Real-Time Detection of Aggressive Behaviors with Alert System

DELCO, Kuh-kai A.
Electronics Engineering Department
Technological University of the Philippines

FERIDO, Angelo
Electronics Engineering Department
Technological University of the Philippines

FRANCO, Kelyn
Electronics Engineering Department
Technological University of the Philippines

MANADONG, Jovert
Electronics Engineering Department
Technological University of the Philippines

SANTIAGO, Brian Jamille
Electronics Engineering Department
Technological University of the Philippines

SALVADOR, Jean Prynce
Electronics Engineering Department
Technological University of the Philippines

ENGR. Mark Melegrito, PECE
Faculty
Electronics Engineering Department
Technological University of the Philippines

I. INTRODUCTION

Injuries contribute to 10% of global mortality and 15% of disability, yet data on injuries in developing countries, where two-thirds of injury deaths occur, are scarce. This report is the first to examine the issue of injuries in the Philippines, a developing country in Southeast Asia. It aims to define the burden of injuries and highlight priority areas for national health research. A review of 35 years of data (1960-1995) reveals a significant increase in injury fatality rates, rising by 196% from 14.3 per 100,000 in 1960 to 42.3 per 100,000 in 1995. One in 11 deaths in the Philippines is due to injuries, with intentional injuries accounting for 48% of all injury deaths. The report emphasizes the need for improved injury surveillance and documentation of non-fatal injury outcomes. It also calls for research into risk factors and interventions to prevent intentional injuries, highlighting the importance of addressing this issue in the Philippines (Consunji & Hyder, 2004).

Video surveillance systems are vital for public security, detecting suspicious behaviors, analyzing crowd dynamics, managing traffic, and tracking vehicles. Manual monitoring of these activities is challenging, leading to research efforts to automate information extraction using machine learning and deep learning. However, the lack of large, annotated video datasets poses a common barrier in this domain. Existing videos are often untrimmed, unannotated, and may contain ambiguous data. This research aims to address these limitations by proposing a machine learning-based transfer learning approach. The goal is to accurately detect violent crowd behavior through surveillance systems (Liyanage & Fernando, 2021).

II. BACKGROUND OF THE STUDY

In surveillance, detecting human abnormal behavior is crucial for public safety, but it's challenging due to lengthy video datasets and occasional occurrences of abnormal activities. This often requires extensive manpower to verify video streams, making the process expensive, inefficient, and time-consuming (Kim et. al., 2021).

The aim of surveillance is to prevent unwanted incidents and respond to them as needed, depending on the situation. An AI-based automated system is necessary to process videos and enhance societal well-being. Recognizing human abnormal behavior through automated surveillance can improve safety for the elderly and patients, reduce criminal activity and theft, decrease workplace harassment and violence (Kim et. al., 2021).

III. STATEMENT OF THE PROBLEM

Maintaining a safe and secure environment in barangays requires vigilance. Aggressive behaviors, from verbal arguments to physical threats, can create fear and disrupt the peace within a community. Traditional methods for addressing such behaviors often rely on resident reports, or security personnel intervention. These methods are reactive that leaves residents feeling vulnerable and may not be sufficient to prevent violence or escalation of tensions. TANAW can identify threatening situations, its real-time detection capability relies on machine learning algorithms that analyzes video feeds from strategically placed cameras. By identifying aggressive behaviors like kicking, strangling, punching, and hair pulling in real-time, TANAW empowers authorities to intervene before situations escalate further.

Early detection is crucial in preventing violence and maintaining a peaceful barangay. Upon detecting aggressive behavior, the alert system of TANAW will be triggered notifying the barangay. This allows for a quick and targeted response before they escalate further. This system and approach to the community promotes a more secure environment for the barangay.

IV. OBJECTIVES

This research aimed to create a Real-Time Detection of Aggressive Behaviors with Alert System

1. To develop a real-time monitoring and detection system for emergency hand gesture recognition with alert system.
2. Implement and improve the real-time detection capability of the system to identify aggressive behaviors and enable authorities for timely intervention in barangays.

