ISSN 2277-2685
IJESR/June-2024/ Vol-14/Issue-2s/403—415
Dr. B. Rama *et. al.,* / International Journal of Engineering & Science Research

IMPROVING WEAPON RECOGNITION FROM IMAGES: APPLICATION OF DEEP LEARNING TECHNIQUE INNOVATION

Dr. B. Rama[1], Mohit Krishna[2], K. Sai chandra[2], N.Abhinash Reddy[2]
[1]Professor, [2]UG Scholar, [1,2]Department of CSE (AI&ML)
[1,2]Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana

ABSTRACT
The recognition of weapons from images plays a pivotal role in ensuring public safety and national security. In the current landscape of increasing firearm-related incidents, the need for accurate and efficient weapon detection systems is more crucial than ever. Conventional weapon recognition systems often rely on handcrafted features and traditional computer vision methods, which are limited in their ability to adapt to the diverse range of weapons and environmental conditions. These drawbacks result in reduced accuracy and the potential for false negatives or positives, endangering lives. In response to these challenges, our proposed system leverages state-of-the-art deep learning algorithms to automatically learn discriminative features from weapon images. The proposed system's versatility is demonstrated through its ability to detect various types of weapons, including handguns, rifles, and knives, while also adapting to different lighting conditions and backgrounds. This research represents a significant step towards leveraging machine learning to bolster public safety and security measures, ultimately reducing the potential risks associated with weapon-related incidents.

Keywords: Weapon Detection, Conventional Weapon Recognition, Deep Learning Algorithms, Machine Learning

1. INRODUCTION

The recognition of weapons from images is an essential aspect of modern security and law enforcement efforts. With the proliferation of firearm-related incidents worldwide, the need for accurate and efficient weapon detection systems has become increasingly apparent. Traditional methods often rely on manually crafted features and conventional computer vision techniques, which may struggle to adapt to the diverse array of weapons and environmental conditions encountered in real-world scenarios. As a result, these approaches may suffer from diminished accuracy and are susceptible to false positives or negatives, thereby compromising public safety and national security. In response to these challenges, this research aims to develop a robust weapon recognition system that harnesses the power of deep learning algorithms. By leveraging state-of-the-art Convolutional Neural Networks (CNNs), the proposed system seeks to automatically learn discriminative features directly from weapon images, thus circumventing the limitations of handcrafted feature extraction. Moreover, the system's adaptability will be demonstrated through its ability to detect various types of weapons, including handguns, rifles, and knives, across different lighting conditions and backgrounds.

2. LITERATURE SURVEY

Santos et.al [1] proposed that the number of crimes with weapons grew on a large scale worldwide, mainly in locations where enforcement was lacking or possessing weapons was legal. It was necessary to combat this type of criminal activity to identify criminal behavior early and allow police and law enforcement agencies immediate action. Despite the human visual structure being highly evolved and able to process images quickly and accurately, if an individual watched something very similar for a long time, there was a possibility of slowness and lack of attention. In addition, large surveillance systems with numerous equipment required a surveillance team, which increased the cost of operation. There were several solutions for automatic weapon detection based on computer vision; however, these had limited performance in challenging contexts. The use of realistic images and synthetic data showed improved performance.

Mukto et.al [2] proposed that criminal activities and Crime Monitoring System had long been a research topic. In that paper, an effective Crime Monitoring System (CMS) was proposed to detect a crime in real-time using a camera surveillance system. The CMS was designed to counterbalance human weaknesses such as inattention, slow reaction, and slacking, for example, in detecting crimes. The proposed CMS detected crime scenes by combining the mechanisms and functionalities of closed-circuit television (CCTV) cameras with various deep-learning methods and image-processing techniques. The CMS operated in three stages: weapon detection, violence detection, and face recognition. They used models to detect weapons, violence, and used face recognition algorithms to recognize faces.

Manikandan et.al [3] proposed that security was one of the important factors and a challenging task to tackle in a crowded place. The crucial thing to address in the crowded place was to identify any abnormal activity. Even though security surveillance was kept in many public places, it failed to identify the abnormal activities problem and to intimate the security admin in advance. Many algorithms and techniques were proposed to tackle that problem, but the environment, time, and place where the crowd was surrounded did not make it possible to obtain any feasible solution. To overcome that problem, a novel method was introduced by using Deep Learning where the input received from the surveillance camera was partitioned into multiple frames. Each of the frames was categorized into normal and abnormal frames. If any abnormal activity was identified, an intimation was generated to the security admin.

Gu et.al [4] proposed that they developed a complete robotic system to enable the robot to fulfill object-independent cooperative grasp tasks. The proposed grasp detection model was evaluated on two public grasping datasets and a set of casual objects. The best model variant could achieve an accuracy of 97.8% and 96.6% on image-wise splitting and object-wise splitting tests on the Cornell Grasp Dataset, respectively. The Jacquard Dataset accuracy was 93.9%.

