
DEEP LEARNING-BASED WEAPON CLASSIFICATION FOR INTELLIGENT SURVEILLANCE APPLICATIONS USING RESNET-50

Kenneth Okhiria^{1,*}, Joseph Samuel Mshelizah¹, Saidi Abdulraimi

¹Nigerian Army University Biu

Email: kennyhnsom@gmail.com*, Samueljoseph9272@gmail.com, Saheeday12@gmail.com

*Corresponding author: Kenneth Okhiria

DOI: <https://doi.org/10.67203/sp.ICESID.2026>

ABSTRACT

Weapon-related crimes pose significant threats to public safety and security, creating a growing need for intelligent surveillance systems capable of automatically identifying potential threats. Deep learning has demonstrated remarkable success in object recognition and image classification, making it well suited for automated weapon recognition. This study develops and evaluates a transfer learning-based approach for classifying two common weapon categories, AK-47 rifles, and knives, using a pretrained ResNet-50 convolutional neural network. A dataset comprising 2,001 labeled images was collected from Google Images, manually cleaned, preprocessed, and divided into training and validation sets using an 80:20 ratio. The pretrained ResNet-50 model was adapted using transfer learning and trained for 50 epochs. Model performance was evaluated using precision, recall, F1-score, and overall accuracy. Experimental results demonstrated high classification performance on the validation dataset, with the AK-47 class achieving a precision of 0.99, recall of 1.00, and F1-score of 0.99, while the Knife class achieved precision, recall, and F1-score values of 0.99, resulting in an overall classification accuracy of 0.99. Additional inference on previously unseen weapon images produced accurate predictions with consistently high confidence scores, providing supplementary evidence of the model's predictive capability beyond the validation samples. These findings demonstrate the effectiveness of the proposed approach as a proof-of-concept classification framework, highlighting its foundational potential for future integration into real-time intelligent surveillance and public safety applications.

Keywords: Weapon; Resnet-50; Transfer- Learning; Classification

1. INTRODUCTION

Video surveillance systems have become an integral component of modern public safety and security, playing a vital role in crime prevention and incident monitoring. Security cameras are widely deployed in transportation infrastructure, public housing complexes, shopping districts, border control areas, and public buildings to monitor activities and enhance security. Despite their widespread deployment, the monitoring of surveillance systems and the identification of potential threats still rely heavily on continuous human observation by security personnel. As the number of surveillance cameras continues to increase, maintaining effective real-time monitoring has become increasingly challenging, with operators often reviewing recorded footage only after incidents have occurred to reconstruct events and gather evidence. Furthermore, prolonged surveillance tasks can result in operator fatigue and human error, reducing the effectiveness of manual monitoring. These limitations highlight the growing need for intelligent and proactive surveillance systems capable of supporting timely threat identification (Kaya *et al.*, 2021).

Weapon-related violence remains a significant public safety concern worldwide, posing serious risks to human life and property. Early identification of dangerous weapons can significantly improve the ability of law enforcement and security agencies to respond promptly and mitigate potential harm. To overcome the limitations of manual surveillance, researchers have increasingly explored automated surveillance systems capable of recognizing weapons before violent incidents occur. Consequently, artificial intelligence (AI)-based surveillance solutions have been developed to detect and recognize dangerous objects within surveillance imagery. By providing faster and more consistent analysis than manual observation alone, AI can support security personnel, reduce operator workload, alleviate workforce shortages, and improve the overall efficiency of surveillance operations (Lim *et al.*, 2019; Xu & Hung, 2020; Mehta *et al.*, 2020; Gelana & Yadav, 2019; Dwivedi *et al.*, 2019).

Recent advances in artificial intelligence have been largely driven by deep learning, which has significantly improved the ability of computer systems to analyze and interpret visual information. Among deep learning models, Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in image classification, object detection, and image segmentation by automatically learning hierarchical feature representations directly from image data. Their ability to extract discriminative visual features without relying on handcrafted feature engineering has made CNNs the dominant architecture for computer vision and intelligent surveillance applications (Verma & Dhillon, 2017; Burgaz & Budak, 2022; Kaya *et al.*, 2021; Oura *et al.*, 2021; Hashmi *et al.*, 2021; Ağdaş *et al.*, 2021). Consequently, CNN-based approaches have substantially improved the automatic recognition of weapons and other dangerous objects, thereby enhancing situational awareness and supporting more effective security operations (Lim *et al.*, 2019).

A reliable weapon classification system can significantly enhance video surveillance, smart home security, intrusion detection, and smart city monitoring by providing timely alerts that support rapid decision-making and emergency response (Warsi *et al.*, 2019). Early recognition of dangerous weapons enables security personnel to intervene before situations escalate, thereby reducing potential risks to public safety (Yadav *et al.*, 2023). Among commonly encountered weapons in security-related incidents, firearms such as the AK-47 and bladed weapons such as knives present significant security concerns because of their potential use in violent attacks. Consequently, this study focuses on the classification of AK-47 rifles and knives to support intelligent surveillance applications (Dwivedi *et al.*, 2019; Ingle & Kim, 2022).