V. RELATED STUDIES

According to the study of Hassner & Kliper-Gross (2012), they emphasize the critical need for real-time detection of violent outbreaks in crowded events using surveillance cameras. It introduces a novel approach called ViF (Violent Flows) for efficient crowd violence detection, which outperforms existing techniques by focusing on the magnitudes of optical-flow fields. The proposed method considers how flow-vector magnitudes change over time and uses this information to classify scenes as either violent or non-violent using linear SVM. Additionally, the study provides a unique dataset of real-world surveillance videos and benchmarks to evaluate violent/non-violent classification and real-time detection accuracy. By comparing their method to state-of-the-art techniques, the study demonstrates the effectiveness of their approach in detecting aggressive behaviors in crowded environments, aligning with efforts in real-time detection of aggressive behaviors with alert systems.

The study of Sumon, et al. (2019) focuses on detecting violent crowd flows using deep learning algorithms applied to a small dataset of violent and non-violent videos. The researchers found that a convolutional neural network (CNN) using transfer learning performed better than other CNN and long short-term memory network (LSTM) models. Combining CNN with LSTM improved accuracy but still didn't surpass the transfer learning model. In future studies, they plan to make the model more lightweight through pruning and deploy it on an unmanned aerial vehicle (UAV) for real-time monitoring. They also aim to create an API for easy access to the model via a web server. This research contributes to real-time detection of aggressive behaviors, aligning with efforts in Real-Time Detection of Aggressive Behaviors with Alert Systems.

VI. METHODOLOGY

This section outlines for developing a real-time detection of aggressive behaviors. The methodology includes Dataset collection, Image Annotation and Augmentation, Model Training and Evaluation, and Deployment. The other procedures expound about public dataset gaining, custom dataset creation, image annotation using Roboflow, model training and configuration, evaluation with false positive/negative analysis, image augmentation, addressing class imbalance and hyperparameter tuning, different training epochs, and deployment for real-time video processing.

A. Dataset Collection

Researchers searched for publicly available video datasets containing examples of aggressive behavior, such as hair pulling, kicking, punching, and strangling. It was critical to find datasets with a balanced representation of aggressive and non-aggressive interactions to avoid bias in the model. This balanced representation would help the model distinguish between the two types of behavior more effectively. Additionally, the search prioritized datasets captured by CCTV cameras. This ensured that the model was trained on data closely resembling the real-world scenarios it would be deployed, considering factors like video quality and recording angles. Using CCTV footage, the model could adapt to the specific characteristics of these surveillance systems, improving its effectiveness.

B. Data Processing

The videos were split into individual frames for frame-by-frame analysis of the behavior. These extracted frames were then resized to a uniform size for efficient processing by the model. Finally, the preprocessed frames (RGB data) were stored in a TFRecord format to train the model. TFRecord is a format optimized for storing large image datasets used in TensorFlow.

C. Model Training and Evaluation:

The system relies on TensorFlow to analyze individual frames extracted from video footage. This program comprises three essential parts working together: The first focuses on reducing the amount of information within each video frame, concentrating on the most critical details relevant to identifying aggressive behavior. The second part helps the program become more robust against differences in video quality often encountered in CCTV footage. This is achieved by introducing slight artificial changes to the training data, such as adjusting lighting or camera angles. The final part uses specialized filters and analysis techniques to pinpoint frames containing aggressive behavior within the video footage. Researchers collected videos showcasing aggressive behavior to train this program effectively and then separated them into individual frames. The program fed these frames into three groups: training, testing, and validation. The training allows the program to learn and identify patterns associated with aggressive behavior. The testing evaluates the program's accuracy in real-world scenarios. Finally, the validation plays a crucial role in preventing the program from becoming overly specific to the training data, ensuring it generalizes well to real-world CCTV footage.

