Corresponding author: Raddam Sami Mehsen (e-mail: radd1980@mtu.edu.iq).

ABSTRACT When applied to an entire field, automation and autonomous systems are among the rare creative superpowers capable of catapulting progress at an exponential rate. The arrival of machine intelligence will give such automated machines the intelligence to perform their tasks with power of outcome, drastically reducing the need for human intervention in redundant processes. Large-scale technological progress can be traced back to responsibilities that are simplified and, as a result, more easily distinguished by means of automation. In accordance with these guidelines, we propose creating a product that eliminates or significantly reduces the need for human intervention in primary issue statements that can be automated and processed. The public safety infrastructure of today relies on surveillance cameras, but these devices are merely video recorders; they have no intelligence of their own. Automated video streams are now required for automatic event detection thanks to the massive amount of data produced by surveillance cameras. The project's main objective is to increase public safety through the mechanization of crime measurement and review using actual Closed-Circuit Television footage (CCTV). This is achieved by assigning the task of recognizing criminal behavior to a system that can do so automatically, allowing for more precise tracking. In this study, we present a model with a precision of 0.95 for assault and 0.97 for abuse.

KEYWORDS deep learning; data forensic; digital security; digital crime; criminal activity.

I. INTRODUCTION

THE need to solve and mechanize supervised classification of live-streaming data in real-time prompted the presentation of the machine vision problem of image classification. Due to the novelty of the issue, there may be workarounds that have not been tried and tested. Beyond that, such applications provide a wide range of solutions, including the initial detection of significant sports actions or everyday activities occurring in a scene, as well as various security and health activities [1]. The goal of our study is to increase public safety by developing an automated system for measuring and analyzing criminal activity. This system will be able to distinguish between normal and criminal behavior based on inferred patterns, freeing humans from the burden of identifying criminal behavior. The current state of detection technology has several drawbacks that prevent it from working with today's widely available infrastructure [2-4]. The need for a watchful supervisor to review footage and ensure that any unusual activity is properly detected and addressed is a major weakness of conventional surveillance systems. Inaccuracies may occur when a person reviews CCTV footage [3, 38]. The proposal not only eliminates the need for extra guidance in order to reduce human input and labor, but it also instantly recognizes the type of crime that is taking place, notes the people involved, and takes immediate steps to start mitigation strategies at the crime scene. Deep learning techniques, particularly CNN architectures like Residual models, can be used to automatically identify many crimes. That can detect the location of the perpetrator and the weapon in a video. Creating a system of automated activity analyzers to increase public safety is simplified by the availability of large datasets like the UCF-crime collection [3] and the RWF-2000 database [4].

II. LITERATURE REVIEW

Numerous related papers and examples of their use in practice are discussed here. In order to better detect criminal activity, researchers have built a pipeline to recognize firearms from photographs by training them on classifier models. VGGNet 19 was used as the pre-trained model for detection, and results showed an accuracy of 69% and a recall of 75%. A technique was proposed for classifying the presence of weapons in surveillance footage and using that information to establish whether or not a crime took place. Region-based Convolutional neural network (RCNN) and faster region-based convolutional network (FRCNN) models were trained using the researchers' data. This research used low-quality films that accurately predicted high crime rates. As an introduction to common techniques for locating and identifying action at sporting events, researchers provide an overview of the field. The authors proposed segmenting the activity recognition pipeline into three distinct stages: feature extraction, deep learning depiction of clips, and sport classification [5-7]. Using the UCF-supplied Sports dataset as a benchmark, they evaluated the contentious issue. Researchers in their investigation of the CNN structures proved the effectiveness of video categorization when using convolutional neural networks (CNNs) [8]. Researchers outperformed other techniques in terms of reliability and ability to comprehend strong features from sparsely labeled data. Studies of transfer learning demonstrate the generalizability of various categorization tasks and imply that the acquired characteristics are generic. Researchers look into a similar objective by studying how to predict and classify criminal behavior using techniques like Decision Trees and Naive Bayes Classifiers, which are part of machine learning. Datasets here make use of geographical information. For data collected in Los Angeles and Denver, they achieved an accuracy of 54% and 51%, respectively [9, 39].

