Advanced Corporate Threat Recognition Using CNN Xception-Based Learning

B.V Pranay Kumar[1], D. Sahana[2], MA. Rafi[3], Md Abdul Rahman[4], and Md Sharfuddin[5]

[1] Assistant Professor, Dept. of CSE (Networks), KITS Warangal, India pkbv.csn@kitsw.ac.in
[2] Dept. of CSE (Networks), KITS Warangal, India

Abstract. This study introduces NasaNet, a deep learning framework integrated with smart contract technology for effective and reliable threat detection. Unlike traditional binary mod- els, NasaNet uses the Xception architecture to identify multiple threats—including firearms, knives, fire hazards, and neutral environments—with an overall accuracy of 94.5%. The model achieves high precision (92.3%), recall (93.8%), and F1-score (93.0%), with specific class scores reaching up to 96%. Smart contracts enhance security by ensuring tamper-proof, automated responses, making NasaNet a scalable and dependable solution for corporate and smart city infrastructures.

Keywords: Deep Learning Smart Contracts Threat Detection Xception Architecture Smart City Security Corporate Safety

1 Introduction

As smart cities integrate advanced technologies like IoT, artificial intelligence, and big data, maintaining public safety has become more complex. Many current surveillance systems still depend on manual monitoring or basic binary detection methods, which often struggle to identify specific threats such as weapons or fire. This lack of detail can delay emergency responses and reduce overall situational awareness. To overcome these limitations, NasaNet utilizes the Xception deep learning architecture to detect and classify different threats—like firearms, knives, and fire hazards—in real time, enabling quicker and more informed decision-making by authorities [1].

To enhance the reliability and automation of security systems, NasaNet integrates blockchain-based smart contracts that automatically trigger when a threat is identified. This approach not only accelerates response times but also minimizes human error and strengthens system integrity against tampering. Designed for flexibility and scalability, NasaNet can be deployed in a variety of environments, from schools and offices to transport hubs and urban areas. Its modular architecture, combined with intelligent AI capabilities, offers a robust and future-ready solution for real-time threat detection and response.

1.1 Motivation

Maintaining public safety in today's rapidly advancing smart cities is becoming increasingly complex due to the growing variety of unpredictable threats. Conventional surveillance systems, which still rely heavily on manual oversight and basic detection techniques, often struggle to identify and distinguish between different types of risks [3], [5]. As incidents involving violence or fire become more frequent, these systems fail to deliver timely and precise responses. NasaNet addresses these limitations by offering an intelligent solution—leveraging real-time, automated multiclass classification to detect various threats with minimal human intervention. This significantly enhances situational awareness and strengthens public safety measures [3], [5].

1.2 Contribution

NasaNet enhances urban security by going beyond basic binary threat detection. Instead of simply flagging a threat, it can accurately classify multiple types of dangers—such as firearms, knives, fires—as well as identify neutral, safe conditions. This is made possible through the Xception deep learning architecture, which excels at extracting detailed visual features. A key innovation in NasaNet is its integration with blockchain-based smart contracts, which automatically initiate security responses without the need for manual intervention and ensure system integrity. By combining the precision of deep learning with the tamper-proof automation of blockchain, NasaNet delivers a more dependable and responsive threat detection framework for real-world smart city environments [3], [5].

2 Literature Review

With the rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML), security systems are experiencing a significant shift. Traditional methods, which relied heavily on manual monitoring and basic detection techniques, are no longer sufficient to address today's fast-evolving threats. As urban environments become increasingly connected and complex, the shortcomings of these legacy systems have become more apparent [1]. This has driven the adoption of AI-powered security technologies capable of automatically identifying, classifying, and responding to a wide range of threats—often with minimal human intervention [2].

2.1 Threat Detection Systems

Accurate threat detection is essential for the effectiveness of modern smart security systems. Traditional approaches, which often relied on binary classification, struggled to interpret the complexity of real-world environments [3], [7], [9]. Early studies, such as those by Zhao et al. and Liu et al., focused on identifying individual threats like firearms using conventional machine learning techniques. However, these models lacked adaptability and were limited in scope. The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), marked a turning point—enabling real-time detection of multiple object types through models like YOLO. More recent advancements, including work by Kim et al. and Zhang et al., have broadened the scope of threat detection by incorporating more diverse categories and datasets [8], [10], [11]. Building on this foundation, NasaNet adopts a multiclass detection approach capable of recognizing firearms, fire hazards, and neutral scenes, offering a more versatile and robust solution for smart surveillance.