Tian et.al [5] proposed that the deep neural network algorithm at that time still stayed in the end-to-end training supervision method like Image-Label pairs, making it difficult for traditional algorithms to explain the reason for the results, and the prediction logic was difficult to understand and analyze. That paper proposed a multi-level hierarchical deep learning algorithm, composed of a multi-level hierarchical deep neural network architecture and a multi-level hierarchical deep learning framework. The experimental results showed that the proposed algorithm could effectively explain the hidden information of the neural network.

Mwaffo et.al [6] proposed that the US Navy explored replacing manned systems in the carrier air wing with uncrewed platforms to increase flexibility, range, and endurance. However, autonomous air refueling lacked safety certification. That paper built on previous research, funded by the Office of Naval Research, using Deep Neural Network (DNN) models to detect the refueling drogue in various conditions. That study presented a

Deep Neural Network (DNN) trained with a limited set of real flight test data to precisely identify and track sections of the refueling drogue. That DNN went beyond simple object detection, utilizing instance segmentation to dynamically trace the drogue's shape.

Kuma et.al [7] Violent attacks were one of the hot issues in recent years. In the presence of closed-circuit televisions (CCTVs) in smart cities, there was an emerging challenge in apprehending criminals, leading to a need for innovative solutions. In this paper, a model was proposed aimed at enhancing real-time emergency response capabilities and swiftly identifying criminals. This initiative aimed to foster a safer environment and better manage criminal activity within smart cities. That proposed architecture combined an image-to-image stable diffusion model with violence detection and pose estimation approaches. The diffusion model generated synthetic data while the object detection approach used YOLO v7 to identify violent objects like baseball bats, knives, and pistols, complemented by Media Pipe for action detection. Thus, that study could handle violent attacks and send alerts in emergencies.

Xu et.al[8] proposed the convolutional neural network to detect key points of swimming posture. In swimming posture analysis, key point detection could provide information about posture accuracy and evaluation. However, due to changes in illumination conditions, reductions in image quality, and occlusion, the accuracy and stability of key point detection were limited. In order to solve that problem, an optical image enhancement method based on a convolutional neural network was proposed in that paper. First, a specific convolutional neural network architecture was used to extract features from input images. Then, optical image enhancement technology was used to preprocess the input image to improve the image quality. Next, the trained convolutional neural network model was used to detect key points.

Liu et.al [9] proposed that for the low probability of intercept (LPI) radar signal recognition problem, recognition algorithms based on deep learning usually used time-frequency analysis to convert the signal into a two-dimensional feature image for classification and recognition. However, these methods often had problems with large network size, high computational complexity, and huge memory consumption, making them difficult to apply on small devices with limited computational power and storage space. That paper proposed an LPI radar signal recognition method based on a lightweight complex convolutional neural network, named CV-LPINet. It used a complex convolutional module to complete data fusion of IQ sampled signals, used a deep separable convolutional module to extract features and reduce dimensions, and introduced residual structures to improve network training. Experiments showed that the average recognition accuracy of that method was 91.76% with a signal-to-noise ratio in the range of −6 to 10 dB, and its recognition accuracy was similar to that of typical algorithms

Ogundunmade et.al [10] proposed that terrorist attacks were one of the major problems Nigeria was facing presently. Seeking techniques to understand the different factors involved in terrorism and how to deal with those factors in order to completely eradicate or reduce terrorist activities was the topmost priority of the government in any country. That study used a Bayesian Neural Network (BNN) model for predicting the nature of attacks of terrorists in Nigeria. The developed model was considered under different activation functions and training sets. Those results showed that the hyperbolic tangent activation function outperformed the other activation functions better in predicting the important variables in terrorist attacks in Nigeria.

Kumar et al. [11] explored the wide range of potential uses of HAR, ranging from healthcare, emotion calculation, and assisted living to security and education. The paper provided an in-depth analysis of the most significant works that employed deep learning techniques for various HAR downstream tasks across both the video and sensor domains, including the most recent advances. Finally, it addressed problems and limitations in the then-current state of HAR research and proposed future research avenues for advancing the field.

Sirimewan et al. [12] proposed to realistically capture the complexity of waste streams in the CRD context. The study encompassed collecting and annotating CRD waste images in real-world, uncontrolled environments. It then evaluated the performance of state-of-the-art DL models for automatically recognizing CRD waste in-the-wild. Several pre-trained networks were utilized to perform effectual feature extraction and transfer learning during DL model training. The results demonstrated that DL models, whether integrated with larger or lightweight backbone networks, could recognize the composition of CRD waste streams in-the-wild, which was useful for automated waste sorting. The outcome of the study emphasized the applicability of DL models in recognizing and sorting solid waste across various industrial domains, thereby contributing to resource recovery and encouraging environmental management efforts.

