Global Scientific JOURNALS

GSJ: Volume 11, Issue 12, December 2023, Online: ISSN 2320-9186

www.globalscientificjournal.com

SURVEILLANCE SYSTEM FOR DETECTION OBJECTS AND EMOTIONS OF PERSONS USING MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

BY

Aguwamba, Rowland Chidi
Department of Office Technology and Management
Akanu Ibiam Federal Polytechnic, Unwana, Ebonyi, Nigeria

Okoroudoh, Itoro Sunday
Information and Communication Technology Department
Akanu Ibiam Federal Polytechnic Unwana, Ebonyi State, Nigeria
&
Nwanhele, Udo Noble
Department of Office Technology and Management
Akanu Ibiam Federal Polytechnic Unwana, Ebonyi State, Nigeria

ABSTRACT

Surveillance against criminal attacks and recognition of objects/persons moving within a building has become a lot more challenging, especially for security personnels. It has become increasingly difficult to continuously monitor the condition of the building against the person coming in or going out. Thus, surveillance system for Detection Objects and Emotions of Persons Using Machine Learning and Artificial Intelligence is implemented to overcome the monitoring issues and provide intimation to house owner or security personnel in real-time. Through the use of TensorFlow, DeepFace and Threading python libraries being comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in Machine Learning (ML) and developers easily build and deploy Machine Learning (ML) powered applications. The libraries provide access to machine learning algorithms and models in the browser, building on top with no other external dependencies making the camara module to be triggered thereby taking the system to capture and analyze different video captures to detect human being thereby sounding an alarm and sending an sms alert through twilio function to a predefined agents and furthermore, making predictions of the emotion of the person capture on the camera.

INTRODUCTION

Digitization has always been the hub of focus in this 21st century. Though it cuts jobs but at the same time it provides the better and standard way of living in the society. In this dispensation, because of emerging trend of innovations, local businesses are folding and leaving the race for intelligent businesses. A business that consists of highly advanced systems for providing and taking

decisions for the organizations. Most houses and business centers spends money in order for some security men to move around the vicinity of the building to check and report the movements and intensions of people around the vicinity but as humans many of them will be tired along the line, they can sleep on duty which may be the very time the person with criminal intension comes, some security men may be distracted or send on different errand that takes them away from their core essence. In the light of this, there is need to bring about highly advanced system that can handle the stated challenges thereby proffering solutions to this menace. What would be the case if we would be able to implement a real-world application that can send signal directly through their expression to the house owner. What if we would be able to develop a system that can detect any anti-social activities before it happens through the emotion of the person. Various field of life make use of Human facial Expression recognition. Signal of person movement on camera and the emotion are being send through sms to the house owner with an alarm for people standing by. This work concentrates mostly to conversion of grayscale form matrix. The data array passed between layer of nodes is called tensor. Most used open-source libraries which come readymade with n numbers of algorithms, known as TensorFlow is used.

It is primordial, engaging many security personnels in the building to detect the movement of people and their motives. Infact, detecting their motives involve going close to them to engage them in questions which may bring insult, reprisal attack and lots more. In a situation of large building, involving many people may be highly cost effective and what if the security personnel dose off while on duty and the criminal enters at that time. However, external factors can easily affect these systems or methods because this system only follow preprogrammed algorithms which may varies based on object size, color or distance. Consequently, accurately determining objects movement on camera is somewhat challenging, depending on interval data response. Moreover, the lack of integration between automatic identification system and visual data yields non-intuitive insights. Thus, the exclusion of many individuals, especially those farther from the camera.

Kumar (2020) observed that the basic idea of system using TensorFlow is to be able to create dataflow graphs. These graphs have nodes and edges. It uses Python as front end and runs it efficiently in optimized C++.

TensorFlow, DeepFace and threading libraries has APIs available in several languages both for constructing and executing a graph. The Application Programming Interface (API) in Python is very enormous and efficient in performing task compared to other language APIs which may be easier to integrate but difficult in achieving great task, but python offer some performance advantages in graph execution. Among the open-source software library which TensorFlow one of it can be used for numerical computation using data flow graphs. mathematical operations are denoted by Nodes in the graphs, while the graph edges represent the multidimensional data arrays called tensors which communicated between them. This architecture being flexible allows you to deploy computation in numeral CPUs and GPUs such as in desktop, server, or mobile device with a single API.