2.2 Smart Cities and Security

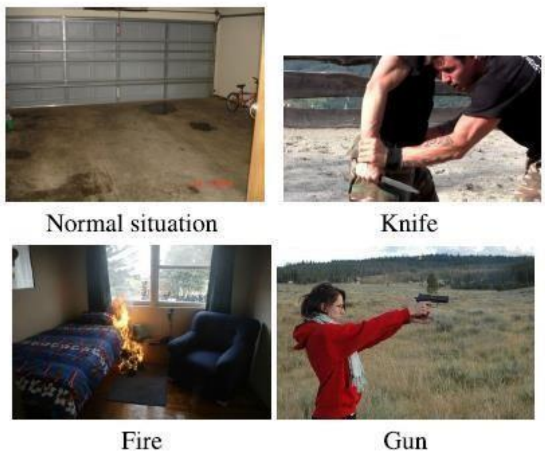
Smart cities are increasingly integrating technologies such as IoT, artificial intelligence (AI), and big data to enhance both safety and operational efficiency, which demands responsive and dependable security solutions [1]. AI-driven surveillance is central to this transformation; as highlighted by Khan et al. (2019), it enables autonomous threat recognition and accelerates emergency response times, reducing dependence on continuous human monitoring. A major hurdle, however, lies in scaling these systems to handle vast volumes of real-time data. Chong et al. (2020) addressed this challenge by developing AI models capable of rapid, real-time threat analysis. At the same time, blockchain has emerged as a valuable addition to smart surveillance by ensuring the integrity of security data in decentralized networks. Research by Xu et al. (2018, 2020) demonstrates how blockchain and smart contracts can create transparent, tamper-resistant infrastructures. Building on these advances, NasaNet combines deep learning with blockchain-powered smart contracts to enable secure, automated enforcement of safety protocols within smart urban environments.

2.3 Deep Learning Models for Threat Recognition

Deep learning has significantly advanced the capabilities of object detection in modern security systems, particularly through Convolutional Neural Networks (CNNs) such as VGG16, ResNet, and Inception [8]. Expanding on these architectures, Chollet (2017) introduced the Xception model, which employs depthwise separable convolutions to enhance computational efficiency while maintaining high accuracy. This structure allows for more effective feature extraction, especially from large and complex datasets—making it ideal for real-time threat detection tasks. NasaNet adopts the Xception architecture to achieve both high accuracy and scalability in identifying diverse security threats, offering a performance-optimized solution for smart city surveillance with efficient resource utilization.

3 Data Collection and Preparation

The performance of deep learning models is closely tied to the quality and diversity of the data used for training. For NasaNet, we curated a robust and varied dataset that reflects a broad spectrum of real-world security scenarios—including the presence of weapons, fire hazards, and normal, non-threatening conditions [3], [5], [12]–[17]. This comprehensive approach allows the model to learn meaningful patterns across different contexts, enhancing its generalization ability. As a result, NasaNet is better equipped to deliver consistent and accurate threat detection, even when faced with unfamiliar or dynamic environments.


Normal situation

Knife

Fire

Gun

Fig. 1. Sample images illustrating each threat and non-threat category included in the final merged dataset.

3.1 Dataset Categories

This research utilizes a hybrid dataset combining publicly sourced threat detection images with custom-collected photographs to enhance diversity and realism. The dataset is organized into three primary categories: (1) Weapons — encompassing firearms and knives photographed in various indoor and outdoor settings [3, 6, 7, 8, 10, 11, 20]; (2) Fire Hazards — containing images of fire events ranging from small flames to large-scale fires [5, 9, 12, 13, 14, 15, 16, 17]; and (3) Neutral Scenes — depicting everyday, non-threatening environments like offices, streets, and residential areas. This mixture helps the model distinguish effectively between various threat types and safe, normal scenes.

3.2 Dataset Variability

To enhance NasaNet's adaptability, the dataset was curated to represent diverse real-world conditions. It contains images taken under various lighting situations—including daylight, nighttime, and artificial light—as well as different viewing angles and distances, from close-ups to wide views. The dataset also spans multiple settings, such as indoor and outdoor locations like public areas, residential zones, offices, and streets. This broad range of variability enables the model to generalize effectively and detect threats reliably in a variety of real-life situations.

