Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/40101

Optimized ML Model with Explainable AI for Threat Detection at Kuwait International Airport

Abdallah S. Mohamed[1], *Adel A. Sewisy*[1], *Khaled F. Hussain*[1] *and Ahmed I. Taloba*[2,*]

[1] Department of Computer Science, Faculty of Computers and Information, Assiut University, Assiut, Egypt
[2] Information System Department, Faculty of Computers and Information, Assiut University, Assiut, Egypt

Received: 30 Jul. 2024, Revised: 15 Sep. 2024, Accepted: 23 Nov. 2024
Published online: 1 Jan. 2025

Abstract: Airport security measures have grown in terms of importance, especially in relation to protection of passengers and airport employees, in addition to looking for efficient ways and means of threat identification. Currently, conventional X-ray image processing methods can generate numerous challenges when it comes to detection of concealed threats like weapons and explosives, and other prohibited items likely to compromise security. Threat identification using X-Ray has been a very important factor in the security aspect of airports worldwide including Kuwait International Airport, this study seeks to design a CNN-GRU hybrid model integrated with Firefly Optimization (FO) and Explainable AI (XAI) to enhance threat recognition accuracy in X-ray images among the passengers passing through Kuwait International Airport. Thus, the objective in threat modeling is to improve threat identification while maintaining the system comprehensible to security personnel. The proposed research is innovative because it for the first time presents a framework utilizing CNN to extract spatial features from the images, GRU for modeling temporal dependencies into the images featuring Firefly Optimization for hyperparameter tuning and SHAP for explainability. This strategy proves both high detection rates and, from the operator's perspective, observers' standpoints, providing the necessary transparency and, therefore, trust to the operating AI models. Moreover, the integration of SHAP-based Explainable AI helps the model to analyse and explain areas of the X-ray images that inform detection of possible threats. Both of these elements, performance and transparency, are important to make the system reliable for its everyday practical use in airports. The experimental results demonstrate that the proposed CNN-GRU-FO-XAI yields accuracy of 99%, which is higher compared with other models. It also shows the increased precision, recall, and F1-score leaving no doubt that the model has high outcomes and only a few false positives. The SHAP analysis also indicates the most important areas in the X-ray images concerning the threat detection as well as offers more openness and interpretability. This results in almost perfect detection performance and also provides the explanations of the decisions made by the model which is beneficial to increase trust in AI-driven systems and make use of this tool in high-risk environments as airports.

Keywords: Airport Security, X-ray Image Analysis, Firefly Optimization, Kwait International Airport, SHAP (Shapley Additive Explanations), Gated Recurrent Unit (GRU), Threat Detection.

1 Introduction

Security is paramount in the daily operations of the global transport systems particularly airports whom are leading entry points in international transport of people and goods [1]. Due to the large number of people and baggage passing through airports, the problem of protecting passengers' lives while maintaining operating efficiency has always been one of the m1.ain concerns of security agencies [2]. The current common airport security measures involve action by personnel in checking bags and use of detectors which suffer some drawbacks in detection of hidden and State-of-art threats [3]. It is here that sophisticated technologies to include X-ray imaging systems have thus emerged useful in scanning the luggage and the passengers for worrysome items including weapons, explosives, drugs and other prohibited items [4]. Analysis of radiographic images is an essential part of contemporary security measures in airports since these systems offer detailed images of the luggage contents, to give an opportunity to officers to examine and singular dangerous items [5]. But, in large numbers of luggage and travellers perusing the security check points each day proved to be a challenge to human involvement as they are likely to overlook certain concealed or even new developed forms of threats [6]. Consequently,

* Corresponding author e-mail: Taloba@aun.edu.eg

the use of automatically accurate image analysing systems that is capable of real-time X-ray imaging that has the ability of detecting possible threats is becoming imperative [7].

Even with advancements in X-ray image technologies, detecting hidden threats in X-ray pictures is still very much a challenge. Conventional processing of X-ray images has limitations in detection of complex shapes. It cannot distinguish an item being harmless or suspicious [8]. Most often these systems employ static and rigid rule-based algorithm implementations or simple machine learning profiles that are not strong enough to deal with the large spectrum of threats that may be hidden in different types of items, bags or containers [9]. More so, these models also do not seem to generalise over varying situations like the same lighting conditions, different kinds of luggage and the presence of a complicated texture or overlapping objects [10]. Manual inspections by security personnel are often subjected to human errors and fatigue from having to screen thousands of bags in a small time. Hence most traditional systems do not have that sort of accuracy and efficacy that is needed for effective threat detection [11]. This demonstrates how urgently new, automated solutions need to be developed for accurate threat detection using X-ray background imaging [12].