One of the major advantages of modern CNN architectures is their ability to support transfer learning, whereby knowledge acquired from large-scale image datasets can be transferred to related tasks with limited training data. Rather than training a model from scratch, transfer learning enables pre-trained networks to be fine-tuned for specific applications, thereby reducing computational cost, shortening training time, and improving model performance when only modest datasets are available. These advantages make transfer learning particularly suitable for intelligent surveillance applications, where obtaining large, well-annotated weapon image datasets can be challenging (Dwivedi *et al.*, 2019; Warsi *et al.*, 2019). Among the available deep learning architectures, ResNet-50 has become one of the most widely adopted models because of its ability to learn rich hierarchical feature representations while mitigating the vanishing gradient problem through residual learning (Kaya *et al.*, 2021).

ResNet-50 has demonstrated excellent performance in image classification and surveillance applications because of its strong feature extraction capability and compatibility with transfer learning. Previous studies have successfully applied ResNet-50 to weapon recognition and firearm classification tasks, demonstrating robust performance under diverse operating conditions (Salido *et al.*, 2021; Ahmed *et al.*, 2022; Kambhatla & Khaled, 2023). However, many existing studies have focused on object detection frameworks, larger datasets, or broader weapon categories, while comparatively fewer have investigated the effectiveness of transfer learning with ResNet-50 for the classification of specific weapon categories using relatively modest datasets. Therefore, this study develops and evaluates a transfer learning-based framework as a proof-of-concept classification study. Rather than real-time localization, this work establishes a diagnostic baseline for distinguishing AK-47 rifles and knives using clean web-scraped image repositories to evaluate supervised image classification performance prior to complex video stream deployment.

2. METHODOLOGY

The proposed methodology consists of four main stages: data acquisition, data preprocessing, model training, and weapon classification. Initially, a dataset containing images of AK-47 rifles and knives was collected and prepared for model development. The images were subsequently preprocessed to improve data quality and ensure compatibility with the deep learning model. During the training stage, a transfer learning approach based on the pre-trained ResNet-50 architecture was employed to learn discriminative features for weapon classification. Finally, the trained model was used to classify unseen query images into their respective weapon categories. The overall methodology is illustrated in Figure 1, while the detailed workflow of the proposed surveillance system is presented in Figure 2.

2.1 Proposed Surveillance System

The proposed intelligent surveillance system follows a sequential workflow comprising four major phases: data acquisition, data preprocessing, model training, and weapon classification. In the data acquisition stage, images representing the target weapon categories were collected to construct the dataset. The preprocessing stage involved resizing and normalizing the images to prepare them for deep learning. The processed images were then used to train the ResNet-50 model through transfer learning, enabling the extraction of discriminative visual features for accurate weapon classification. In the final stage, unseen query images were presented to the trained model, which predicted the corresponding weapon class. The individual stages of the proposed methodology are described in detail in the following subsections.