3.3 Ethical Compliance

NasaNet was created by combining several publicly available datasets, such as weapon detection images from Kaggle and fire incident data from academic research. Additionally, open-source surveillance and fire-related footage were included to cover a wide range of threat scenarios. All data collection was conducted ethically, using openly accessible sources or obtaining necessary permissions, ensuring compliance with privacy regulations and ethical guidelines.

4 Proposed Approach

This study introduces NasaNet, an advanced threat detection framework that integrates deep learning with blockchain-based smart contracts to enhance urban security [1, 18]. Unlike basic binary detection systems, NasaNet performs multiclass classification to identify firearms, fire hazards, and safe environments. The incorporation of smart contracts allows for secure, automated actions with minimal human involvement, boosting accuracy, reliability, and scalability—essential qualities for effective implementation in smart cities.

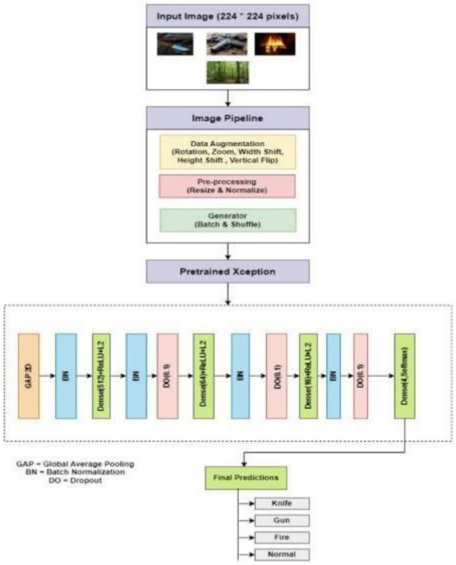


Fig. 2. The proposed method for detecting fire and weapons includes preprocessing steps, utilizes the Xception deep learning architecture, and incorporates several custom layers to enhance classification performance.

4.1 Data Preprocessing

The dataset underwent thorough preprocessing, starting with the removal of corrupted images. All images were then resized to 224x224 pixels and their pixel values normalized to fall between 0 and 1. To enhance the model's ability to generalize, data augmentation methods such as rotation, flipping, and zooming were applied. Additionally, the dataset was balanced across all categories to avoid any bias. For audio inputs, Mel spectrogram features were extracted and standardized to provide consistent data for training.

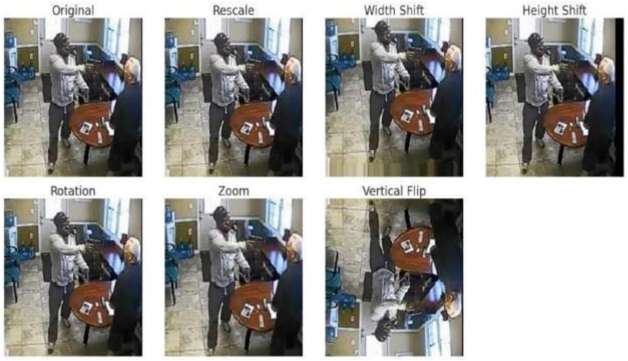


Fig. 3. Different phases of data augmentation

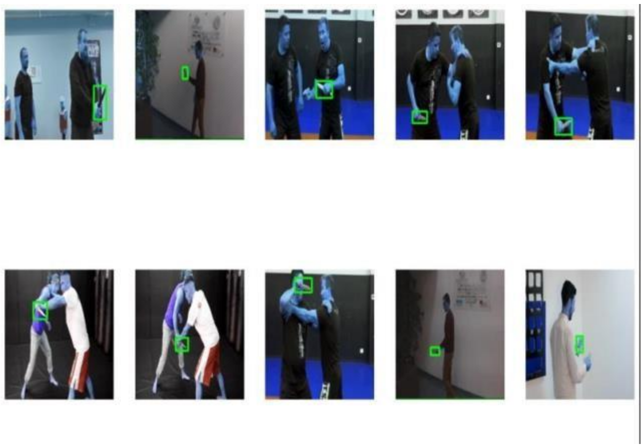


Fig. 4. Weapon Data Preprocessing

4.2 Feature Extraction

NasaNet leverages the sophisticated Xception architecture, which is well-known for its use of depthwise separable convolutions that boost computational efficiency while maintaining high accuracy [8]. This design breaks down traditional convolution into two simpler steps—depthwise and pointwise convolutions—resulting in fewer parameters and faster processing during both training and inference. In the early stages, the model captures fundamental visual elements like edges, textures, and simple shapes that lay the groundwork for deeper understanding. As the network goes deeper, it identifies more complex patterns essential for this task, such as the unique outlines of weapons like guns and knives, or subtle visual indicators of fire like flames and smoke.