III. PROPOSED METHOD

One of the biggest issues with standard surveillance systems is their reliance on an attentive supervisor to watch film and make sure that any unusual activity is properly recorded and addressed. CCTV footage needs to be reviewed by a human, which could result in mistakes [10]. The proposed work aims to eliminate the need for extra supervision in an effort to reduce manual intervention and labor, but it also instantly recognizes the type of crime that is taking place, takes note of the people involved, and in response, it sets off actions that start taking care of the crime scene right away [11].

Identifying criminal activity in surveillance camera feeds is the primary focus of this paper. Further, a separate module utilizes a technique called Triplet Loss, which has already been implemented, to identify faces in these streams. The paper proposes two modules to illustrate the issue. The first option is to use the techniques outlined here to identify individuals in a CCTV feed [12]. Next we will examine how to use deep learning techniques for crime detection. The paper's two modules are depicted in a high-level overview in Figure 1.

The DNN (Deep Neural Networks) module in Open-built-in CV is used by the Face Recognition feature to train a model specifically for facial recognition. Embedding training with the Triplet Loss Function [13, 30] allows us to do this by first locating faces in the input data, then cleaning and preparing the data, and finally training. After calculating these embeddings, the model can be used to identify the faces in question. All that needs to be done is to load the model and use a webcam to identify the face by drawing bounding boxes around it and giving a confidence parameter. This paper accomplished a lot in terms of evaluating this module. The input videos for the crime detection module are first sorted to distinguish between those containing criminal activity and those without. It is trained using a ResNet architecture [14] that already exists. This architecture measures accuracy and other metrics after basic preprocessing steps like adding more data and turning videos into image frames. Finally, the trained model allows crime detection via mobile phone webcam simulation. In this paper, evaluation for a total of six classes is accomplished at an accuracy of over 90%. Altogether, this paper suggests a full pipeline that integrates these two parts to determine the perpetrator simultaneously.

IV. FACE RECOGNITION

The process of face detection involves identifying and returning the position of a face within a picture or video. The next step in face verification is to check if the image of the face being presented matches one already in the database. To do this, we employ distance metrics such as the L2 norm or cosine similarity to determine the degree to which two faces are alike. Finally, face recognition uses both of these methods to extract salient facial features and assign those features to one or more labels from the dataset used to train the model. In this paper, we propose a way to recognize faces by finding faces, computing face embeddings, training a Support Vector Machine (SVM) [15, 40] on the given embeddings, and then finding faces in images or simulated video streams. Figure 2 shows the pipeline that is talked about in this paper. Caffe and Open Face Models are responsible for face detection and feature extraction, respectively.

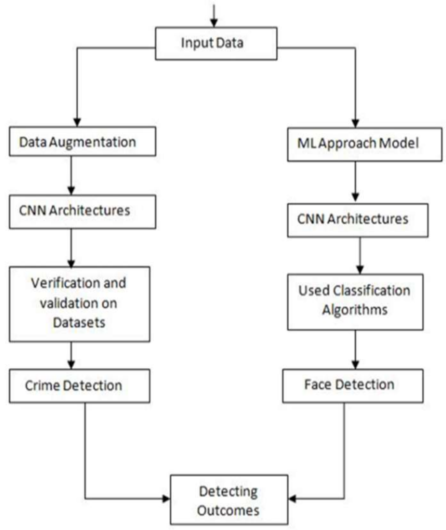

Figure 1. Crime and Face Detection Model

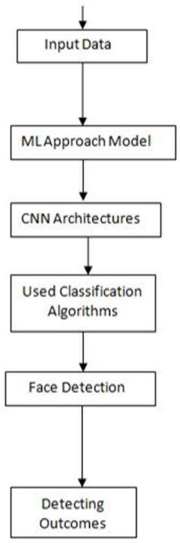

Figure 2. Procedure of Face Detection model

OpenCV uses Single Shot Detector (SSD) architecture [16] and a ResNet to perform deep learning face detection. Single-shot detection is a method whereby multiple objects can be located in a single image with a single model training it. The given image is discretized into the various boxes that are generated around the regions with high confidence feature maps [17, 45]. After determining the level of certainty associated with each box, their sizes are modified until the best possible detection fit is achieved. When a face is recognized, the final bounding boxes are shown in Figure 3.