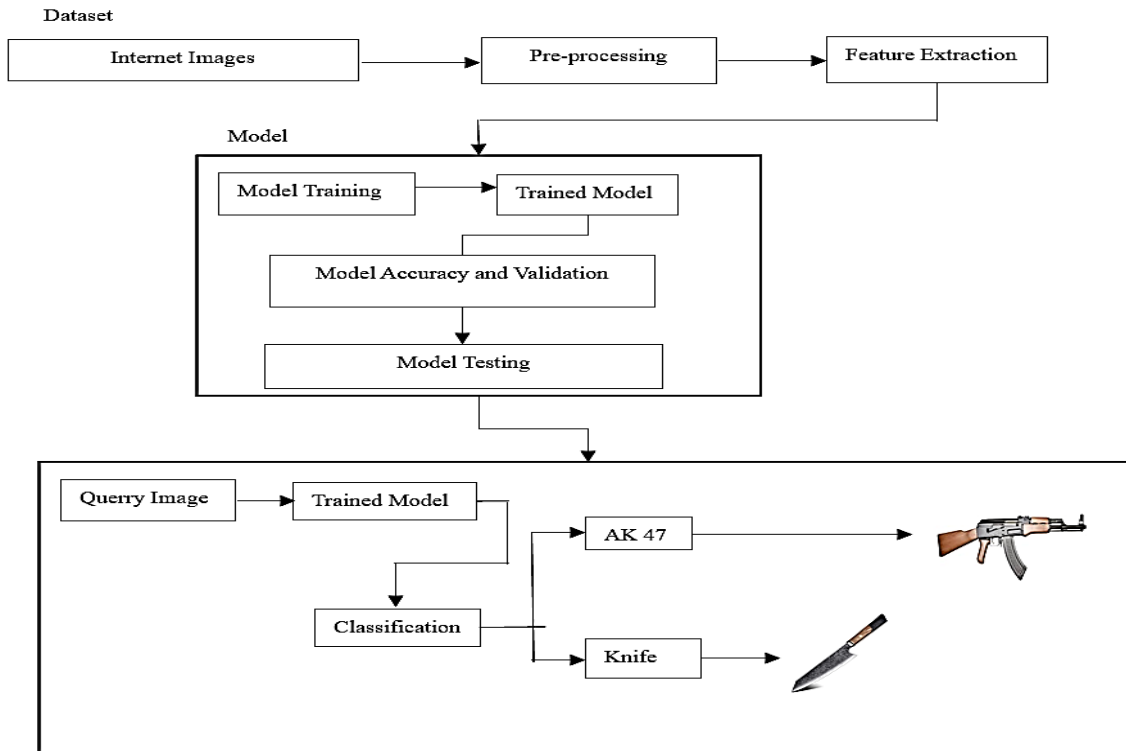


Figure 1: Block diagram of proposed methodology

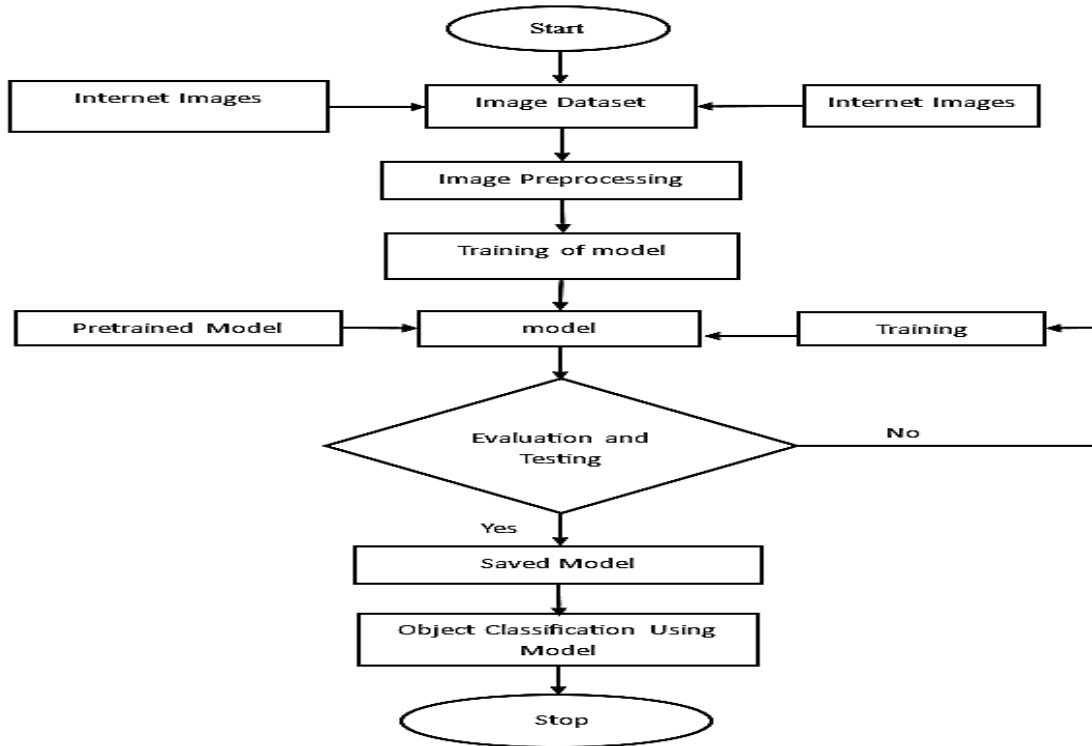


Figure 2: Methodology flowchart

2.2 Data Acquisition

During the data acquisition phase, a dataset comprising 2,001 images of AK-47 rifles and knives was collected from Google Images to support the development of the proposed weapon classification model. Search queries related to the two weapon categories were used to retrieve images captured under different viewpoints, backgrounds, lighting conditions, and object orientations in order to improve dataset diversity. The collected images were manually inspected to ensure their relevance

to the target classes before undergoing preprocessing. Representative samples of the acquired images are presented in Figure 3.



Figure 3: Image Dataset

2.3 Data Preprocessing

Before model training, the collected images underwent a series of preprocessing operations to ensure that the dataset was properly organized, standardized, and optimized for the ResNet-50 architecture. The preprocessing pipeline comprised image collection and cleaning, dataset organization, image resizing, image normalization, and dataset splitting. These operations improved data consistency, enhanced training efficiency, and supported stable model training.

2.3.1 Image Collection, Cleaning and Dataset Organization

The dataset comprised images belonging to two weapon categories: AK-47 rifles and knives. These images were collected from online sources, primarily Google Images, using search queries related to the target weapon categories. To ensure dataset quality and relevance, all collected images were manually inspected, and duplicate, irrelevant, corrupted, and low-resolution samples were removed. This cleaning process ensured that only clear and representative images of each weapon category were retained for model development.

Following the cleaning process, the images were organized into two class-specific subdirectories, namely AK47 and Knife, within a parent directory named ResNet_Dataset. This hierarchical directory structure enabled the PyTorch ImageFolder utility to automatically assign class labels during data loading, thereby facilitating supervised learning. Proper dataset organization also ensured that each image was correctly associated with its corresponding weapon category throughout the training and validation processes, providing a consistent and efficient data management framework for model development.

2.3.2 Image Resizing

All images were resized to 224×224 pixels, corresponding to the input dimensions required by the ResNet-50 architecture. Uniform image dimensions ensured consistency across all samples, reduced computational complexity, and enabled the direct use of ImageNet pre-trained weights without modifying the network architecture.