The intent of this research is to represent a substance hybrid CNN-GRU model that takes advantage of both Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) for advanced X-ray image processing. This model will also be optimized using Firefly Optimization (FO) and Explainable Artificial Intelligence (XAI) techniques, particularly SHAP (SHapley Additive exPlanations), for hyperparameter tuning and providing insights into model decision making. The intended outcome of the CNN-GRU fusion model is to achieve better results through the CNN module for spatial features (edges, textures, shapes) and the GRU for any sequential patterns (object movement, changes over time). Firefly Optimization will thus be applied to fine-tuning the hyperparameters of the CNN-GRU model for optimal performance. Meanwhile, XAI methods like SHAP will enhance the interpretability of the model enabling security personnel to understand and trust the model's decision which is very critical in environments like airports. This study thus aims to maximize the use of advanced AI techniques to raise the bar as far as threat detection is concerned across X-ray images, efficiency with reliability and transparency in airport security. The relevance of this research base is pinpointed primarily in its capacity to bring a breakthrough in utilizing contemporary IT solutions in airport security and threat identification.

Using such techniques as the hybrid CNN-GRU model, this study brings the possibility of extending the analysis of X-ray images with high levels of accuracy as compared to conventional and categorical approaches. In hyperparameter tuning application, Firefly Optimization is used for model optimization to enhance its performance while Explainable AI gives transparency of the model to the operators. This approach can cut the time it takes to flag threats and at the same time increase the accuracy of threat detection so as to minimize on false positives or false negatives. As the pressure on facilities to deploy higher levels of security in airports rises, this research shows how utilising knowledge from the extended field of AI is not only beneficial in increasing activity rates in airports, but importantly is a substantial step to improving passenger's safety. Additionally, the are solution based that addresses another important question regarding the use of AI in security context: the ability to have confidence in the automated decision- making process to have confidence in AI- based decisions that are made. This can result in enhancing the trust within the operator to the AI security systems such that the human can rely more on machines during critical decision-making processes. Finally, the findings could lead to the utilization of AI solutions in the airport security systems and other security risk sectors where accuracy, transparency and high efficiency is desirable. The key contribution for this study was outlined below:

1. A novel hybrid architecture of concrete CNN-GRU is presented for threat detection in X-ray images by taking benefit of spatial feature extraction and temporal dependency of the identified threats across multiple frames in an accurate way.
2. The model's hyperparameters are optimized using Firefly Optimization for better performance to make effective and accurate predictions while making threat detection in real time.
3. This study also adds transparency to the model's decision-making through incorporation of SHAP (Shapley Additive Explanations), such that security personnel may, through training, learn about the features and areas that steer detection of threats, thus earning trust and enabling human verification.
4. This suggested CNN-GRU model optimized with Firefly and enhanced with XAI has outperformed all existing models and proved its effectiveness in detecting different threats in X-ray images with better metrics.
5. This work provides a model specifically built for airport security with implementation implications for improving threat detection, efficiency of operations, and safeguarding passengers in high-risk environments.

2 Literature Review

Wang et al. [13] presents an AATR approach to counter the shortcoming of the traditional threat recognition system by focusing on material signatures. Through the combination of adaptive machine learning, a 3D multi-scale computed tomography image segmentation technique is incorporated with a multiclass support vector machine classifier. This

approach enables the system to focus on the new and emerging threats which can recognize materials and features not featured in the training dataset. The results when compared in controlled experimental conditions prove to be above 90% with false alarm rate below 20% depending on varying detection rates even on unknown material. Overall, this study highlights the importance of using flexible machine learning in improving on the performance of screening devices in dynamic environments.

Zaliskyi et al. [14] improve X-ray baggage screening by addressing a main challenge of high false alarm rates in existing systems. Using geometric approach and Beer-Lambert low a shadow image processing is designed for identification of simple object images. These images are then used to identify even more complicated objects including weapons such as firearms. By adapting spectral analysis for the detection of shadow images, the proposed method enhances the detection accuracies of X-ray systems. The outcomes of the study show that such strategy can effectively contribute to the development of algorithmic solutions for aviation security, improving the mechanisms of automated screening of baggage prohibited and dangerous items.