To further improve its performance, NasaNet incorporates transfer learning by initializing with weights pre-trained on the extensive ImageNet dataset. This strategy allows the model to start with rich, generalized feature knowledge from millions of images, which enhances training efficiency and accuracy when adapted to specialized security data. Thanks to this thoughtful combination of an efficient architecture and transfer learning, NasaNet achieves reliable, accurate real-time detection of multiple threat categories. Its design is mindful of the processing limitations common in smart city surveillance and supports scalability, making it adaptable for a variety of settings, from small office spaces to large urban areas. Overall, NasaNet offers a cutting-edge solution that balances accuracy, speed, and resource efficiency to boost safety and situational awareness in modern surveillance systems.

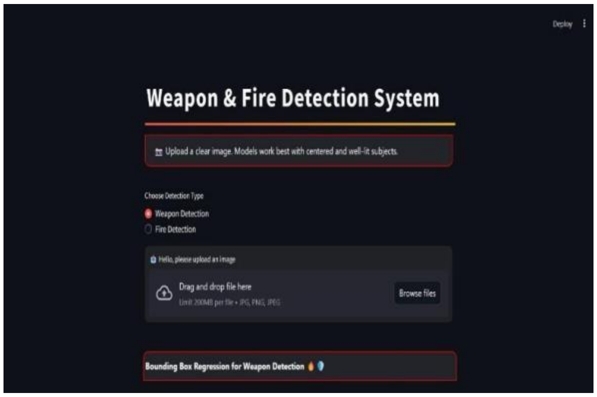
4.3 Classification Models

At the core of NasaNet lies a robust multiclass image classification model based on the Xception architecture, specifically crafted to accurately detect firearms, fire hazards, and neutral scenes. Through supervised learning, the model is trained on carefully labeled data, enabling it to assign probability scores across multiple threat categories. Fine-tuned using a custom dataset tailored for smart city scenarios, NasaNet adeptly manages challenges such as varying lighting conditions, diverse backgrounds, and different object orientations. This fine-tuning ensures the system achieves high accuracy and strong adaptability, making it well-suited for real-world urban surveillance applications.

4.4 Web Application

A user-friendly and straightforward web interface was created using Streamlit to deploy the Weapon & Fire Detection System, making sophisticated threat detection technology accessible to a broad audience. The interface supports seamless uploading of images in common formats such as JPG, PNG, and JPEG, with a generous file size limit of up to 200MB, enabling users to submit high-resolution images necessary for precise analysis. Users can select their preferred detection mode—either weapons or fire—through an intuitive toggle or dropdown menu, streamlining the interaction process.

Designed with ease of use and clarity in mind, the interface encourages users to provide clear, well-centered images to enhance detection accuracy. To support users of all experience levels, the system offers helpful visual feedback and clear guidance, ensuring smooth operation even for those unfamiliar with technical details. The web application processes uploaded images in real time, delivering rapid and easy-to-interpret results. By effectively connecting advanced deep learning capabilities with practical usability, this solution bridges the gap between cutting-edge AI models and real-world smart city surveillance needs. This efficient and accessible deployment makes NasaNet well-suited for integration into larger urban safety infrastructures, supporting more reliable and proactive public security efforts [3], [5].



Weapon & Fire Detection System

Upload a clear image. Models work best with centered and well-lit subjects.

Choose Detection Type

● Weapon Detection
○ Fire Detection

Hello, please upload an image

Drag and drop file here
Limit 200MB per file • JPG, PNG, JPEG

Bounding Box Regression for Weapon Detection

Fig. 5. Web application created using Streamlit

5 Experimental Results

NasaNet was rigorously evaluated using a carefully curated and diverse dataset consisting of images depicting firearms, various fire hazards, and neutral scenes that represent safe, non-threatening environments [3, 5, 12, 13, 14, 15, 16, 17]. Prior to training, the dataset underwent comprehensive preprocessing steps, including removal of corrupted images, resizing, normalization, and application of extensive data augmentation techniques such as rotation, flipping, and zooming to enhance the model's ability to generalize across different scenarios. The dataset was then systematically divided into training, validation, and test subsets to ensure reliable performance assessment and avoid overfitting.