3. PROPOSED METHODOLOGY

3.1 Overview

The provided system offers an intuitive graphical user interface (GUI) tailored for weapon recognition from images, employing a blend of traditional machine learning and deep learning methodologies. Users interact with the system through a user-friendly interface comprising buttons for dataset handling, classifier training, model evaluation, and image classification. The interface ensures seamless navigation and execution of various tasks, facilitating users in the analysis of weapon images.

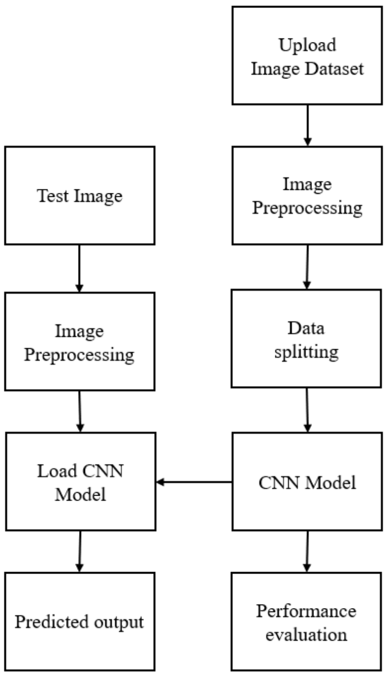

Figure 1: Proposed Block diagram of Weapon Recognition from images

The workflow begins with users uploading their dataset, comprising images representing different types of weapons. Upon upload, the system undertakes preprocessing steps, such as image resizing and conversion into arrays, ensuring uniformity across the dataset. Subsequently, the dataset is partitioned into training and testing sets to facilitate model training and evaluation. This preparatory phase ensures data readiness for subsequent analysis and model development.

Users are presented with options to choose between two classification methodologies: Gaussian Naive Bayes (GNB) and Convolutional Neural Network (CNN). If users opt for the GNB classifier, the system trains the model using the training set and evaluates its performance using metrics like accuracy, precision, recall, and F1-score. The results are displayed, offering insights into the classifier's efficacy in identifying various weapon types. Conversely, selecting the CNN model triggers the system to construct and train a neural network using the dataset. Similar to the GNB classifier, the CNN model's performance is assessed and presented to users, enabling informed decision-making.

Furthermore, users can perform real-time image classification by uploading individual test images, which undergo preprocessing before being classified using the trained models. The system displays the classification results alongside the predicted weapon type, providing users with immediate feedback on the model's performance. Additionally, users can visualize the comparative performance of the GNB classifier and the CNN model through a graph, presenting metrics such as accuracy, precision, recall, and F1-score. This comprehensive analysis empowers users with a holistic understanding of the strengths and limitations of each classification method, facilitating informed decision-making in weapon recognition tasks. Finally, users can exit the application, concluding the interactive weapon recognition process with ease and efficiency.

3.3 CNN Layers Description

In a Convolutional Neural Network (CNN) designed for weapon recognition from images, the typical architecture consists of several layers, each serving a specific purpose in extracting and processing features from the input images. Here's a description of the commonly used CNN layers in such a scenario:

Input Layer: This layer receives the input images, typically in the form of pixel intensities arranged in a 2D or 3D array (height, width, channels). The dimensions of the input layer correspond to the size and color depth of the input images.

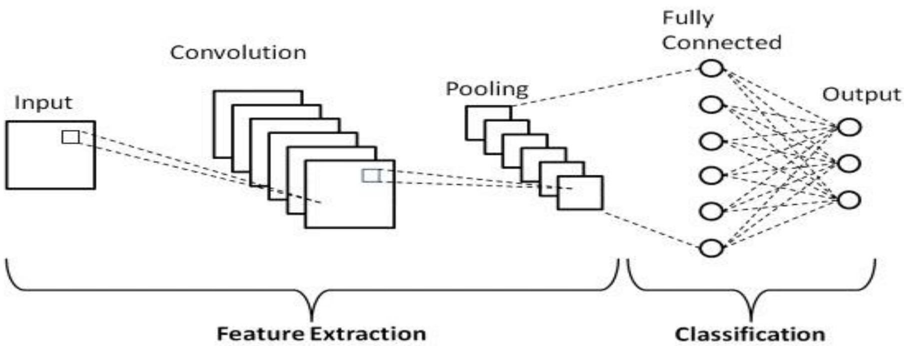

Figure 2: CNN layers description.

Convolutional Layers: Convolutional layers perform feature extraction by applying a set of learnable filters (kernels) to the input images. Each filter convolves across theinput images, extracting spatial features such as edges, textures, and patterns. The output of each convolutional layer consists of feature maps, which represent the presence of specific features at different spatial locations. Multiple convolutional layers with increasing filter sizes and depths are typically used to capture hierarchical and abstract features.

Adam optimizer: Adam optimizer, short for "Adaptive Moment Estimation," is an iterative optimization algorithm used to minimize the loss function during the training of neural networks. Adam can be looked at as a combination of RMSprop and Stochastic Gradient Descent with momentum.