Figure 3. Bounding Box for Localization of a given face

More so, we can preprocess images and carry out face alignment on datasets with enhanced outcomes by using the dlib library to identify facial markers, including the mouth, right and left brows, eyes, nose, and jawline [18]. After performing some basic editing and alignment on the provided face, we feed it into the proposed neural network. Every input batch needs to have three images: a Positive Picture (another image of person "A"), a Negative Image (the current perception of person "A"), and an Anchor Image (any other image that is not person "A"). The neural network uses triplet loss to calculate the face embedding and fine-tune the weights. As a result, the embedding of the "Anchor" and "Positive" photos are relatively close together, however the embedding of the "Negative" image is more apart. The facial recognition pipeline begins with the computation of face embedding using a convolutional neural network (CNN) (Caffe) model; these embeddings are sufficiently distinct from one another to enable the training of a classifier on top of the computed face embedding (such as Random Forests, SGD Classifiers, SVMs, and so on) [19].

A. DATA AUGMENTATION:

In order to maximize the usefulness of a little amount of data, our research suggests data augmentation techniques [20] to identify patterns in our data. Images can now be off flipped, rotated, zoomed, translated, scaled, cropped, moved along the x and y axes, subjected to shearing, skewing, filtered in black and white, and blurred, among other effects. Figure 4 displays our dataset augmented version.

Figure 4. Normal Image, Rotated Left, Rotated Right

B. CRIME-RELATED VIDEO CATEGORIZATION

The purpose of this component is to identify outliers from typical behavior and video footage. Because it is the only dataset with recordings of many sorts of crimes, the UCF Crimes Dataset is used for training [21, 44], each of which has its own unique characteristics. The dataset includes 13 different types of incidents: accidents, fights, burglaries, shoplifting, robberies, shootings, abuse, arrests, arsons, assaults, thefts, explosions, and vandalism. In all, there are about 1,900 pieces of actual stuff in there. Figure 5 and Figure 6 show sample frames.

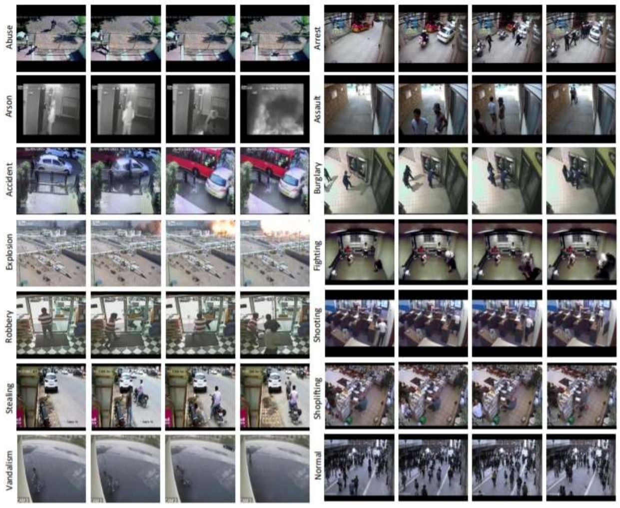

Figure 5. Video from UCF dataset being turned into photos for preparation

Figure 6. Sample anomalous frames from UCF Crime Dataset (a) Abuse (b) Arson (c) Explosion (d) Fight (e) Road Accident (f) Shooting

- Preparing and preprocessing the dataset:

Each scenario uses one of three methods: converting, enriching, and augmenting the data. Abuse, assault, fighting, normal, robbery, and vandalism are the six broad categories used to organize the criminal acts examined in this research. The authors reduced the number of dataset classes because the provided movies were too large. With the help of video editing and trimming, five-minute-long videos were condensed to forty-five seconds, with the emphasis laid on the time of the actual incident, rather than on irrelevant or misleading material. Some parts of the crime scene were highlighted in low-resolution videos that were cropped and sharpened. Each video clip of a criminal act is individually tagged by hand. After that point, the rest of the video feed can be considered canonical.

Finally, data augmentation was carried out by increasing the variety of input data for a given model's training. An illustration of data augmentation is shown in Figure 7.