STATEMENT OF THE PROBLEM/JUSTIFICATION

a) There is lack of understanding and interpretation when viewing large dataset, Professionals, Companies and individuals fail in their methods projects because of insufficient knowledge about data based on the information made available from research data.

b) Few critical problems in the surveillance space which include tracking of movement of people in a particular area, occupancy parking lots etc. Many companies have spent considerable effort in building a system that can tract movement of people with a locality or building

c) The big data age: The amount of data being saved in research centers, security central control systems like Network Video Recorder NVR which has imposed higher data traffic for consumption for end users.

d) The concern and confusion of selecting tools that supports multi-directional data storage or handle issues in Big data visualization.

OBJECTIVES OF THE STUDY

The followings are the objectives of the study; which are primarily to:

1. Build an efficient Model capable of detecting objects and making deep decisions.
2. Train and test the model in areas of surveillance using 10 epochs.
3. Implement the model using Tensorflow and DeepFace.
4. Comparing real time model object detections and human observation.

LITERATURE REVIEW

When interacting and socializing with other people, oftentimes emotions, such as anger, joy, surprise, fear, sadness, anxiety are often expressed. Tuomi (2018) opined that studying how to read them is useful for a variety of purposes, including investigations, interviews, and can be a tough task, so technology that deals with emotion recognition is used to do the job of detecting and recognizing different facial expressions using Facial Expression Analysis (Oh et al., 2018). This allows authorities to detect the emotions of a person. Face, object, and event recognition capabilities are enhanced by the use of AI in security and surveillance. TensorFlow is an open-source framework for deep learning dataflow and contains application programming interfaces (APIs) computer vision (Yoon, H., Lee, S. H., Park, M., 2020). It is AI or "Artificial Intelligence" that detects and studies different facial expressions to use them with additional information presented to them.

Due to its efficiency and accuracy, TensorFlow has emerged to be an effective tool in object detection and related tasks (Sanyal, 2022).

Emotion recognition can be realized through different detection methods and different sensors. (Malfaz & Salichs, 2004). Sensors are combined with advanced algorithm models and rich data to form human-computer interaction systems or robot systems (Ogata & Sugano 1999; Rattanyu, Ohkura & Mizukawa, 2010). In the field of medical and health care, Hasnul, Aziz, Alelyani, Mohana and Aziz (2021) noted that emotion recognition can be used to detect the patient's psychological state or adjuvant treatment, and improve medical efficiency and medical experience. In the field of Internet education, emotion recognition can be used to detect students' learning status and knowledge acceptance, and cooperate with relevant reminders to improve learning efficiency (Feidakis, Daradoumis & Caballé, 2011).. In the field of criminal interrogation, emotion recognition can be used to detect lies (authenticity test) (Saste & Jagdale, 2017). In the field of

intelligent cockpits, it can be used to detect the drowsiness and mental state of the driver to improve driving safety (Zepf, Hernandez, Schmitt, Minker & Picard, 2020).

Kumar (2020) observed that the basic idea of TensorFlow is to be able to create dataflow graphs. These graphs have nodes and edges. The array (data) passed along from layer of nodes to layer of nodes is known as tensor. TensorFlow Object Detection API is the tool that conducts efficient and swift object detection. It is an open-source framework that blends in machine learning and deep learning models and algorithms. It uses Python as a convenient front-end and runs it efficiently after optimizing it with C++. The platform allows developers to create a graph of computations to perform. There are several steps involved in detecting objects in real-time using this platform. But with the help of the OpenCV software, professionals can do it efficiently, instantly, and accurately. Currently, TensorFlow is a popular software library. There are several real-world applications of deep learning that make TensorFlow popular. Apart from that being an open-source library for deep learning and machine learning, the platform plays the role of a text-based application, image recognition, voice search, and many more. Recently, object detection has become a critical task for various reasons. The development of object detecting AI models in different commercially beneficial domains helped automate production lines, produce self-driving cars, assisting referees in sports, manufacturing lines, and other smaller projects. And deploying this technology has only emerge as a growing trend, with no signs of slowing down.