D. Deployment

The trained LSTM model will be deployed for real-time analysis of network traffic data streams. Incoming network traffic will be continuously preprocessed and fed into the model to predict potential anomalies. The system will be configured to trigger alerts based on detected anomalies, allowing network security personnel to investigate and take appropriate actions. The deployed model's performance will be monitored continuously to ensure its effectiveness. Retraining or fine-tuning may be necessary as network traffic patterns evolve.

V. REFERENCES

Consunji, R., & Hyder, A. A. (2004). The burden of injuries in the Philippines: implications for national research policy. Accident Analysis & Prevention, 36(6), 1111–1117. https://doi.org/10.1016/j.aap.2004.05.002

Liyanage P., Fernando, P. (2021). Suspicious Human Crowd Behaviour Detection – a transfer learning approach. (2021, December 2). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/9774784

Kim, D., Kim, H., Mok, Y., & Paik, J. (2021). Real-Time surveillance system for analyzing abnormal behavior of pedestrians. Applied Sciences, 11(13), 6153. https://doi.org/10.3390/app11136153

Hassner, T., Itcher, Y., & Kliper-Gross, O. (2012). Violent flows: Real-time detection of violent crowd behavior. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. doi:10.1109/cvprw.2012.6239348

Sumon, S. A., Shahria, M. T., Goni, M. R., Hasan, N., Almarufuzzaman, A. M., & Rahman, R. M. (2019). Violent Crowd Flow Detection Using Deep Learning. Lecture Notes in Computer Science, 613–625. doi:10.1007/978-3-030 14799-0_53

Real-time Detection of Deadly Weapons and Enhancing Public Safety using YOLOv5 with Alert System

DELCO, Kuh-kai A.
Electronics Engineering Department
Technological University of the Philippines

FERIDO, Angelo
Electronics Engineering Department
Technological University of the Philippines

FRANCO, Kelyn
Electronics Engineering Department
Technological University of the Philippines

MANADONG, Jovert
Electronics Engineering Department
Technological University of the Philippines

SANTIAGO, Brian Jamille
Electronics Engineering Department
Technological University of the Philippines

SALVADOR, Jean Prynce
Electronics Engineering Department
Technological University of the Philippines

ENGR. Mark Melegrito, PECE
Faculty
Electronics Engineering Department
Technological University of the Philippines

I. INTRODUCTION

In 2016, gun-related incidents globally surpassed 250,000, with over 1,013,000,000 firearms, mainly in civilian hands (N. Kurek, L. A. Darzi and J. Maa, 2020).. Real-time processing and fast response times are crucial for weapon detection systems, necessitating research to enhance processing speed. Challenges like similar shapes of non-weapon objects lead to high false negatives and false positives. Detecting weapons accurately while avoiding false alarms is critical, especially in videos where existing models struggle. To address these issues, this study proposes a model leveraging fast detection speeds of modern techniques like YOLOv5. The key contribution lies in reducing false positives and negatives in weapon detection, expanding the range to include rifles (R. Olmos, S. Tabik and F. Herrera.,2018).

Traditional weapon detection systems using sensors are expensive, insecure, and inefficient, often unable to cover large surveillance areas. To address these limitations, machine learning algorithms, particularly the YOLO (You Only Look Once) algorithm, have been deployed. Unlike sensors, YOLO analyzes entire input images, simplifying the detection process. It highlights regions of interest with bounding boxes, enhancing real-time detection accuracy. YOLO, especially the YOLOv5 algorithm, enables easy detection of illegal activities, particularly in crowded areas. It identifies both high and low-level objects like weapons and unusual items, improving security measures significantly (Mukto, M., et al, 2024).

II. BACKGROUND OF THE STUDY

Crime is difficult to predict in advance, but signs of potential criminal activity can be observed by monitoring suspect behavior and surroundings. Surveillance cameras are commonly used to monitor individual activities for any abnormal or criminal incidents in an area. These cameras transmit captured videos over IP networks for security personnel to view. However, continuous, and careful monitoring by security personnel is necessary to identify unusual incidents in a timely manner. This monitoring can be laborious and have negative effects on mental health. Therefore, there is a need for automated systems that utilize machine learning and artificial intelligence to monitor human actions and predict the occurrence of crime incidents before they happen (Motiian et al., 2017).