2.3.3 Image Normalization

Following image resizing, pixel values were normalized to improve training stability and enhance model convergence. The pixel intensities were first scaled to the range **[0, 1]** and subsequently standardized using the ImageNet mean and standard deviation values, consistent with the preprocessing requirements of the pre-trained ResNet-50 model. This normalization process reduced the effects of illumination and contrast variations across the dataset, improved numerical stability during optimization, and ensured compatibility with the feature representations learned from the ImageNet dataset.

2.3.4 Data Augmentation

To improve the variability of the training samples and enhance the model's ability to recognize weapons under different image conditions, data augmentation was applied dynamically during the training process. The augmentation operations included random resized cropping, random horizontal flipping, random rotation within a specified range, and colour-jitter adjustment of brightness and contrast. These transformations generated varied representations of the original training images during model training without creating or permanently storing additional image files. The validation images were not subjected to these random augmentation operations; instead, they were processed using deterministic resizing and centre cropping before normalization. Thus, the dataset remained at 2,001 original images, while augmentation increased the variability of the samples presented to the model during training.

2.3.5 Dataset Splitting

Following preprocessing, the prepared dataset of 2,001 images was randomly partitioned into training and validation subsets using an approximately 80:20 ratio to facilitate model development and performance evaluation. A total of 1,600 images were allocated to the training set, while the remaining 401 images were reserved for validation, as summarized in Table 1. The partitioning process maintained class representation between the AK-47 and Knife categories, with 800 AK-47 and 800 Knife images used for training, and 201 AK-47 and 200 Knife images used for validation.

The validation set was used to monitor model performance during training and to assess its classification performance after training. A separate held-out test set was not created because of the moderate dataset size of 2,001 images, allowing a larger proportion of the available data to be retained for model development. Consequently, the performance metrics reported in this study are based on the validation set and may be influenced by model-development and hyperparameter-selection decisions. The absence of an independent test set is therefore acknowledged as a limitation of the study. Future work should evaluate the model using a separate unseen test set or an external dataset to provide a more independent estimate of its generalization performance.

Table 1: Data Set Distribution

Classification	Training Set	Validation Set	Total
AK47	800	201	1001
Knife	800	200	1000
Total	1600	401	2001

2.4 Transfer Learning and ResNet-50 Architecture

The proposed weapon classification framework was developed using the ResNet-50 deep convolutional neural network through a transfer learning approach. This section describes the adopted ResNet-50 architecture and the transfer learning strategy employed to adapt the pre-trained network to the binary weapon classification task.

2.4.1 ResNet-50 Architecture

ResNet-50 was adopted as the backbone convolutional neural network for this study because of its strong feature extraction capability, computational efficiency, and proven performance in image classification tasks. The architecture comprises 50 learnable layers organized into residual blocks with skip connections that facilitate efficient gradient propagation during training. The model was initialized with ImageNet pre-trained weights to provide robust feature representations for the weapon classification task. The standard ResNet-50 architecture used in this study is summarized in Table 2 based on the architecture described by (He *et al.*, 2016).

Table 2: ResNet-50 Architecture

Layer Name	Output size	50-layer
Conv 1_×	112 x 112	7 x7, 64, stride 2 3 x 3 max pool, stride 2
Conv 2_×	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv 3_×	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv 4_×	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv 5_×	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax

2.4.2 Transfer Learning Strategy

A transfer learning strategy was employed to adapt the pre-trained ResNet-50 model to the binary weapon classification task. The model was initialized with ImageNet pre-trained weights, while the final fully connected layer was replaced with a new layer comprising two output neurons corresponding to the AK-47 and Knife classes. During transfer learning, the pretrained ResNet-50 feature-extraction layers were kept frozen, while the newly replaced fully connected classification layer was trained for the two-class weapon classification task. This approach reduced training time and supported efficient adaptation of the pretrained model to the target classification task.

2.5 Model Training

Following dataset preprocessing and configuration of the ResNet-50 architecture, the modified network was trained to classify images into the AK-47 and Knife categories. The training process involved configuring the model hyperparameters, optimizing the network using an appropriate loss

function and optimization algorithm, and monitoring its convergence throughout training. The training procedure and adopted hyperparameters are described in the following subsections.

2.5.1 Training Procedure

Following dataset preparation and partitioning, the modified ResNet-50 model was trained using the training subset, while the validation subset was used to monitor the model's learning progress and performance throughout the training process. The dataset was processed in mini batches of 32 images, and the model was trained for 50 epochs. During each epoch, both the training and validation losses were monitored to assess convergence, evaluate model stability, and minimize the risk of overfitting.

2.5.2 Training Parameters

The principal hyperparameters adopted during model training are summarized in Table 3. These parameters were selected to achieve stable convergence while maintaining computational efficiency throughout the training process.

Table 3: Training Parameters

Related Parameter	Value
Batch size	32
Learning rate	1×10^{-4}
Optimizer	Adam
Number of Epochs	50
Loss Function	Cross Entropy Loss
Backbone model	Pre-trained model, ResNet-50 (ImageNet)

2.6 Performance Evaluation

The performance of the proposed ResNet-50 model was evaluated using four widely adopted classification metrics: accuracy, precision, recall, and F1-score. These metrics provide complementary information regarding the model's overall classification performance, prediction reliability, and ability to correctly distinguish between the AK-47 and Knife classes. The mathematical expressions of the evaluation metrics are presented in Table 4.