The definition of emotion is the basis of emotion recognition. The basic concept of emotion was proposed by Ekman (1970). At present, there are two mainstream emotion models: the Discrete emotion model and the dimensional emotion model. The sensors used for emotion recognition mainly include visual sensors, audio sensors, radar sensors, and other physiological signal sensors, which can collect signals of different dimensions and achieve emotional analysis through some algorithms. Different sensors have different applications in emotion recognition. Each of these sensors for emotion recognition has comparative advantages and disadvantages

Emotion recognition based on visual sensors is one of the most common emotion recognition methods. It has the advantages of low cost and simple data collection. At present, visual sensors are mainly used for facial expression recognition (FER) to detect emotion (Li & Deng, 2020; Schmid et al., 2011; Sandbach, Zafeiriou, Pantic & Yin, 2012) The accuracies of these methods severely drop as the light intensity decreases. Facial expressions can intuitively reflect the subjective emotions in interpersonal communication, but they are affected by limited lighting, occlusion, small changes in facial expressions, and individual differences. The performances of existing vision-based emotion recognition systems have significantly drop in environments with changing lighting conditions. Self-occlusion due to head rotation or face contact, and occlusion by other people passing in front of the camera, are both common problems. Moreover, individual differences can affect the feature extraction and learning of the model. There are large differences between infants and adults, males and females, and different groups, which makes it challenging to train a FER classifier with strong generalization performance.

METHODOLOGY

This work focuses on the use of Datasets from YOLO, COCO and post training of the model in security threat detections. A decision tree model is developed, trained to learn and present data in a normal tree like structures containing nodes and sub-nodes which has help visualize predictions from the algorithms with confidence score.

1. The system dataset: The proposed dataset was collected from different research libraries. The training and testing were preprocessed before training the model to certain confidence level for better decision making.

2. Data acquisition: Acquiring and merging of data from all the appropriate sources. The process of the data acquisition involves searching for the datasets that can be used to train the machine learning models.

3. Preprocessing: is one of the preliminary phases where data is transformed into computer understandable format. Data from outside world is often inconsistent and lacks some behavioural trend and may likely contain errors. This requires standard data processing to minimize errors since it also requires discrete values.

3.4 Training and testing of DT classifiers: The training and testing of DT model requiring training and testing dataset from YOLO and COCO SSD which contains several key point features of certain classified objects contained in the library.

Algorithm1: Support vector machine(SVM)
Step 1: Start
Step 2: Find candidate_SV with closest pair from classification (SV=>support vector)
Step 3: If there are violating points:
(a) Find violating_points
(b**)** Compute the candidate_SV= candidate_SV + voilating_points)
(c) If there is any $\alpha_p < 0$ due to the addition of c to S that gives negative:
(i) Candidate_SV = candidate_SV
(ii) Repeat module to prune all data points
(d) end_if
Step4: end_if
Step5: Stop
Algorithm2: Decision Tree(DT)
Step 1: Start (Form_DcisionTree)
Step 2: Compute Class Frequency_Value(CFT)
Return a leaf_Node;
Step 3: Create a decision tree of N nodes;
Step 4: For Each Attributes of A
(a) ComputeGainValue(A)
Step 5: N(test) = BestattributeGain
Step 6: If N(test) is continuous
(a) Find ThresholdValue;
Step 7: For Each CFT' in the splitting of CFT
(a) if CFT is Empty:
(i) Child of N_{Node} is a leaf Node
(b) else:
(ii) Child of N_{Node} = FormDecisionTree(DT')
Step 8: ComputeError of N_{Node};
Step 9: return N_{Node}

3.5 System Requirements
3.5 System Requirements

The system requirements are the hardware and software need for implementing the proposed system.

a) Hardware requirement:

1. Processor: at least Core (TM)2 Duo, 2.00 GHZ processor recommended.
2. Hard disk: at least 520GB of available space required on system drive, 25.3 GB of available space required on installations.
3. Display super VGA (1024X768) or higher resolution display with 256 colours.
4. RAM: Minimum requirement of 4GB for best performance.
5. CD-WR Drive and USB port enabled.
6. Webcam for visual input

b) Software requirement:

1. Operating system: Linux, Microsoft windows 7 and above
2. Chromium browser for web view
3. Electron js for compiling to executable file exe, dmg
4. Pycharm IDE
5. P5 js for visualizing training process and confidence.
6. Tensorflow library, DeepFace Library using ML5.