- The residual network (ResNet):

The ResNet layers are created so that they are formulated as learning residual functions with relation to the layer inputs rather than learning unreferenced functions [21–24]. The signal cannot be transmitted from one convolutional layer to the next since each of the 18–152 layers has a short connection between them. These links can carry gradient flows from the first to the last layer of a network, simplifying the training of extremely deep neural networks. Figure 7 for residual block shows how the link suppresses the signal moving from top to bottom (below). The designers were able to resolve the vanishing gradient issue using skip connections [25, 36–37] and the notion of a Residual Network. As a result, it might create a direct connection to the output and skip training on a few levels.

- Technique for Automatically Categorizing Videos:

In order to classify crimes using the proposed pipeline, it is necessary to iterate over each frame of the input video. Similar frames are fed into a convolutional neural network [26], and the results are categorized independently. The model selects the highest probability label before writing the output frame. The previously mentioned solution, which only takes into account a single frame, will not work for us because our problem is sequential. There must be some kind of connection kept alive between successive frames of a single video input for the purposes of crime detection.

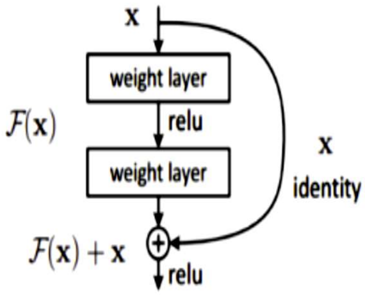

Figure 7. Residual Block

This is accomplished by keeping track of the most recent "N" forecasts and computing the aforementioned for a given time frame. The pipeline [27] takes these into account when determining which label has the highest probability by averaging the last "N" predictions, and then returns the result.

- Using the UCF-Crimes Dataset to train the ResNet:

Here, we go over the steps involved in ResNet's training and testing processes. Finding the image folders used for training and testing is the first order of business [28]. Parameters for training are also provided, including batch size, epoch count, image width and height, and learning rate. The Tensor Flow Image Data Generator [29, 35, 41-43] component is used to produce test and train sets. Each epoch's worth of training data is recorded and plotted afterward. The model and its weights are now preserved. The model's hyper-parameters, such as the learning rate, the number of epochs, and the batch size, are specified before training begins. While training a CNN [30, 46], it is important to keep in mind the optimal values for its hyper-parameters. The final training epochs are displayed in Figure 8. The ResNet Module was supplied with the information listed below [31]:

- Dataset Splitting: 25% of dataset is kept as test data and for model training kept remaining 75%.
- Total Epochs = #50
- Categorical Cross-Entropy = Loss
- Stochastic Gradient Descent Optimizer = 0.0001 (Learning Rate)
- Accuracy as Metric

The following factors are used in the analysis and conclusion [31-33]:

- However, no criminal activity actually takes place at the location, resulting in a **False Positive** in the model.
- One type of error is called a "**false negative**," which occurs when the model does not detect a crime.
- One definition of a "**True Positive**" is when the model reliably detects the offending event.
- The **true negative** scenario is one in which neither the crime nor the failure to detect it occurs.

V. RESULTS

A. FACIAL RECOGNIZATION:

Using roughly 15 photos of each participant, we trained facial recognition software to distinguish between three different faces. For this module, the authors had to do with as little data as possible due to constraints in both their financial and computational capacities; the results are depicted in Figure 8 below, which shows the use of a camera to simulate a CCTV stream. The model accurately identified each face and provided a confidence score and bounding box for its classification.

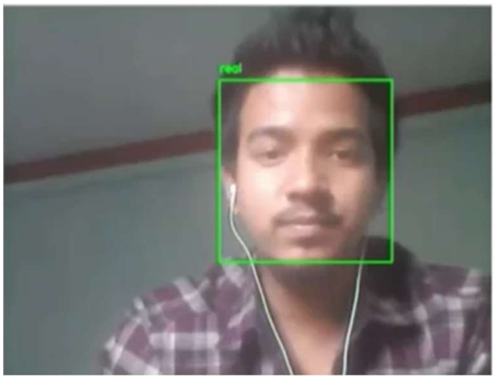

Figure 8. Recognizing Person 'B' through Face (simulating a CCTV)