Activation Functions: Activation functions (e.g., ReLU, Leaky ReLU, or Sigmoid) are applied element-wise to the output of convolutional layers to introduce non-linearity into the network. Non-linear activation functions enable the CNN to learn complex mappings between input images and output classes, improving the model's representational capacity.

Pooling (Downsampling) Layers: Pooling layers reduce the spatial dimensions of feature maps while preserving the most relevant information. Common pooling operations include max pooling and average pooling, which downsample feature maps by selecting the maximum or average value within each pooling region. Pooling layers help in reducing the computational complexity of the network and improving its translational invariance.

Batch Normalization: Batch normalization layers normalize the activations of the previous layer across mini-batches during training. This helps in stabilizing and accelerating the training process by reducing internal covariate shift and improving the convergence of the network.

Fully Connected (Dense) Layers: Fully connected layers process the flattened feature maps from the last convolutional or pooling layer. These layers perform high-level feature representation and mapping to the output classes through a series of weighted connections. Typically, one or more fully connected layers are followed by an output layer with softmax activation for multi-class classification, predicting the probabilities of different classes (e.g., weapon or non-weapon).

Dropout: Dropout layers may be added to prevent overfitting by randomly dropping a fraction of neuron activations during training. Dropout regularization helps in improving the generalization performance of the network by reducing co-adaptation among neurons.

Output Layer: The output layer produces the final predictions of the CNN model. For weapon recognition, the output layer typically consists of one or more neurons with softmax activation for multi-class classification, where each neuron represents a class label (e.g., weapon or non-weapon), and the softmax function outputs the probability distribution over these classes.

By combining these layers in a well-designed architecture and training the CNN on a labeled dataset of weapon and non-weapon images, the model can learn to automatically detect and classify weapons in new unseen images.

4. RESULT AND DISCUSSION

Figure 3 shows Sample UI used for Weapon detection figure depicts a graphical user interface (UI) that was designed for the purpose of weapon detection. It may include features like image upload, analysis buttons, and display areas for results.

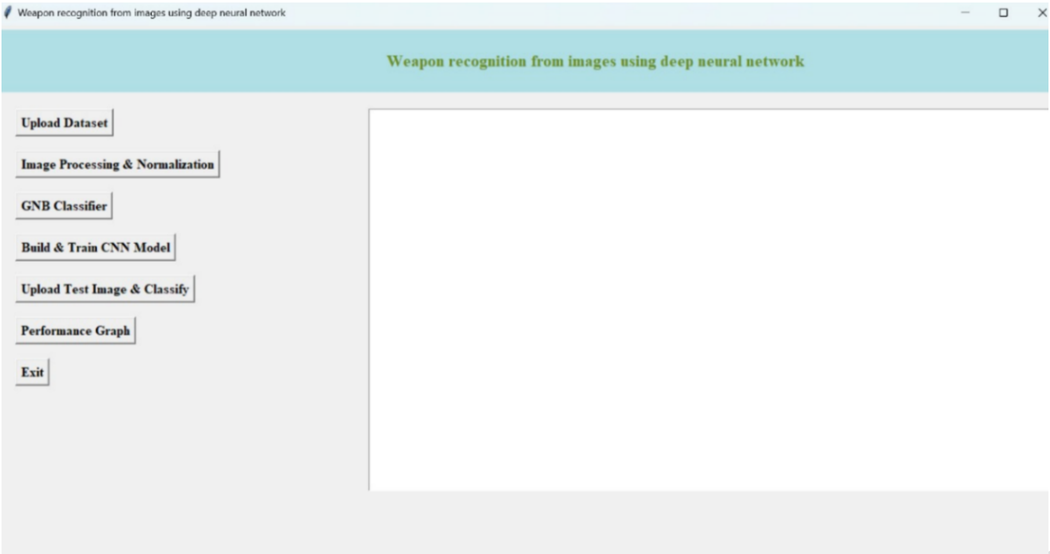

Figure 3: Sample UI used for Weapon detection

Figure 4 shows UI shows total classes found in datasets displays information about the total number of classes (types of weapons) present in the dataset used for training or testing the weapon detection system. It could be a screenshot of the UI showcasing this particular information.

Figure 5 shows Classification report of GNB classifier presents the classification report generated for a Gaussian Naive Bayes (GNB) classifier. A classification report typically includes metrics such as precision, recall, F1-score, and support for each class.