In the evaluation process, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) denote the correctly and incorrectly classified samples used in computing the performance metrics.

Table 4: Performance Evaluation Metrics. Source (Burgaz and Budak, 2022)

Metric	Equation
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1 score	$2 \times \frac{Precision \times Recall}{Precision + Recall}$
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$

3. RESULTS AND DISCUSSION

This section presents and discusses the performance of the proposed ResNet-50-based weapon classification model. Following model training, the network was evaluated using the validation dataset to assess its ability to distinguish between the AK-47 and Knife classes. The experimental results are presented in terms of training performance, classification metrics, and prediction capability on unseen images, demonstrating the effectiveness of the proposed approach for intelligent surveillance applications.

3.1 Training and Validation Performance

The training and validation performance of the proposed ResNet-50 model was evaluated using the loss and accuracy recorded over 50 training epochs. The corresponding learning curves are presented in Figure 4, where Figure 4(a) illustrates the training and validation loss, while Figure 4(b) presents the training and validation accuracy.

As shown in Figure 4(a), both the training and validation loss decreased progressively throughout the training process, indicating that the model progressively learned discriminative features for the classification task. However, the validation loss remained lower than the training loss during most of the training epochs. Similarly, Figure 4(b) shows that both the training and validation accuracy increased progressively before stabilizing at high values, with the validation accuracy remaining higher than the training accuracy. This pattern is attributable, in part, to the different preprocessing procedures applied during training and validation. During training, random resized cropping, horizontal flipping, rotation, and colour-jitter augmentation were applied to the images, introducing additional variability into the samples presented to the model. In contrast, the validation images were processed using deterministic resizing and center cropping without random augmentation. Consequently, the training accuracy and loss reflect the model's performance on dynamically transformed training samples, whereas the validation metrics reflect performance on consistently processed validation samples. The higher validation performance therefore does not, by itself, indicate data leakage or overfitting, but reflects, in part, the different preprocessing conditions applied during training and validation. Overall, the learning curves indicate progressive convergence and high classification performance, with the difference between the training and validation curves being consistent with the different preprocessing conditions adopted in the experiment.

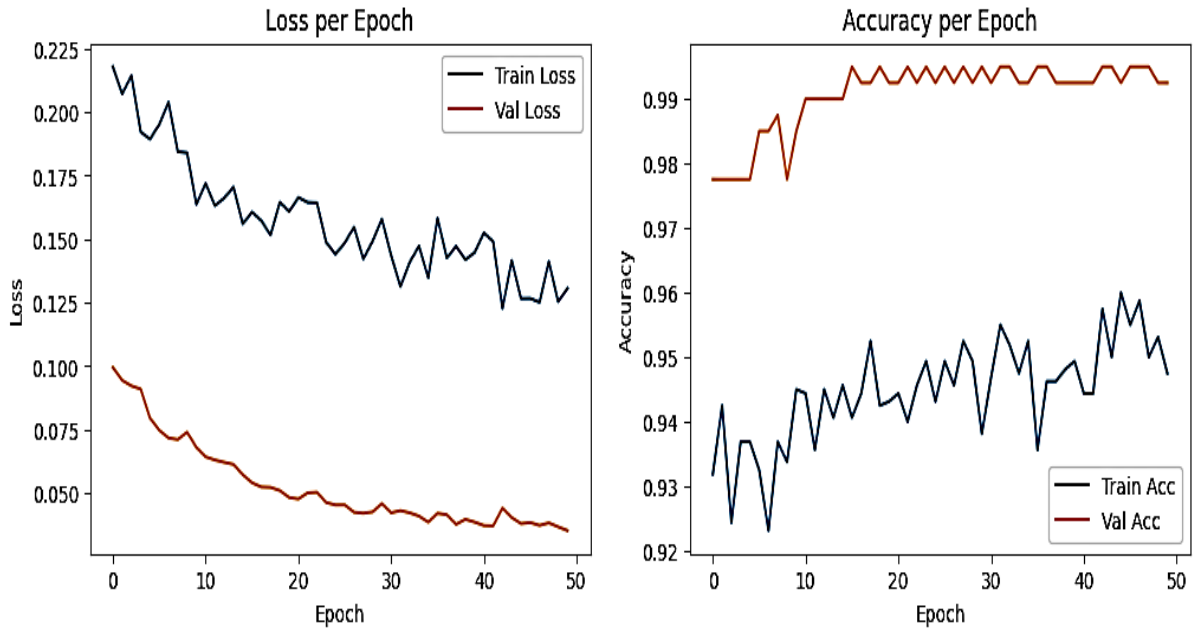


Figure 4: Training and Validation Loss Curves

3.2 Classification Performance

The classification performance of the proposed ResNet-50 model was evaluated using precision, recall, F1-score, and overall classification accuracy. The quantitative performance metrics are presented in Table 5, while the corresponding confusion matrix is shown in Figure 5.

As presented in Table 5, the proposed model achieved excellent classification performance for both weapon categories. For the AK-47 class, the model attained a precision of 0.99, recall of 1.00, and an F1-score of 0.99. Similarly, the Knife class achieved precision, recall, and F1-score values of 0.99, with an overall classification accuracy of 0.99. These results demonstrate that the fine-tuned ResNet-50 model effectively learned discriminative visual features that enabled accurate differentiation between the two weapon categories.