B. METHOD OF CRIME SPOTTING

With respect to training Tables I and Table II display the results of testing the model over 50 epochs. Both the training loss and the validation accuracy were recorded. We evaluated their trained model with training dataset, looking for words like "fighting," "vandalism," and "abuse." This resulted in the model's accurate forecasting of the subsequent events in the stream [34, 47-48]. It is shown below in Figure 9.

```
Epoch 44/50
205/205 [==============================] - 97s 471ms/step - loss: 0.1779 - accuracy: 0.9497 - val_loss: 0.1266 - val_accuracy: 0.9623
Epoch 45/50
205/205 [==============================] - 97s 475ms/step - loss: 0.1821 - accuracy: 0.9469 - val_loss: 0.1257 - val_accuracy: 0.9646
Epoch 46/50
205/205 [==============================] - 98s 476ms/step - loss: 0.1722 - accuracy: 0.9484 - val_loss: 0.1233 - val_accuracy: 0.9651
Epoch 47/50
205/205 [==============================] - 97s 473ms/step - loss: 0.1741 - accuracy: 0.9486 - val_loss: 0.1224 - val_accuracy: 0.9646
Epoch 48/50
205/205 [==============================] - 101s 494ms/step - loss: 0.1669 - accuracy: 0.9503 - val_loss: 0.1207 - val_accuracy: 0.9665
Epoch 49/50
205/205 [==============================] - 105s 514ms/step - loss: 0.1677 - accuracy: 0.9500 - val_loss: 0.1179 - val_accuracy: 0.9678
Epoch 50/50
205/205 [==============================] - 181s 886ms/step - loss: 0.1581 - accuracy: 0.9535 - val_loss: 0.1168 - val_accuracy: 0.9669
```

Figure 9. Training the classifier (for total 50 Epochs)

Table 1 shows the accuracy metrics for the proposed system where Table 2 shows the resultant performance metrics for different classes. Figure 10 shows the training loss and accuracy as well as the validation loss and accuracy that have been recorded.

Table 1 Accuracy Metrics

Classes	Recall	Precision	F1-score	Support
Accuracy	--	--	0.98	2199
Weighted avg	0.97	0.98	0.97	2199
Macro avg	0.92	0.96	0.92	2199

Table 2 Results for the Different Classes

Classes	Recall	Precision	F1-score	Support
Abuse	0.95	**0.97**	0.97	237
Fighting	0.98	0.92	0.93	394
Normal	1.00	1.00	1.00	1010
Robbery	0.99	0.96	0.97	335
Vandalism	1.00	0.98	0.99	173
Assault	0.53	**0.95**	0.68	59

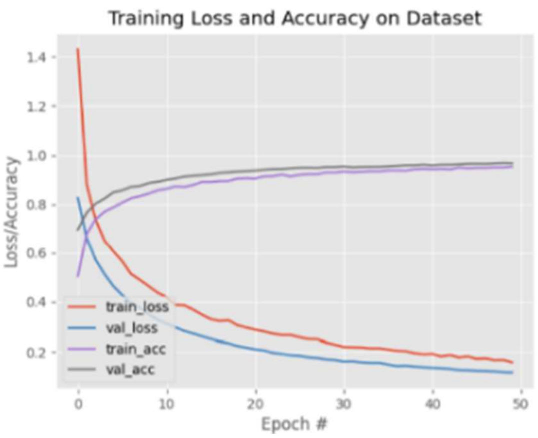

Figure 10. Accuracy and Loss function

(a)

(b)

(c)

(d)

Figure 11. Detecting Crime in the footage

VI. CONCLUSION

The project's face recognition module replicated a facial recognition model and showcased a virtual version of the final prototype. After all the data had been prepared, the Crime Detection module trained the ResNet Model for several classes and showed good results. It is possible that in the future, additional facial classes will be used to train the face recognition model. It would be fascinating to compile a library of these faces for later use in identifying victims in real time from CCTV footage. Examining the differences between the performances of various facial recognition models would be fascinating as well. The model is trained in the way that the person involved in the activity can also be detected and being useful for further investigation. It is also important to detect the criminal activity at real time from CCTV footages. The paper investigates a limited classes due to computational constraints, however aims to cover more in future. The proposed model gives precision of 0.97 for abuse and 0.95 for assault. The Five-minute-long films were trimmed down to forty-five seconds with the aid of video editing and cutting, with the emphasis being placed on the time of the actual incident rather than on unrelated or false information, i.e., the data used is in small amount. This is the novel approach used in this research.