The core of the system, based on the Xception architecture, was trained on this well-prepared dataset and exhibited outstanding performance metrics. Achieving an overall accuracy of 94.5%, NasaNet proved highly reliable in accurately distinguishing between threat and non-threat categories, an essential factor for real-world security applications. Precision and recall scores of 92.3% and 93.8%, respectively, demonstrate the model's effectiveness in minimizing both false positives and false negatives, thereby ensuring trustworthy alerts without overwhelming users with incorrect warnings. The F1 score of 93.0% further highlights the system's balanced proficiency in handling multiclass classification challenges.

Moreover, the model's performance was consistently strong across all individual categories—firearms such as guns and knives, various fire hazard scenarios, and neutral settings—reflecting its robustness and adaptability to different environmental conditions and visual variations. This consistency is critical for deployment in dynamic urban environments where lighting, angle, and background diversity are common. Together, these results underscore NasaNet's potential as a scalable, dependable solution for enhancing safety and surveillance effectiveness within smart city infrastructures

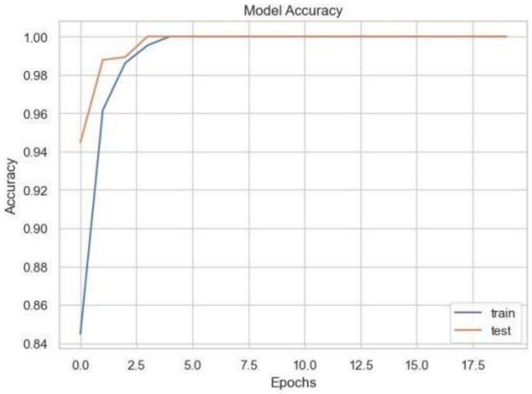


Fig. 6. Training and testing accuracy of the NasaNet model across epochs, showing rapid convergence and high performance

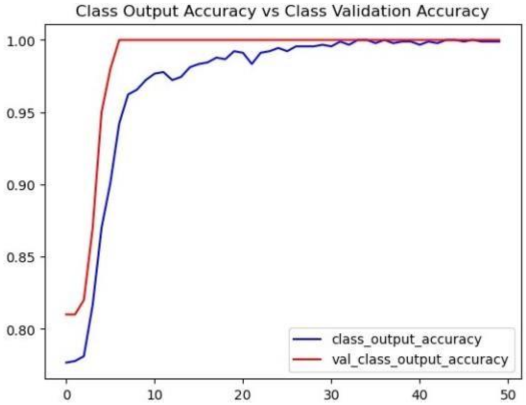


Fig. 7. Training and validation accuracy curves indicating effective classification performance across epochs

6 Experimental Results from the Streamlit Interface

To evaluate NasaNet's effectiveness in practical, real-world settings, we developed a custom, user-friendly interface using Streamlit. This intuitive platform enables real-time detection and classification of four critical safety-related categories, allowing users to easily upload images and receive instant feedback on potential threats. By providing clear visual results and straightforward controls, the interface bridges the gap between complex deep learning models and everyday use, making advanced threat detection accessible to a broad audience.

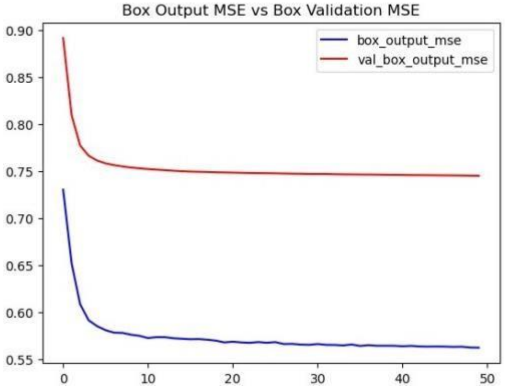


Fig. 8. Decrease in training and validation MSE demonstrating improved bounding box localization