The confusion matrix shown in Figure 5 further validates the classification performance of the proposed model. The model correctly classified 200 AK-47 images and 198 Knife images, with only a small number of misclassifications. Specifically, one AK-47 image was incorrectly classified as a Knife, while two Knife images were misclassified as AK-47. The low number of false positive and false negative predictions is consistent with the high precision, recall, F1-score, and overall accuracy obtained during evaluation. The low number of misclassifications is consistent with the high precision, recall, F1-score, and overall accuracy obtained during evaluation. These results demonstrate the effectiveness of the transfer learning-based ResNet-50 model for the two-class weapon classification task under the image-based conditions investigated.

Table 5: Classification Performance of the Proposed ResNet-50 Model

Object Classifier	Class	Precision	Recall	F1-Score	Accuracy
ResNet-50	AK47	0.99	1.00	0.99	0.99
	Knife	0.99	0.99	0.99	0.99

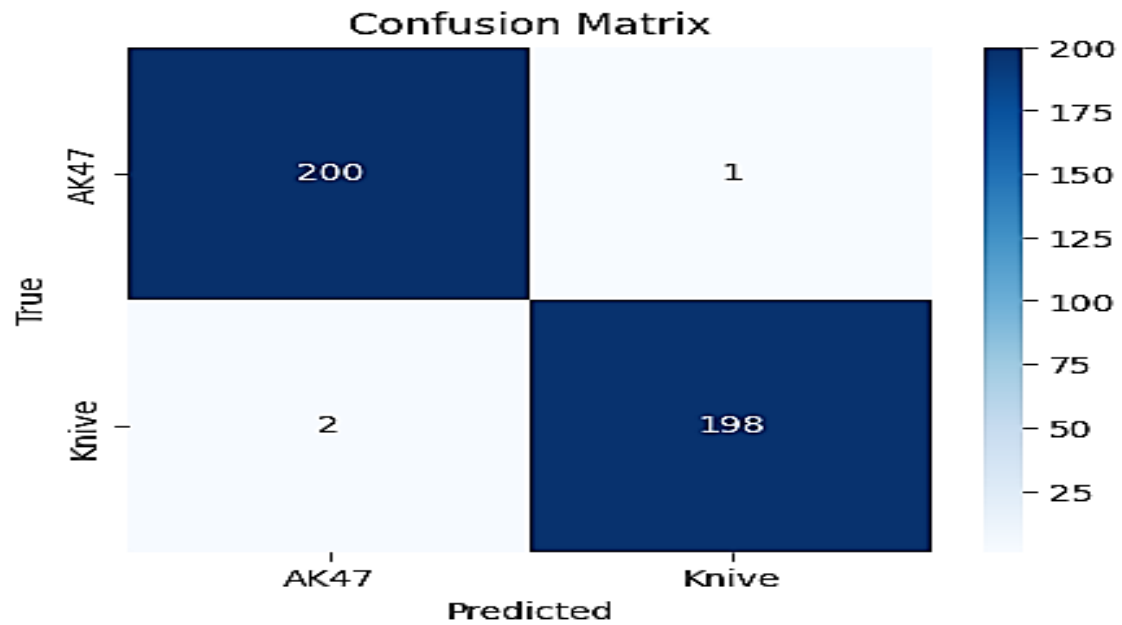


Figure 5: Confusion Matrix of the ResNet-50 Model

3.3 Prediction Results on Validation Images

To further assess the performance of the proposed ResNet-50 model, inference was carried out on representative images from the validation dataset. Sample prediction results are presented in Figure 6 where each image is displayed alongside its predicted class label. The results show that the model correctly identified both AK-47 and Knife images with high confidence, demonstrating its ability to extract discriminative features from images that were not directly used for model parameter optimization.

The prediction results are consistent with the quantitative performance metrics presented in the previous section. The model successfully recognized variations in object orientation, scale, and background while maintaining accurate classification performance. These findings indicate that the trained ResNet-50 model accurately distinguished between the two weapon categories in the evaluated validation samples under the image-based conditions investigated.



Figure 6: Predicted Result on Validation Images

3.4 Prediction Results on Unseen Images

To further evaluate the generalization capability of the proposed ResNet-50 model, inference was performed on completely unseen weapon images that were not included in either the training or validation datasets.

The model accurately classified all the unseen images into their respective categories, namely AK-47 and Knife, with very high confidence scores. Specifically, the AK-47 images were correctly recognized with confidence scores of 99.91% and 99.97%, while the Knife images achieved confidence scores of 99.81% and 99.93%, respectively. These results demonstrate that the model not only identified each weapon category correctly but also exhibited a high degree of confidence in its predictions. The result is shown in figure 7.

The consistently high confidence scores indicate the strong feature extraction capability of the ResNet-50 architecture and the effectiveness of the adopted transfer learning strategy. Furthermore, the model maintained accurate predictions across the examined unseen images, which included variations in image background, illumination, and weapon orientation. These results provide supplementary evidence of the model's predictive capability beyond the training and validation samples. Overall, the inference results demonstrate that the proposed ResNet-50-based classifier can accurately distinguish the two weapon categories under the image-based conditions investigated. The results provide a proof-of-concept baseline for potential integration into intelligent surveillance applications, although further evaluation using real-time surveillance video and camera-captured data is required.