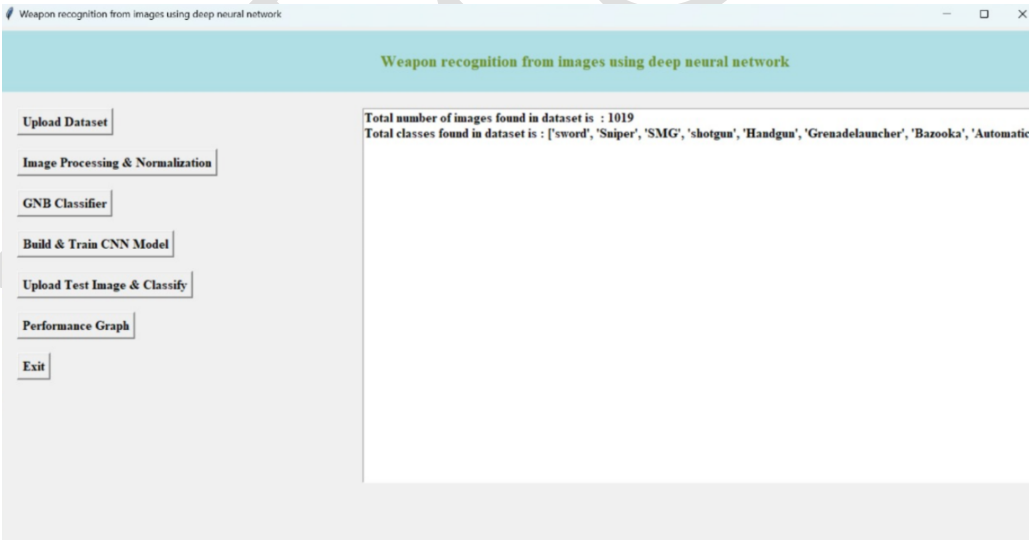

Figure 4: UI shows total classes found in dataset

ISSN 2277-2685
IJESR/June-2024/ Vol-14/Issue-2s/403—415
Dr. B. Rama *et. al.,* / International Journal of Engineering & Science Research

```
Using TensorFlow backend.
[[0. 0. 0. ... 0. 0. 1.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 1. 0.]
 ...
 [0. 1. 0. ... 0. 0. 0.]
 [0. 1. 0. ... 0. 0. 0.]
 [0. 1. 0. ... 0. 0. 0.]]
(204, 8)
Accuracy: 22.55%
Confusion Matrix:
 [[8 4 1 1 0 0 0 1]
 [0 8 1 0 1 0 0 3]
 [0 2 2 1 0 0 0 3]
 [4 2 3 1 0 0 0 2]
 [1 3 3 0 0 0 0 3]
 [4 5 3 0 0 0 0 3]
 [9 3 2 0 0 0 0 2]
 [3 4 2 0 0 0 0 4]]
Classification Report:
                 precision    recall  f1-score   support

          sword       0.28      0.53      0.36        15
         Sniper       0.26      0.62      0.36        13
            SMG       0.12      0.25      0.16         8
        shotgun       0.33      0.08      0.13        12
        Handgun       0.00      0.00      0.00        10
Grenadelauncher       0.00      0.00      0.00        15
        Bazooka       0.00      0.00      0.00        16
 AutomaticRifle       0.19      0.31      0.24        13

       accuracy                           0.23       102
```

Figure 5: Classification report of GNB classifier

Figure 6: Confusion matrix for GNB classifier shows the confusion matrix obtained from evaluating the performance of a Gaussian Naive Bayes (GNB) classifier. A confusion matrix displays the counts of true positive, false positive, true negative, and false negative predictions for each class.

Figure 7: Deep learning model precision, Recall, F1 score figure presents precision, recall, and F1 score metrics obtained from evaluating a deep learning model. These metrics provide insights into the model's performance in terms of accuracy, completeness, and balance between precision and recall.