In order to overcome the limitations of detecting mostly occluded items and actual X-ray baggage scans, (Sara and Mandava [15] suggests a Modified Encoder-Decoder model. The model also includes a Preprocessing step to reduce inherent noise in X-Ray images by enhancing image qualities with the help of a Poisson Noise Reduction technique. The obtained images are further refined for easier segmentation to define the potential threats. In the SIXray and GDXray datasets, extensive assessments show that the presented model performs better, and obtains mAP, IoU, and DC values that equal 97.32%, 73.14%, and 85.12% respectively. This study also achieves relatively high accuracy of 99.17% in sixteen testing samples of SIXray real datasets, thereby demonstrating the reliability and efficiency of the approach in solving current threats to aviation security.

Alansari et al.[16] presents the Multi-Scale Hierarchical Transformer for X-ray Detection (MHT-X) that solves the problem of existing deep learning methods used in baggage screening by providing the overall view of threats. MHT-X is the integration of multi-scale contour mapping and hierarchical feature extraction using visual Transformers. The enhancement of concealed and obstructed threat objects relates to the formulation of precise contour maps and hierarchical spatial attributes. The performance is validated on two datasets, CLCXray and PIDray, and yields mAP of 65.26% and 78.19% respectively; it outperforms existing models. This study suggests that improved feature extraction methods have ways of improving the machinery based automated detection systems in aviation security.

Gota et al.[17] discuss the process of constructing the threat object recognition system that uses convolutional neural networks in the context of X-ray images analysis. This study employs OpenCV and Keras with TensorFlow, developing and testing CNN models for identifying prohibited items in hand luggage. To improve the models' performance and flexibility, the study uses some data augmentation techniques. The experiment shows enhanced detection performance; this highlights the efficiency of CNNs and today's advanced modern machine learning in reducing reliance on manual scrutiny and enhancing the precision of discernment of threat objects in aviation security.

The present literature on baggage screening and threat detection by using X-ray in aviation security presents a range of measures meant to increase accuracy and decrease false alarms. COTS X-ray systems although currently dominating the field have drawbacks such as; higher false positive rate and thus demanding sophisticated image processing and recognition. In Being in business, some of the challenges mentioned above have been or have been addressed by various studies through the following solutions. For example, Wang and collaborators proposed the Adaptive Automatic Threat Recognition (AATR) with enhanced capabilities of dealing with different threat signatures compared to previous versions, with high detection probabilities and low false alarms. Zaliskiyi et al. (2022) proposed the spectral analysis-based shadow image processing method to improve baggage object recognition. Because denoising is an issue when using X-ray images, Sara and Mandava (2024) proposed the Modified Encoder-Decoder model to incorporate Poisson Noise Reduction for enhanced segmentation and threat detection. Moreover, Alansari et al. (2024) proposed the Multi-Scale Hierarchical Transformer-X (MHT-X), which employs multiple scale contour map and hierarchical features extraction, obtaining better threat detection performance against the conventional CNN based models. Gota et al. (2020) explored the possibilities of CNN integration with the machine learning algorithms such as OpenCV and Keras in baggage recognition asserting that with the increasing focus toward the application of AI and Deep learning in threat detection the use and importance of AI cannot be overstated. Collectively, these constitute a progression in the continuing process of improving the performance of baggage screening systems in aviation security.

3 Research Gap

With threats getting more complex in aviation security, the advantages of implementing the manual and automated forms of baggage screening show the necessity of the multi-sensor, accurate and easily interpretable fusion system [18]. Traditional approaches have issues with changing threat signatures, high false positives and low explainability of decision making logic which set the need for a more effective and transparent solution for concealed threat identification on X-ray CT baggage scans [19].Therefore, the study proposed solution harnesses a three part CNN-GRU solution that is