Figure 7: Prediction Result on Unseen images

3.5 Discussion of Findings

The experimental results demonstrate that the proposed ResNet-50-based weapon classification model is capable of accurately distinguishing between AK-47 rifles and knives under the image-based conditions investigated. The high values obtained for precision, recall, F1-score, and overall accuracy indicate that the model learned discriminative visual features for the two weapon categories. The validation results also show that the model achieved high classification performance on images that were not used for parameter updates during training. However, because the study employed a training-validation split without an independent held-out test set, the validation performance should be interpreted as an assessment of model performance within the experimental dataset rather than as an unbiased estimate of performance on completely independent data.

The high classification performance can be attributed to the adoption of the ResNet-50 architecture and the transfer learning strategy employed in this study. Residual learning enables the network to learn hierarchical feature representations while mitigating the vanishing gradient problem associated with deep neural networks. Furthermore, initializing the model with ImageNet pre-trained weights provided useful feature representations that supported the classification task using the available weapon image dataset. These findings are consistent with previous studies that have reported the effectiveness of deep convolutional neural networks and transfer learning for image classification and weapon recognition tasks.

The confusion matrix and inference results on the validation and unseen images provide further evidence of the model's ability to distinguish between the two weapon categories under the image-based conditions investigated. The small number of misclassifications in the validation results indicates that the learned feature representations were sufficiently discriminative for the two-class classification task. The high confidence scores obtained during inference also indicate that the model produced confident predictions for the examined unseen images. However, these results should not be interpreted as evidence of robustness under operational surveillance conditions, since the evaluation did not involve continuous video streams or camera-captured surveillance footage. The ability to produce accurate and confident predictions is nevertheless a desirable characteristic for a model intended for future integration into intelligent surveillance applications.

The learning curves provide additional insight into the training behaviour of the model. Both training and validation loss decreased progressively, while training and validation accuracy increased before stabilizing at high values. The validation accuracy remained higher than the training accuracy, while the validation loss remained lower than the training loss during most of the training process. This pattern is attributable, in part, to the different preprocessing conditions applied during training and validation. Random resized cropping, horizontal flipping, rotation, and colour-jitter augmentation were

applied dynamically to the training images, introducing additional variability into the samples presented to the model. In contrast, the validation images were processed using deterministic resizing and center cropping without random augmentation. The higher validation performance therefore does not, by itself, indicate data leakage or overfitting, but reflects, in part, the different preprocessing conditions applied during training and validation.

Despite the promising performance achieved in this study, certain limitations should be acknowledged. The developed model was trained and evaluated using only two weapon categories, namely AK-47 rifles and knives, and the dataset consisted primarily of images collected from online sources. The model was not evaluated under challenging real-world surveillance conditions such as occlusion, motion blur, low illumination, cluttered backgrounds, varying camera viewpoints, or continuous video streams captured from fixed surveillance cameras or unmanned aerial vehicles. In addition, the study focused on image classification rather than object detection and localization. Future research should therefore consider expanding the dataset to include additional weapon categories, evaluating the model using independent test data and real-time surveillance video, and investigating object-detection architectures and more recent deep learning models to improve robustness and support eventual integration into practical intelligent surveillance applications.

4. CONCLUSIONS

This study presented a deep learning-based weapon classification framework relevant to intelligent surveillance applications using the ResNet-50 architecture and a transfer learning strategy. The pretrained ResNet-50 model was adapted to classify two weapon categories, namely AK-47 rifles and knives. The developed framework demonstrated the application of transfer learning to image-based weapon classification using a relatively modest dataset.

A complete weapon classification pipeline was developed using the PyTorch deep learning framework in the Google Colab environment. The collected dataset was preprocessed, normalized, and organized to facilitate model training and validation. The model exhibited stable convergence throughout the 50 training epochs, as reflected in the training and validation loss and accuracy curves. The validation results further demonstrated high classification performance under the image-based experimental conditions adopted in this study.

Performance evaluation showed that the proposed ResNet-50 model achieved an overall accuracy of 0.99, with precision, recall, and F1-score values of 0.99 or higher for the two weapon classes. The confusion matrix and inference results on unseen images further demonstrated the model's ability to distinguish between the two weapon categories, with prediction confidence scores exceeding 99% for the examined unseen samples. These findings establish the proposed approach as a proof-of-concept image classification framework for automated weapon recognition.

However, the study does not establish the performance of the model in operational intelligent surveillance environments. The experiment was limited to static images collected primarily from online sources and focused on binary image classification rather than real-time object detection and localization. The model was not evaluated using continuous surveillance video, camera-captured data, occlusion, motion blur, low illumination, cluttered backgrounds, or other challenging operational conditions. Therefore, the results demonstrate the potential of the ResNet-50 approach as a baseline for future intelligent surveillance integration rather than direct deployment readiness.

Future research should expand the dataset to include a broader range of weapon categories and establish an independent test set for more robust performance evaluation. Further investigation should also include real surveillance-camera data and continuous video streams, together with object-detection and localization capabilities, to assess the model under realistic surveillance conditions. These developments would provide a stronger basis for progressing from the present image-based proof-of-concept framework toward practical intelligent surveillance applications.