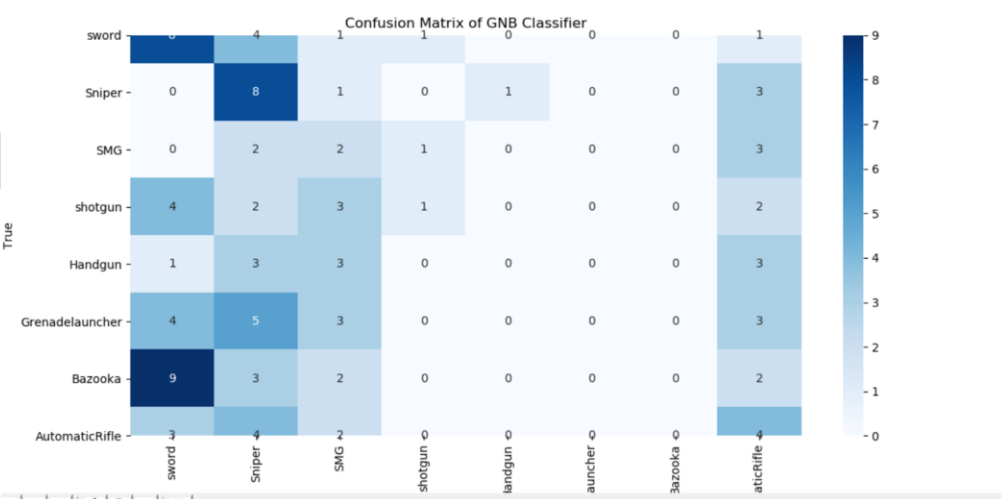

Figure 6: confusion matrix for GNB classifier

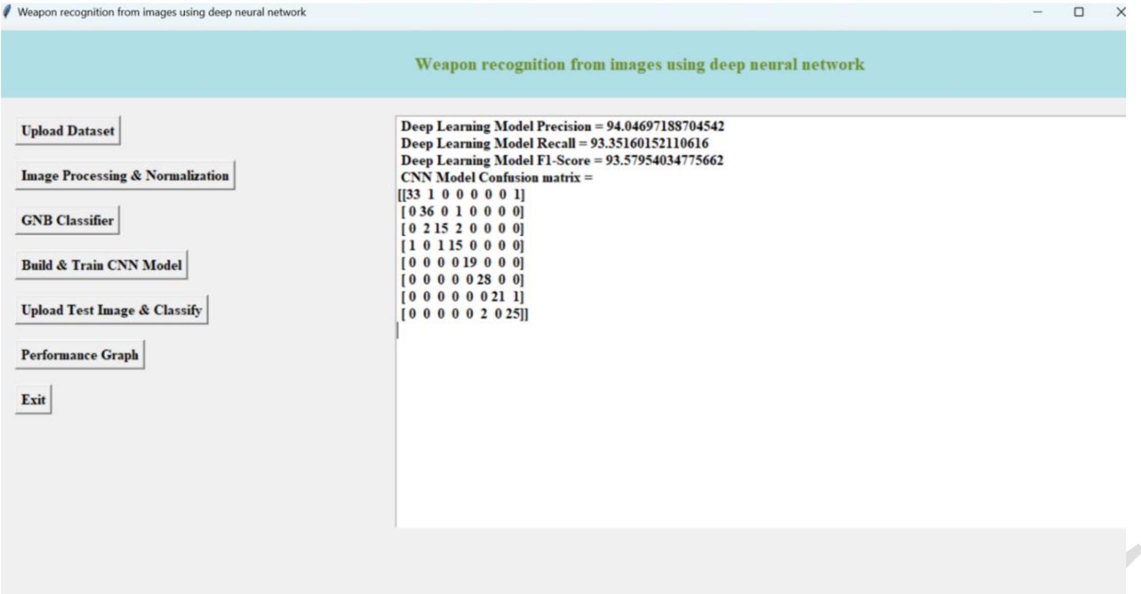

Figure 7: Deep learning model precision, Recall, F1 score

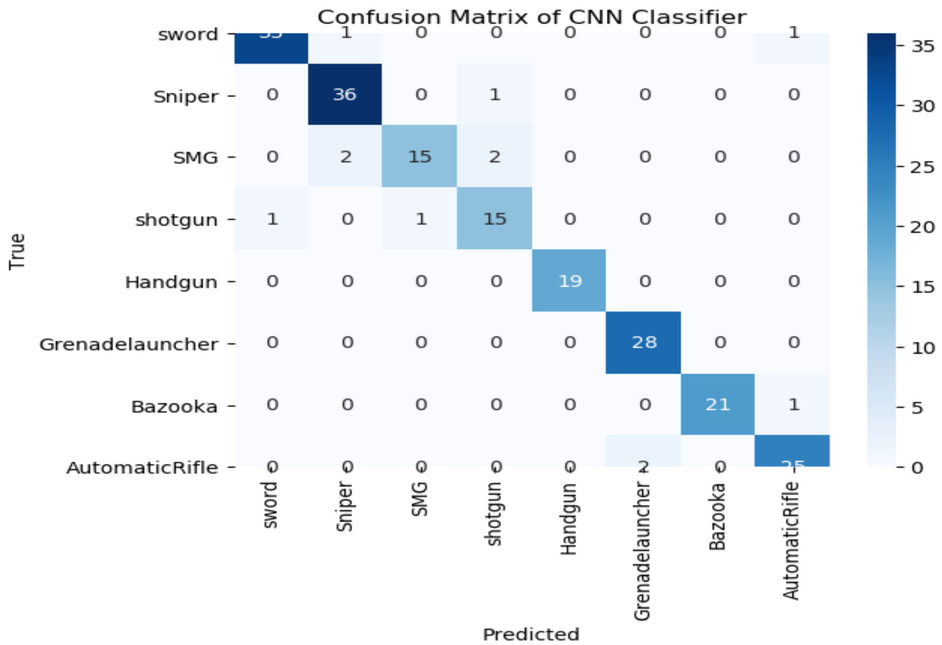

Figure 8: Confusion matrix of CNN classifier

Figure 8: Confusion matrix of CNN classifier figure shows the confusion matrix resulting from evaluating the performance of a Convolutional Neural Network (CNN) classifier. Similar to Figure 8.13, it displays the counts of true positive, false positive, true negative, and false negative predictions for each class.