IMPLEMENTATION

Figure 1: facial emotion expression and recognition system.

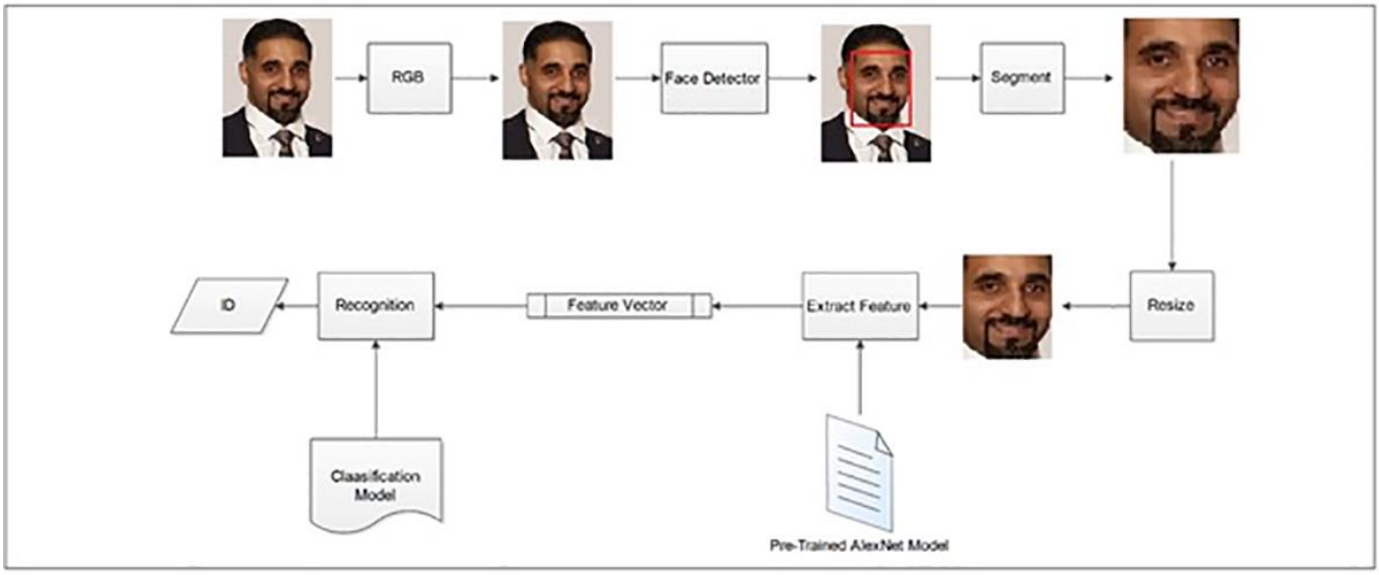

Figure 2: Block diagram of the proposed biometric system

Figure 3: Face images before and after preprocessing

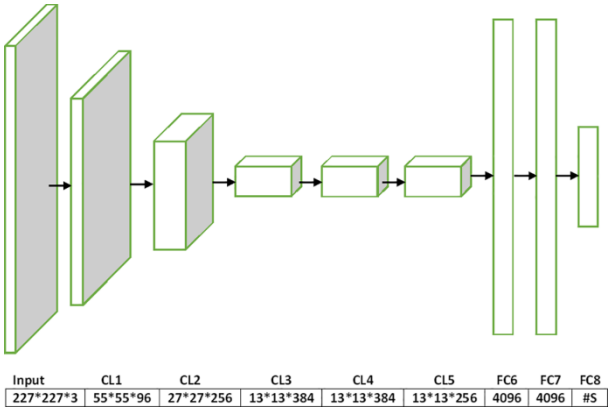

Figure 4: The schema of the modified TensorFlow, where (#S) is the number of subjects in the dataset used during training.