References

[1] O. Abdel-Hamid, M. Abdel-Rahman, H. Jiang, & G. Penn, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing ICASSP'2012*, 2012, pp. 4277-4280, doi: 10.1109/ICASSP.2012.6288864.

[2] D. Shah, R. Dixit, A. Shah, P. Shah, M. Shah, "A comprehensive analysis regarding several breakthroughs based on computer intelligence targeting various syndromes," *Augment Hum Res*, vol. 5, issue 1, pp. 14, 2020. https://doi.org/10.1007/s41133-020-00033-z

[3] W. Sultani, C. Chen, M. Shah, "Real-world anomaly detection in surveillance videos," *Cornell University Library, arXiv:1801.04264*, 2018.

[4] M. Cheng, K. Cai and M. Li, "RWF-2000: An open large scale video database for violence detection," *Proceedings of the 2020 25th IEEE International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2021, pp. 4183-4190, doi: 10.1109/ICPR48806.2021.9412502.

[5] K. Ahir, K. Govani, R. Gajera, M. Shah, "Application on virtual reality for enhanced education learning, military training and sports," *Augment Hum Res*, vol. 5, issue 1, article no. 7, 2020. https://doi.org/10.1007/s41133-019-0025-2

[6] A. Bates, *Stingray: A New Frontier in Police Surveillance*, Cato Institute Policy Analysis, no. 809, 2017.

[7] C. Berghoff, M. Neu and Arndt von Twickel. (2021) "The Interplay of AI and Biometrics: Challenges and Opportunities." Computer 54 (2021): 80-85.

[8] A. L. Blum, R. L. Rivest, "Training a 3-node neural network is NP-complete," *Neural Netw.*, vol. 5, issue 1, pp. 117–127, 1992. https://doi.org/10.1016/S0893-6080(05)80010-3

[9] C. Berghoff, M. Neu and A. von Twickel, "The interplay of AI and biometrics: Challenges and opportunities," *Computer*, vol. 54, no. 9, pp. 80-85, 2021, doi: 10.1109/MC.2021.3084656.

[10] P. Chen, H. Y. Yuan, X. M. Shu, "Forecasting crime using the ARIMA model," *Proceedings of the 5th IEEE International Conference on Fuzzy Systems and Knowledge Discovery*, Ji'nan, 18-20 October 2008, pp. 627-630. https://doi.org/10.1109/FSKD.2008.222.

[11] A. Dey, "Machine learning algorithms: a review," *Int J Comput Sci Inf Technol*, vol. 7, issue 3, pp. 1174–1179, 2016.

[12] T. Fatih, C. Bekir, "Police use of technology to fight against crime," *Eur Sci J*, vol. 11, issue 10, pp. 286–296, 2015.

[13] M. Gandhi, J. Kamdar, M. Shah, "Preprocessing of non-symmetrical images for edge detection," *Augment Hum Res*, vol. 5, issue 1, 10, 2020. https://doi.org/10.1007/s41133-019-0030-5

[14] W. Gorr, R. Harries, "Introduction to crime forecasting," *Int J Forecast*, vol. 19, issue 4, pp. 551–555, 2003. https://doi.org/10.1016/S0169-2070(03)00089-X

[15] A. Gupta, V. Dengre, H. A. Kheruwala, M. Shah, "Comprehensive review of text-mining applications in finance," *Financ Innov*, vol. 6, issue 1, pp. 1–25, 2020. https://doi.org/10.1186/s40854-020-00205-1

[16] K. Jani, M. Chaudhuri, H. Patel, M. Shah, "Machine learning in films: an approach towards automation in film censoring," *J Data Inf Manag*, vol. 2, issue 1, pp. 55–64, 2020. https://doi.org/10.1007/s42488-019-00016-9.

[17] K. Jha, A. Doshi, P. Patel, M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artif Intell Agric*, no. 2, pp. 1–12, 2019. https://doi.org/10.1016/j.aiia.2019.05.004.

[18] E. E. Joh, "The undue influence of surveillance technology companies on policing," *N Y Univ Law Rev*, vol. 92, pp. 101–130, 2017. https://doi.org/10.2139/ssrn.2924620.