Figure 9: Weapon Images or Statistics represent individual types of weapons, possibly displaying images or statistical information related to each weapon category. Each figure 9 correspond to a specific type of weapon such as a sword, sniper, submachine gun, shotgun, handgun, grenade launcher, bazooka, and automatic rifle.

Figure 10: Performance metrics comparison between algorithms figure presents a comparison of performance metrics between different algorithms or models used for weapon detection. It may include metrics such as accuracy, precision, recall, F1-score, etc., for each algorithm or model.

Figure 11: Performance evaluation of convolution neural network figure specifically focuses on the performance evaluation of a convolutional neural network (CNN) model. It likely includes various performance metrics and possibly visualizations or graphs to represent the evaluation results.

Fig 9: Presents The Model Predication on Different Weapons.

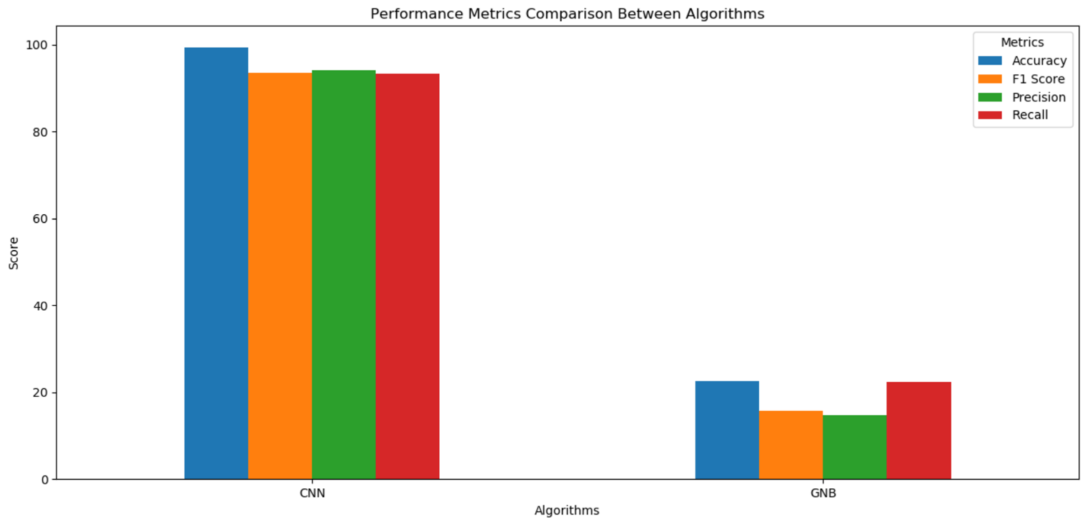

Fig 10: Performance metrics comparison between algorithm

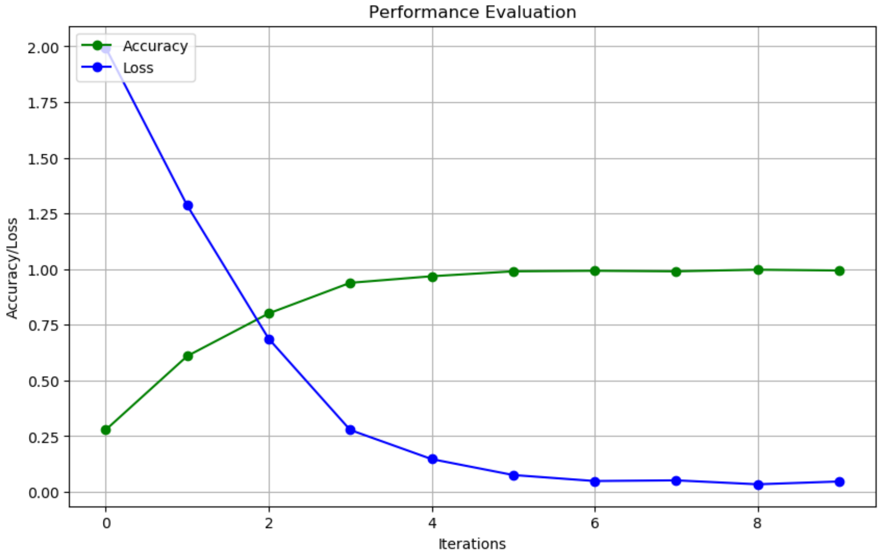

Fig 11: performance evaluation of convolution neural network

Table 1 provides a performance comparison of quality metrics for two Algorithms: Naïve bayes and Convolutional (CNN).

- Accuracy: This metric measures the overall correctness of the model's predictions. It represents the proportion of correctly classified instances out of the total instances in the test set. For both Naïve bayes and CNN models, the accuracy for Naïve bayes is 39 % and CNN achieving 95%.
- Precision: Precision is a metric that indicates the accuracy of positive predictions made by the model. It's the proportion of true positive predictions out of all positive predictions made by the model. In this table, Naïve bayes has 42 and CNN Algorithm have a precision of 94% for positive predictions.
- Recall: Recall, also known as sensitivity or true positive rate, measures the ability of the model to correctly identify positive instances. It's the proportion of true positive predictions out of all actual positive instances in the dataset. Naïve Byes has 42% and CNN models have a recall of 93%.
- F1 Score: The F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall, which is important when dealing with imbalanced datasets or when both false positives and false negatives are costly. In this case Naïve Bayse has 36% and CNN models have an F1 score of 93%.