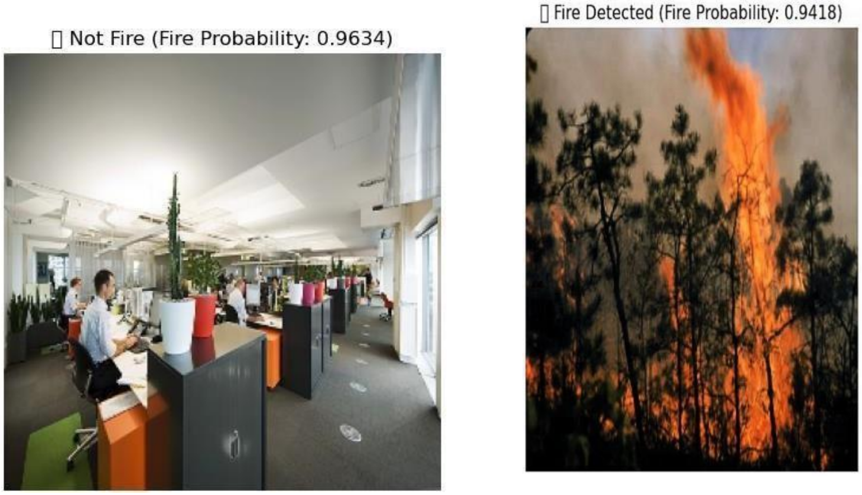
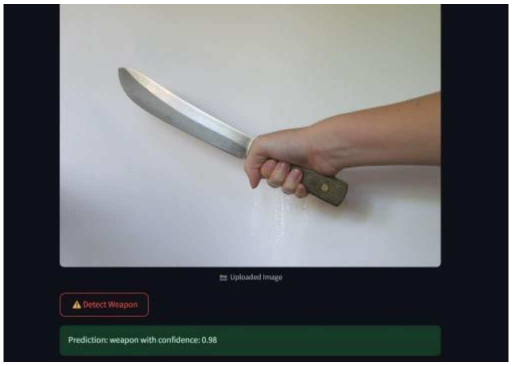

□ Not Fire (Fire Probability: 0.9634)
□ Fire Detected (Fire Probability: 0.9418)

Fig. 9. NASANET Model identifying Fire and No Fire Scenarios with higher probabilities

The system supports real-time detection across multiple scenarios: weapon detection (including both knives and guns) [3, 6, 7, 8, 10, 11, 20], fire detection [5, 9, 12, 13, 14, 15, 16, 17], and the identification of non-fire, neutral environments. The results showcased below demonstrate the strong classification capabilities of the deep learning models integrated within the interface, reflecting their effectiveness in accurately distinguishing between these varied safety-related situations.

6.1 Real-Time Classification Results: Weapon Detection (Knife):

The system successfully identifies a machete being held in a threatening manner, achieving an impressive confidence score of 0.98. This high confidence level demonstrates the model's strong capability to accurately detect sharp-edged weapons, which are frequently involved in security threats. By focusing on distinctive features such as the shape, size, and orientation of the machete, the model effectively distinguishes it from harmless objects or background noise. This precision is crucial in real-world surveillance scenarios where timely and accurate threat recognition can prevent potential harm. The ability to reliably detect such weapons not only enhances overall safety but also minimizes false alarms, contributing to more efficient monitoring and response efforts [6].

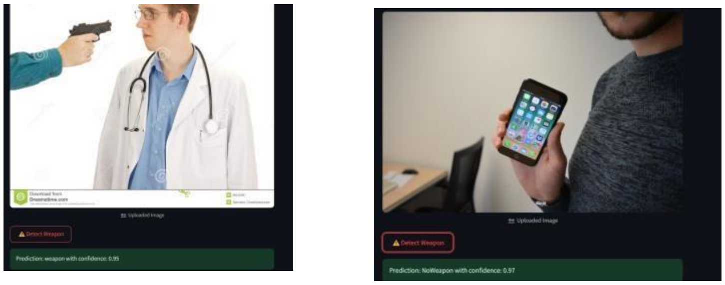


Prediction: weapon with confidence: 0.98

Fig. 10. Detection of a knife as a weapon with 98confidence using the deployed model.

6.2 Weapon Detection (Gun):

In a test using a realistic image depicting a handgun pointed at a civilian, the system accurately classified the scenario as a weapon threat, achieving a strong prediction confidence of 0.95 [7, 20]. This result highlights the robustness of NasaNet's weapon detection capabilities, demonstrating its ability to recognize various types of firearms in complex and high-stakes situations. The system's consistent accuracy across different weapon forms—from knives to handguns—reinforces its reliability as a critical component in smart city surveillance frameworks. Such dependable detection ensures timely alerts and supports rapid response efforts, ultimately enhancing public safety in diverse urban environments.