CONFLICT OF INTEREST

There is no conflict of interest with any person / institution in the article prepared

REFERENCES

- Ağdaş, M.T., Türkoğlu, M. and Gülseçen, S. (2021) “Deep Neural Networks Based on Transfer Learning Approaches to Classification of Gun and Knife Images,” 4(1). Available at: <https://doi.org/10.35377/saucis.04.01.891308>.
- Ahmed, S. *et al.* (2022) “Development and Optimization of Deep Learning Models for Weapon Detection in Surveillance Videos,” *Applied Sciences (Switzerland)*, 12(12). Available at: <https://doi.org/10.3390/app12125772>.
- Burgaz, M. and Budak, C. (2022) “Detection of Object (Weapons) with Deep Learning Algorithms from Images Obtained by Unmanned Aerial Vehicles,” 2, pp. 263–270. Available at: <https://doi.org/10.24012/dumf.1116534>.
- Dwivedi, N., Singh, D.K. and Kushwaha, D.S. (2019) “Weapon classification using deep convolutional neural network,” *2019 IEEE Conference on Information and Communication Technology, CICT 2019*, pp. 1–5. Available at: <https://doi.org/10.1109/CICT48419.2019.9066227>.
- Gao, M. *et al.* (2021) “A Transfer Residual Neural Network Based on ResNet-34 for Detection of Wood Knot Defects.”
- Gelana, F. and Yadav, A. (2019) *Firearm Detection from Surveillance Cameras Using Image Processing and Machine Learning Techniques, Advances in Intelligent Systems and Computing*. Springer Singapore. Available at: https://doi.org/10.1007/978-981-13-2414-7_3.
- Hashmi, T.S.S. *et al.* (2021) “Application of Deep Learning for Weapons Detection in Surveillance Videos,” *2021 International Conference on Digital Futures and Transformative Technologies, ICoDT2 2021* [Preprint]. Available at: <https://doi.org/10.1109/ICoDT252288.2021.9441523>.
- He, K. *et al.* (2016) “Deep residual learning for image recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December, pp. 770–778. Available at: <https://doi.org/10.1109/CVPR.2016.90>.
- Ingle, P.Y. and Kim, Y.G. (2022) “Real-Time Abnormal Object Detection for Video Surveillance in Smart Cities,” *Sensors*, 22(10). Available at: <https://doi.org/10.3390/s22103862>.
- Kambhatla, A. and R Khaled, A. (2023) “Real Time Deep Learning Weapon Detection Techniques for Mitigating Lone Wolf Attacks,” *International Journal of Artificial Intelligence & Applications*, 14(04), pp. 1–16. Available at: <https://doi.org/10.5121/ijaia.2023.14401>.
- Kaya, V., Tuncer, S. and Baran, A. (2021) “Detection and classification of different weapon types using deep learning,” *Applied Sciences (Switzerland)*, 11(16). Available at: <https://doi.org/10.3390/app11167535>.
- Lim, J. *et al.* (2019) “Gun detection in surveillance videos using deep neural networks,” *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2019*, (November), pp. 1998–2002. Available at: <https://doi.org/10.1109/APSIPAASC47483.2019.9023182>.
- Mehta, P., Kumar, A. and Bhattacharjee, S. (2020) “Fire and Gun Violence based Anomaly Detection System Using Deep Neural Networks,” *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, (Icesc), pp. 199–204. Available at: <https://doi.org/10.1109/ICESC48915.2020.9155625>.
- Oura, P., Junno, A. and Junno, J.A. (2021) “Deep learning in forensic gunshot wound interpretation—a proof-of-concept study,” *International Journal of Legal Medicine*, 135(5), pp. 2101–2106. Available at: <https://doi.org/10.1007/s00414-021-02566-3>.
- Salido, J., Lomas, V. and Ruiz-santaquiteria, J. (2021) “Automatic Handgun Detection with Deep Learning in Video Surveillance Images,” *Applied Sciences (Switzerland)*, 11(13).
- Verma, G.K. and Dhillon, A. (2017) “A Handheld Gun Detection using Faster R-CNN Deep Learning,” *ACM International Conference Proceeding Series*, pp. 84–88. Available at: <https://doi.org/10.1145/3154979.3154988>.
- Warsi, A. *et al.* (2019) “Gun Detection System Using Yolov3,” *2019 IEEE 6th International Conference on Smart Instrumentation, Measurement and Application, ICSIMA 2019*, (August),

- pp. 1–4. Available at: <https://doi.org/10.1109/ICSIMA47653.2019.9057329>.
- Xu, S. and Hung, K. (2020) “Development of an AI-based System for Automatic Detection and Recognition of Weapons in Surveillance Videos,” *ISCAIE 2020 - IEEE 10th Symposium on Computer Applications and Industrial Electronics*, pp. 48–52. Available at: <https://doi.org/10.1109/ISCAIE47305.2020.9108816>.
- Yadav, P., Gupta, N. and Sharma, P.K. (2023) “A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods,” *Expert Systems with Applications*, 212. Available at: <https://doi.org/10.1016/j.eswa.2022.118698>.