[19] S. Judd, "On the complexity of loading shallow neural networks," *J Complex*, vol. 4, issue 3, pp. 177–192, 1988. https://doi.org/10.1016/0885-064X(88)90019-2.

[20] V. Kakkad, M. Patel, M. Shah, "Biometric authentication and image encryption for image security in cloud framework," *Multiscale Multidiscip Model Exp Des*, vol. 2, issue 4, pp. 233–248, 2019. https://doi.org/10.1007/s41939-019-00049-y.

[21] C. M. Katz, D. E. Choate, J. R. Ready, L. Nuňo, "Evaluating the impact of officer worn body cameras in the Phoenix Police Department," *Center for Violence Prevention & Community Safety*, Arizona State University, Phoenix, 2014, pp 1–43.

[22] K. Kundalia, Y. Patel, M. Shah, "Multi-label movie genre detection from a movie poster using knowledge transfer learning," *Augment Hum Res*, vol. 5, issue 1, 11, 2020. https://doi.org/10.1007/s41133-019-0029-y.

[23] T. L. Le, M. Q. Nguyen, T. T. M. Nguyen, "Human posture recognition using human skeleton provided by Kinect," Proceedings of the 2013 IEEE International Conference on Computing, Management and Telecommunications, Ho Chi Minh City, 2013, pp. 340-345. https://doi.org/10.1109/ComManTel.2013.6482417.

[24] S. Marsland, *Machine Learning: an Algorithmic Perspective*, CRC Press, Boca Raton, 2015, pp. 1–452. https://doi.org/10.1201/b17476-1.

[25] L. McClendon, N. Meghanathan, "Using machine learning algorithms to analyze crime data," *Mach Lear Appl Int J*, vol. 2, issue 1, pp. 1–12, 2015. https://doi.org/10.5121/mlaij.2015.2101.

[26] G. S. McNeal, "Drones and aerial surveillance: Considerations for legislators," Brookings Institution: The Robots Are Coming: The Project on Civilian Robotics, November 2014, Pepperdine University Legal Studies Research Paper, no. 2015/3, https://www.brookings.edu/wp-content/uploads/2016/07/Drones_Aerial_Surveillance_McNeal_FINAL.pdf.

[27] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini et al, "An overview on application of machine learning techniques in optical networks," *IEEE Commun Surv Tutorials*, vol. 21, issue 2, pp. 1381–1408, 2019. https://doi.org/10.1109/COMST.2018.2880039.

[28] B. Naik, A. Mehta, M. Shah, "Denouements of machine learning and multimodal diagnostic classification of Alzheimer's disease," *Vis Comput Ind Biomed Art*, vol. 3, issue 1, 26, 2020. https://doi.org/10.1186/s42492-020-00062-w.

[29] S. Panchiwala, M. Shah, "A comprehensive study on critical security issues and challenges of the IoT world," *J Data Inf Manag*, vol. 2, issue 7, pp. 257–278, 2020. https://doi.org/10.1007/s42488-020-00030-2.

[30] R. Pandya, S. Nadiadwala, R. Shah, M. Shah, "Buildout of methodology for meticulous diagnosis of K-complex in EEG for aiding the detection of Alzheimer's by artificial intelligence," *Augment Hum Res*, vol. 5, issue 1, 3, 2020. https://doi.org/10.1007/s41133-019-0021-6.

[31] P. Parekh, S. Patel, N. Patel, M. Shah, "Systematic review and meta-analysis of augmented reality in medicine, retail, and games," *Vis Comput Ind Biomed Art*, vol. 3, issue 1, 21, 2020. https://doi.org/10.1186/s42492-020-00057-7.

[32] V. Parekh, D. Shah, M. Shah, "Fatigue detection using artificial intelligence framework," *Augment Hum Res*, vol. 5, issue 1, 5, 2020. https://doi.org/10.1007/s41133-019-0023-4.

[33] D. Patel, D. Shah, M. Shah, "The intertwine of brain and body: a quantitative analysis on how big data influences the system of sports," *Ann Data Sci*, vol. 7, issue 1, pp. 1–16, 2020. https://doi.org/10.1007/s40745-019-00239-y.