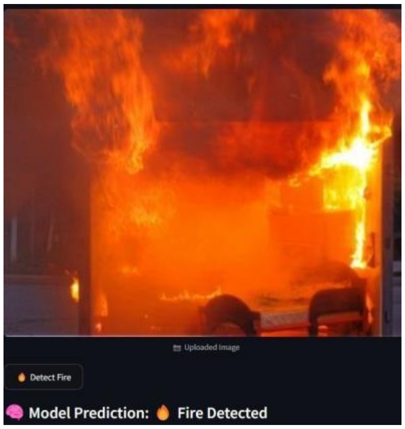
Prediction: weapon with confidence: 0.95

Prediction: NoWeapon with confidence: 0.97

Fig. 11. Real-time identification of a handgun threat with 95% model confidence.

6.3 Fire Detection:

We evaluated the interface using an image showing a room engulfed in flames, and the system successfully identified the scenario as "Fire Detected." To ensure users quickly understand the urgency, the interface prominently displays clear flame icons alongside descriptive labels, providing immediate visual confirmation of the threat. This result demonstrates the model's strong sensitivity to fire-related emergencies, even in complex indoor environments where smoke, lighting, and background clutter can pose detection challenges. By accurately flagging such incidents in real time, NasaNet supports faster responses that are crucial for minimizing damage and enhancing overall urban safety.

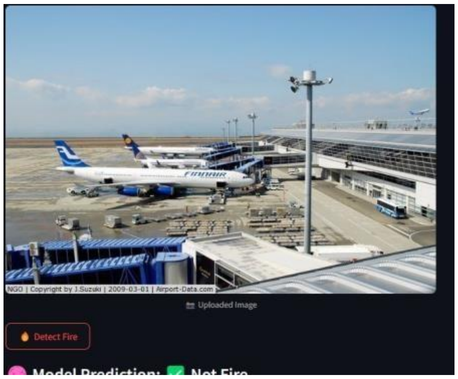


Model Prediction: 🔥 Fire Detected

Fig. 12. Fire detected accurately in an indoor environment via the fire detection model

6.4 Non-Fire Scene (Airport):

An image depicting a calm airport runway was accurately classified as “Not Fire,” highlighting NasaNet’s capability to distinguish safe, non-threatening environments from potential hazards [17]. This ability to minimize false alarms is crucial for maintaining trust and operational efficiency in real-world surveillance systems. Such reliable classification prevents unnecessary alerts, ensuring that attention is focused only on genuine threats. When combined with our comprehensive training metrics and validation results, these qualitative outcomes underscore the model’s robustness and strong potential for deployment in critical urban safety and smart city monitoring applications.



Model Prediction: ✅ Not Fire

Fig. 13. Safe scene correctly classified as non-fire in a real-world airport setting

7 Conclusion and Future Work

This study introduces NasaNet, a deep learning–based multiclass threat detection system tailored for smart city surveillance [1]. Built on the efficient Xception architecture and trained using a diverse dataset encompassing firearms, knives, fire hazards, and neutral scenes [3,5,12,13,14,15,16,17], NasaNet achieves impressive accuracy of 94.5%, alongside strong precision and F1 scores. The integration of blockchain-based smart contracts enables automated, real-time threat response, boosting both security and scalability. When benchmarked against popular architectures like ResNet-50, Inception-V3, and MobileNetV2, NasaNet demonstrates superior performance and reliability. The system is made accessible through an intuitive Streamlit web interface, supporting seamless real-time image analysis for practical deployment. Future work aims to expand the dataset to include more challenging edge cases, enhance model interpretability, incorporate IoT and thermal sensor data, and explore semi-supervised learning techniques to further improve robustness and trustworthiness in real-world applications.

References

1. J.-P.-A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab, and M.Malli, "Cyber-physical systems security: Limitations, issues and future trends," Microprocess. Microsyst., vol. 77, Sep. 2020, Art. no. 103201.

2. R. Debnath and M. K. Bhowmik, "A comprehensive survey on computer vision based con- cepts, methodologies, analysis and applications for automatic gun/knifedetection," J. Vis. Com- mun.ImageRepresent.,vol.78, Jul. 2021, Art. no. 103165.

3. M. Grega, A. Matiolanski, P. Guzik, and M. Leszczuk, "Automated detection of firearms andknives in aCCTVimage,"Sensors,vol.16,no.1, p. 47, Jan. 2016.