Gun and knife-related incidents are rising due to insufficient security checks. While CCTV cameras are being installed in many places, constant surveillance is challenging for humans. There is a pressing need for automated weapon detection to curb these incidents. The proposed solution aims to develop an automated system using YOLOv5 deep learning model to detect firearms and knives. This system is designed to enhance security measures and prevent potential threats more effectively [2].

III. STATEMENT OF THE PROBLEM

Public safety is a concern from schools, workplaces, and public spaces. The presence of deadly can turn a situation dangerous putting lives at risk. Traditional methods of weapon detection often rely on physical checks or surveillance footage review. These methods are important, but it has limitations, Physical check can be time consuming, while footage review can lead to delays in identifying threats. This is concerning in crowded environments where immediate action is important. Unlike traditional methods, TANAW incorporates a weapon detection model built upon YOLOv5. This model allows for real-time identification of a variety of deadly weapons, including firearms and knives, this eliminates the delays associated with physical check or video review. This allows for a faster response to potential threat, saving lives, and preventing dangerous situations from escalating.

The system integrates an automated alert system that triggers instantly upon identifying weapon. This notification system directly informs the barangay to allow security officers to intervene quickly.

IV. OBJECTIVES

This research aimed to create a Detection of Deadly Weapons and Enhancing Public Safety using YOLOv5 with Alert System

1. Develop and Implement a Real-Time Weapon Detection System Based on YOLOv5.
2. Enhance Public Safety and Security Measures Through Automated Weapon Detection and Immediate Alert Systems.

V. RELATED STUDIES

Researchers (Grega et al., 2016) built algorithms to automatically detect dangerous situations in CCTV footage, like spotting knives and guns held in hands. Their knife detection was especially good at finding these weapons (high sensitivity) while rarely mistaking other objects for them (high specificity). It even worked well on grainy videos. While the gun detection was very good at not raising false alarms (high specificity) for non-dangerous situations, it sometimes missed actual guns (lower sensitivity) in videos with dangerous objects. They're working on improving these algorithms to work better in different situations, combine them, and reduce false alarms. They also want to add new ways to identify dangerous tools in difficult conditions, ultimately creating a complete system for CCTV operators.

According to Olmos et al. (2018), an innovative automatic handgun detection system designed for surveillance and control purposes addresses the issue of false positives by leveraging deep Convolutional Neural Networks (CNNs). By refining the detection problem and training data based on CNN classifier results, the researchers evaluated different classification models using sliding window and region proposal approaches. The most promising results were achieved with a Faster R-CNN-based model trained on their new database, showing potential even in low-quality videos like those on YouTube and functioning effectively as an automatic alarm system. The study also introduced a new metric, Alarm Activation Time per Interval (AATpI), to evaluate detection model performance in real-time video surveillance scenarios.

Dwivedi et al. (2019) focuses on weapon detection and classification using video surveillance to enhance public safety, particularly in places like schools and malls. The researchers propose two new deep CNN architectures based on the VGG16 model, fine-tuned by training them with images of knives, guns, and no-weapons classes. They achieved a maximum accuracy of 98.41% with Model A using a 0.5 dropout rate, highlighting the effectiveness of their approach. Interestingly, they found that simply increasing the number of neurons doesn't always improve accuracy; Model A with fewer neurons outperformed Model B with a higher dropout rate in fully connected layers. This study underscores the importance of optimizing model architecture and dropout rates for accurate weapon detection in real-time scenarios,

VI. METHODOLOGY

This section outlines for developing a real-time gun and knife detection. The methodology includes Dataset collection, Image Annotation and Augmentation, YOLOv5 Model Training and Evaluation, and Deployment. The other procedures expound about public dataset gaining, custom dataset creation, image annotation using Roboflow, YOLOv5 model training and configuration, evaluation with false positive/negative analysis, image augmentation, addressing class imbalance and hyperparameter tuning, different training epochs, and deployment for real-time video processing.