RESULTS

At the end of the project the following results were obtained as shown in the images provided

- The system was able to capture, and interpret video feeds.
- Learning of objects around the camera view during training sessions.
- Real time visual tags of objects in the video feed.

RECOMMENDATIONS

a.) A model developed will be capable of giving accurate information on surveillance, detecting objects and emotions of persons to track unacceptable occurrences and maintain a record of evidence in the event of a crime.
b.) The model will be capable to eliminate hidden patterns and possible errors, filter noise and make up missing values in order to have relevant information from captured scenes.
c.) Benefit of supporting and handling large dataset that deals with multi-dimension image display obstacles i.e. Staff, Students and other visitors to the Institution.
d.) User speed and accuracy to act on visual findings for better decision making.

REFERENCES

1. Kumar, S. (2020). Human Facial Expression Recognition Using TensorFlow And OpenCV https://doi.org/10.13140/RG.2.2.19218.89288
2. Tuomi, I., (2018). The Impact of Artificial Intelligence on Learning, Teaching, and Education. https://doi.org/10.2760/12297
3. Oh Y. H., See J., Le Ngo A. C., Phan R. C. W. & Baskaran V. M. (2018). A survey of automatic facial micro-expression analysis: databases, methods, and challenges. Front. Psychol. https://doi.org/10.3389/fpsyg.2018.01128
4. Yoon, H., Lee, S. H. & Park, M., (2020). TensorFlow with user friendly Graphical Framework for object detection API.
5. Sanyal, S. (2022). Object Detection by TensorFlow: An Emerging Transformative Trend. Retrieved from https://www.analyticsinsight.net/object-detection-by-tensorflow-an-emerging-transformative-trend/
6. Malfaz, M. & Salichs, M.A. (2004, July). A new architecture for autonomous robots based on emotions. The Fifth IFAC Symposium on Intelligent Autonomous Vehicles. Lisbon, Portugal.
7. Ogata T. & Sugano S. (1999). Emotional communication between humans and the autonomous robot which has the emotion model; Proceedings of the 1999 IEEE International Conference on Robotics and Automation. Detroit, MI, USA. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10007272/ pp. 3177–3182.
8. Rattanyu K., Ohkura M. & Mizukawa M. (2010). Emotion monitoring from physiological signals for service robots in the living space. Gyeonggi-do, Republic of Korea. 27–30. pp. 580–583.
9. Hasnul M.A., Aziz N.A.A., Alelyani S., Mohana M., Aziz & A.A.J.S. (2021). Electrocardiogram-based emotion recognition systems and their applications in healthcare—A review. *Sensors*. https://doi.org/10.3390/s21155015.
10. Feidakis M., Daradoumis T., & Caballé S.(2011). Emotion measurement in intelligent tutoring systems: What, when and how to measure.; Proceedings of the 2011 Third International Conference on Intelligent Networking and Collaborative Systems; Fukuoka, Japan. pp. 807–812.
11. Saste S.T. & Jagdale S.(2017). Emotion recognition from speech using MFCC and DWT for security system; Proceedings of the 2017 international Conference of Electronics, Communication and Aerospace Technology (ICECA); Coimbatore, India. ; pp. 701–704.
12. Zepf S., Hernandez J., Schmitt A., Minker= W. & Picard R. (2020). Driver emotion recognition for intelligent vehicles: A survey. https://doi.org/10.1145/3388790.
13. Li S. & Deng W. (2020). Deep facial expression recognition: A survey. https://do.org/ 10.1109/TAFFC.2020.2981446.
14. Sandbach G., Zafeiriou S., Pantic M., & Yin L.J.I. (2012). Computing V. Static and dynamic 3D facial expression recognition: A comprehensive survey. 683–697. https://doi.org/10.1016/j.imavis.2012.06.005.
15. Schmid P.C., Mast M.S., Bombari D., Mast F.W., Lobmaier J. (2011). How mood states affect information processing during facial emotion recognition: An eye tracking study. Swiss J. Psychol. https://doi.org/ 10.1024/1421-0185/a000060.g fbm