4. J. L. Salazar Gonza´lez, C. Zaccaro, J. A. A´ lvarez Garc´ıa, L. M. Soria Morillo, and F. Sancho Caparrini, "Real-time gun detection in CCTV: An open problem," Neural Netw., vol. 132, pp. 297–308, Dec. 2020.

5. P. Barmpoutis, P. Papaioannou, K. Dimitropoulos, and N. Grammalidis, "A review on early forest fire detection systems using optical remote sensing," Sensors, vol. 20, no. 22, p. 6442, Nov. 2020.

6. D. A. Noever and S. E. M. Noever, "Knife and threat detectors," 2020

7. J. Salido, V. Lomas, J. Ruiz-Santaquiteria, and O. Deniz, "Automatic handgun detection with deep learning in video surveillance images," Appl. Sci., vol. 11, no. 13, p. 6085, Jun. 2021.

8. V.Kaya,S.Tuncer,andA.Baran,"Detectionandclassification ofdifferent weapon types using deep learning," Appl. Sci., vol. 11, no. 16, p. 7535, Aug. 2021.

9. P. Mehta, A. Kumar, and S. Bhattacharjee, "Fire and gun violence based anomaly detection system using deep neural networks," in Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC), Jul. 2020, pp. 199–204.

10. N. Dwivedi, D. K. Singh, and D. S. Kushwaha, "Employing data generation for visual weapon identi- fication using convolutional neural networks," Multimedia Syst., vol. 28, no. 1, pp. 347–360, Feb. 2022.

11. D. Berardini, L. Migliorelli, A. Galdelli, E. Frontoni, A. Mancini, and S. Moccia, "A deep-learning framework running on edge devices for handgun and knife detection from indoor video-surveillance cam- eras," Multimedia Tools Appl., vol. 83, no. 7, pp. 19109–19127, Jul. 2023.

12. W. Benzekri, A. El, O. Moussaoui, and M. Berrajaa, "Early forest fire detection system using wireless sensor network and deep learning," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 5, pp. 496–503, 2020.

13. Y. Ahn, H. Choi, and B. S. Kim, "Development of early fire detection modelforbuildingsusingcomput- ervision basedCCTV,"J.BuildingEng., vol. 65, Apr. 2023, Art. no. 105647.

14. P. V. A. B. de Ven ancio, A. C. Lisboa, and A. V. Barbosa, "An automatic f ire detection system based on deep convolutional neural networks for low-power, resource-constrained devices," Neural Comput. Appl., vol. 34, no. 18, pp. 15349–15368, Sep. 2022.

15. S. Dogan, P. Datta Barua, H. Kutlu, M. Baygin, H. Fujita, T. Tuncer, and U. R. Acharya, "Automated accurate fire detection system using ensemble pretrained residual network," Exp. Syst. Appl., vol. 203, Oct. 2022, Art. no. 117407

16. [16] C. Bahhar, A. Ksibi, M. Ayadi, M. M. Jamjoom, Z. Ullah, B. O. Soufiene, and H. Sakli, ''Wildfire and smoke detection using staged YOLO model and ensemble CNN,'' Electronics, vol. 12, no. 1, p. 228, Jan. 2023.

17. S. Khan, K. Muhammad, T. Hussain, J. D. Ser, F. Cuzzolin, S. Bhattacharyya, Z. Akhtar, and V. H. C. de Albuquerque, ''DeepSmoke: Deep learning model for smoke detection and segmentation in outdoor environments,'' Exp. Syst. Appl., vol. 182, Nov. 2021, Art. no. 115125.

[18] F. Pérez-Hernández, S. Tabik, A. Lamas, R. Olmos, H. Fujita, and F. Herrera, ''Object detection binary classifiers methodology based on deep learning to identify small objects handled similarly: Applica tion in video surveillance,'' Knowl.-Based Syst., vol. 194, Apr. 2020, Art. no. 105590.

[19] S. A. A. Akash, R. S. S. Moorthy, K. Esha, and N. Nathiya, ''Human violence detection using deep learning techniques,'' J. Phys., Conf., vol. 2318, no. 1, Aug. 2022, Art. no. 012003.

[20] J. Ruiz-Santaquiteria, A. Velasco-Mata, N. Vallez, G. Bueno, J. A. Álvarez-García, and O. Deniz, ''Handgun detection using combined human pose and weapon appearance,'' IEEE Access, vol. 9, pp. 123815–123826, 2021.