A. Dataset Collection

The researchers searched for publicly available image sets containing guns and knives in various settings, including viewpoints, lighting, and backgrounds. This ensured the system could adapt to real-world scenarios under different conditions. A custom dataset with images and videos specific to the barangay was created to further enhance the model and capture local weapons. Maintaining a balanced representation of different weapon types within the dataset was essential.

B. Image Annotation and Augmentation

Researchers used Roboflow to label each weapon instance in the images with bounding boxes. This process was like circling the weapons in a photo lineup. The accuracy of these labels was essential for training the program to recognize weapons effectively.

C. YOLOv5 Model Training and Evaluation:

The researchers used YOLOv5 to identify guns and knives in videos. The videos combined the public and custom datasets collected. Then, the combined data was split into training, validation, and testing. During training, YOLOv5 was taught to recognize weapons using training data. Based on this data, an optimization algorithm was employed to improve the model's internal settings. Validation data was essential to improve the training process and prevent overfitting by monitoring the model's performance and adjusting the learning rate and training epochs (100, 200, and 300). Testing data evaluated YOLOv5's accuracy in detecting weapons and avoiding errors like mistaking harmless objects for real weapons. Metrics such as precision, recall, F1-score, and false positives/negatives were used for evaluation.

D. Deployment

The final stage included deploying the trained program for real-time usage through a structured video processing pipeline. This involved initializing the YOLOv5 model, capturing video, and setting necessary parameters. Each video frame was read, resized, and preprocessed for efficient processing, followed by predicting weapons within a specific area of interest (ROI) in the frame, focusing on areas with a higher likelihood of weapons presence, such as building entrances. Detected weapons were annotated on the frame for clear visualization, and the processed video stream with weapon annotations overlaid was displayed, allowing real-time monitoring. Deployment is required to ensure the hardware can handle real-time video processing with YOLOv5 for efficient operation.

V. REFERENCES

N. Kurek, L. A. Darzi and J. Maa. (2020). “A Worldwide perspective provides insights into why a US surgeon general annual report on firearm injuries is needed in America,” Current Trauma Reports, vol. 6, pp. 36–43.

R. Olmos, S. Tabik and F. Herrera. (2018). “Automatic handgun detection alarm in videos using deeplearning,” Neurocomputing, vol. 275, no. 9, pp. 66–72.

Mukto, M. M., Hasan, M., Mahmud, M. M. A., Haque, I., Ahmed, M. A., Jabid, T., Ali, M. S., Rashid, M. R. A., Islam, M. M., & Islam, M. I. (2024). Design of a Real-Time crime monitoring system using deep learning techniques. Intelligent Systems With Applications, 21, 200311.https://doi.org/10.1016/j.iswa.2023. 00311

Real-Time Detection of Knives and Firearms using Deep Learning. (2022, October 21). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/abstract/documen /997295

Grega, M., Matiolański, A., Guzik, P., & Leszczuk, M. (2016). Automated detection of firearms and knives in a CCTV image. Sensors, 16(1), 47. https://doi.org/10.3390/s16010047

Olmos, R., Tabik, S., & Herrera, F. (2018). Automatic handgun detection alarm in videos using deep learning. Neurocomputing (Amsterdam), 275, 66–72. https://doi.org/10.1016/j.neucom.2017.05.02

Dwivedi, N., Singh, D. K., & Kushwaha, D. S. (2019). Weapon Classification using Deep Convolutional Neural Network. 2019 IEEE Conference on Information and Communication Technology. doi:10.1109/cict48419.2019.90662

Dwivedi, N., Singh, D. K., & Kushwaha, D. S. (2019). Weapon Classification using Deep Convolutional Neural Network. 2019 IEEE Conference on Information and Communication Technology. doi:10.1109/cict48419.2019.906622