[34] D. Patel, Y. Shah, N. Thakkar, K. Shah, M. Shah, "Implementation of artificial intelligence techniques for cancer detection," *Augment Hum Res*, vol. 5, issue 1, 6, 2020. https://doi.org/10.1007/s41133-019-0024-3.

[35] H. Patel, D. Prajapati, D. Mahida, M. Shah, "Transforming petroleum downstream sector through big data: A holistic review," *J Pet Explor Prod Technol*, vol. 10, issue 6, pp. 2601–2611, 2020. https://doi.org/10.1007/s13202-020-00889-2.

[36] M. Pathan, N. Patel, Hıo Yagnik, M. Shah, "Artificial cognition for applications in smart agriculture: A comprehensive review," *Artif Intell Agric*, vol. 4, pp. 81–95, 2020. https://doi.org/10.1016/j.aiia.2020.06.001.

[37] A. Rani, S. Rajasree, "Crime trend analysis and prediction using mahanolobis distance and dynamic time warping technique," *Int J Comput Sci Inf Technol*, vol. 5, issue 3, pp. 4131–4135, 2014.

[38] A. Rummens, W. Hardyns, L. Pauwels, "The use of predictive analysis in spatiotemporal crime forecasting: Building and testing a model in an urban context," *Appl Geogr*, vol. 86, pp. 255–261, 2017. https://doi.org/10.1016/j.apgeog.2017.06.011.

[39] K. Shah, H. Patel, D. Sanghvi, M. Shah, "A comparative analysis of logistic regression, random forest and KNN models for the text classification," *Augment Hum Res*, vol. 5, issue 1, 12, 2020. https://doi.org/10.1007/s41133-020-00032-0.

[40] N. Shah, S. Engineer, N. Bhagat, H. Chauhan, M. Shah, "Research trends on the usage of machine learning and artificial intelligence in advertising," *Augment Hum Res*, vol. 5, issue 1, 19, 2020. https://doi.org/10.1007/s41133-020-00038-8.

[41] N. Shah, N. Bhagat, M. Shah, "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention," *Vis. Comput. Ind. Biomed. Art*, vol. 4, article no. 9, 2021. https://doi.org/10.1186/s42492-021-00075-z.

[42] A. Simon, M. S. Deo, S. Venkatesan, D. R. Babu, "An overview of machine learning and its applications," *Int J Electr Sci Eng*, vol. 1, issue 1, pp. 22–24, 2016.

[43] J. Stanley, "Police body-mounted cameras: with right policies in place, a win for all," 2015. [Online]. Available at: https://www.aclu.org/police-body-mounted-cameras-right-policies-place-win-all.

[44] A. Sukhadia, K. Upadhyay, M. Gundeti, S. Shah, M. Shah, "Optimization of smart traffic governance system using artificial intelligence," *Augment Hum Res*, vol. 5, issue 1, 13, 2020. https://doi.org/10.1007/s41133-020-00035-x.

[45] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer-Verlag, Berlin, 2010, pp. 979.

[46] T. Talaviya, D. Shah, N. Patel, H. Yagnik, M. Shah, "Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides," *Artif Intell Agric*, no. 4, pp. 58–73, 2020. https://doi.org/10.1016/j.aiia.2020.04.002.

[47] A. Vedaldi, B. Fulkerson, "Vlfeat: an open and portable library of computer vision algorithms," Proceedings of the 18th ACM International Conference on Multimedia, Firenze, 2010, pp. 1469–1472. https://doi.org/10.1145/1873951.1874249.

[48] A. Vredeveldt, L. Kesteloo, P. J. Van Koppen, "Writing alone or together: police officers' collaborative reports of an incident," *Crim Justice Behav*, vol. 45, issue 7, pp. 1071–1092, 2018. https://doi.org/10.1177/0093854818771721.

***RADDAM SAMI MEHSEN** Graduated Middle Technical University, Technical Institute of Baqubah. Received bachelor from al Mustansiriyah University, and Master degree of Computer science, 2011, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MH, India. Specialist in computer science, expert in many programming languages HTML, PHP, JavaScript, C ++, Visual Basic. Experience in project management and client relations. He has a good experience in designing programs and databases for people and institutions, correcting their mistakes, and making modification operations.*