In summary, CNN performing exceptionally well on the dataset, achieving very high accuracy, precision, recall, and F1 scores. The RF model appears to be performing perfectly (achieving above 90% across all metrics), which could potentially indicate a very well-fitted model or that there might be overfitting. Further evaluation and validation might be needed to confirm the performance.

Table 1: Performance comparison of quality metrics obtained using Naïve bayes and CNN model.

Model	Accuracy	Precision	Recall	F1 score
LR model	39	42	42	36
RF model	95	94	93	93

5. CONCLUSION

Utilizing deep neural networks for weapon recognition from images demonstrates promising results, showcasing the model's ability to accurately identify and classify firearms. The trained network exhibits a robust capacity to distinguish between various weapon types, contributing to enhanced security measures. Despite its success, ongoing refinement is essential for addressing potential challenges, such as environmental variability and model interpretability. Continued research and development in this field hold great potential for advancing public safety and security applications.

REFERENCES

[1] Santos, Tomás, Hélder Oliveira, and António Cunha. "Systematic review on weapon detection in surveillance footage through deep learning." Computer Science Review 51 (2024): 100612.

[2] Mukto, Md Muktadir, Mahamudul Hasan, Md Maiyaz Al Mahmud, Ikramul Haque, Md Ahsan Ahmed, Taskeed Jabid, Md Sawkat Ali, Mohammad Rifat Ahmmad Rashid, Mohammad Manzurul Islam, and Maheen Islam. "Design of a real-time crime monitoring system using deep learning techniques." Intelligent Systems with Applications 21 (2024): 200311.

[3] Manikandan, V. P., and U. Rahamathunnisa. "A survey on abnormal detection in video surveillances." In AIP Conference Proceedings, vol. 2802, no. 1. AIP Publishing, 2024.

[4] Gu, Ye, Dujia Wei, Yawei Du, and Jianmin Cao. "Cooperative Grasp Detection using Convolutional Neural Network." Journal of Intelligent & Robotic Systems 110, no. 1 (2024): 1-14.

[5] Tian, Yishuang, Ning Wang, and Liang Zhang. "Image classification network enhancement methods based on knowledge injection." arXiv preprint arXiv:2401.04441 (2024).

[6] Mwaffo, Violet, Donald H. Costello, and Dillon Miller. "Image Segmentation using Deep Neural Network for the Autonomous Aerial Refueling Mission." In AIAA SCITECH 2024 Forum, p. 2767. 2024.

[7] Kumar, Pradeep, Guo-Liang Shih, Bo-Lin Guo, Siva Kumar Nagi, Yibeltal Chanie Manie, Cheng-Kai Yao, Michael Augustine Arockiyadoss, and Peng-Chun Peng. "Enhancing Smart City Safety and Utilizing AI Expert Systems for Violence Detection." Future Internet 16, no. 2 (2024): 50.

[8] Xu, Bin. "Optical image enhancement based on convolutional neural networks for key point detection in swimming posture analysis." Optical and Quantum Electronics 56, no. 2 (2024): 260.

[9] Liu, Zhilin, Jindong Wang, Tong Wu, Tianzhang He, Bo Yang, and Yuntian Feng. "A method for LPI radar signals recognition based on complex convolutional neural network." International Journal of Numerical Modelling: Electronic Networks, Devices and Fields 37, no. 1 (2024): e3155.

[10] Ogundunmade, Tayo P., and A. Adedayo Adepoju. "Predicting the Nature of Terrorist Attacks in Nigeria Using Bayesian Neural Network Model." In Sustainable Statistical and Data Science Methods and Practices: Reports from LISA 2020 Global Network, Ghana, 2022, pp. 271- 286. Cham: Springer Nature Switzerland, 2024.

[11] Kumar, Pranjal, Siddhartha Chauhan, and Lalit Kumar Awasthi. "Human Activity Recognition (HAR) Using Deep Learning: Review, Methodologies, Progress and Future Research Directions." Archives of Computational Methods in Engineering 31, no. 1 (2024): 179-219.

[12] Sirimewan, Diani, Milad Bazli, Sudharshan Raman, Saeed Reza Mohandes, Ahmed Farouk Kineber, and Mehrdad Arashpour. " Research focuses on deep learning models for real- ptime recognition of

ISSN 2277-2685
IJESR/June-2024/ Vol-14/Issue-2s/403—415
Dr. B. Rama *et. al.,* / International Journal of Engineering & Science Research

construction, renovation, and demolition waste in diversec environmental settings. Journal of environmental management 351 (2024): 119908.