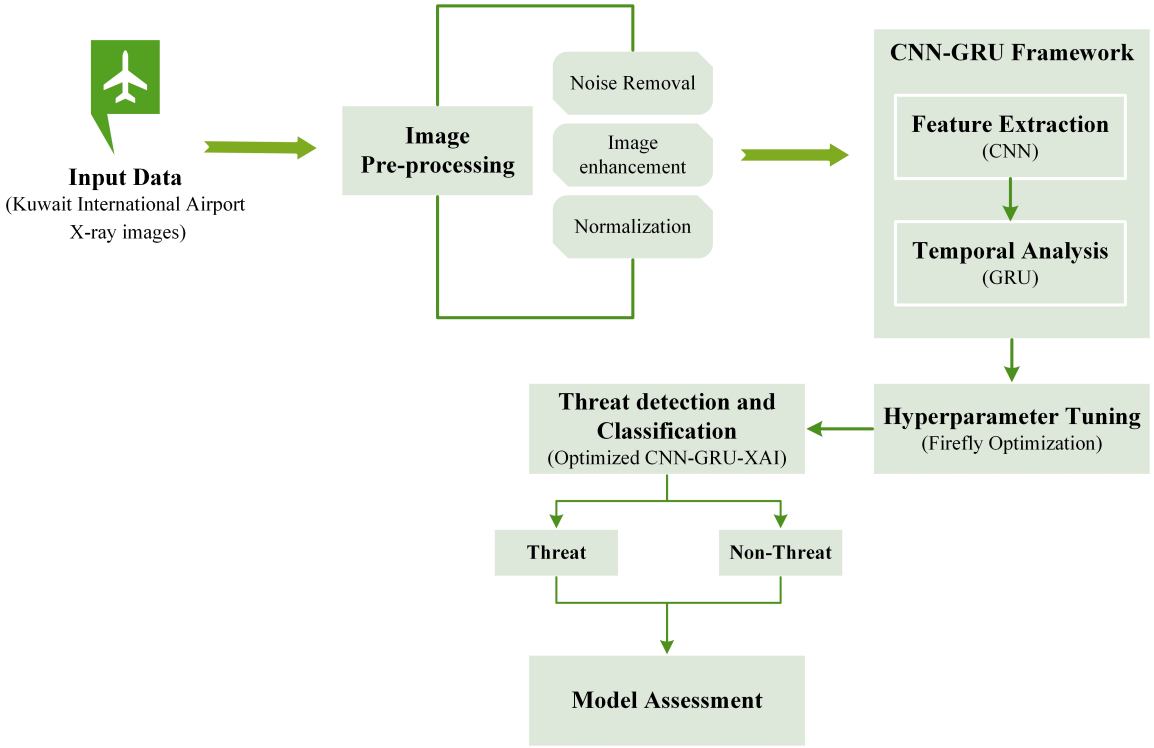

Fig. 1: Workflow of the suggested approach.

Firefly optimised to detect concealing threats in baggage scans acquired from X-ray CT for improved operational efficiency and to engender end-user XAI has been integrated into the framework to ensure that the insights derived are easily explainable and can be trusted while being used by security personnel. This solution also meets the dynamic threats' requirement through the application of adaptive machine learning strategies and high-quality preprocessing, feature selection and temporal analysis for higher accuracy and practical usability.

4 Research Mechanism

The proposed method starts by cleaning and preparing the data that will be used to extract features from the X-ray images as shown in figure 1. These steps include noise removal by employing filters such as Gaussian or media filtering for eradicating unimportant distortions. The next step is normalisation to allow pixel intensity values to have consistent values thus enhancing the convergence of the model. Enhancement techniques are then applied to the image to enhance features such as the contrast and the sharpness of the image which are special areas that are easily spotted when enhancing the threat. Besides, rotation, flipping and scaling are also applied to the images in order to increase the variability of the dataset, which is important for model stability. Subsequently, a combined convolutional neural network and gated recurrent unit model is used for feature extraction and temporal analysis. The CNN extracts spatial characteristics from X-ray images, and the GRU recognizes the temporal patterns in multiple frames to detect threats. Here instead Firefly Optimization is used to adjust the hyperparameters of the model to the best values. Finally, the proposed Explainable AI (XAI) methods such as SHAP are implemented to enable the security staff to comprehend the working of the model. The stages of pre-processing, deep learning, optimisation and explainability provide high accuracy of threat detection and high transparency, making the methodology suitable for real time applications in airport security.

4.1 Data Collection

As for the improvement of airport security by using AI methods, the X-ray image datasets must be real, which can be collected from public domain or can be synthesized. The datasets accessible to the public with the help of academic

Fig. 2: Hi-SCAN 6040 CTiX computed tomography (CT) X-rays Machine.

institutions or security organizations contain X-ray images of different objects, including possible threats, labelled, and can be used as a basis for training and checking machine learning models. Additional data from synthesis will improve the dataset and allow for expanding the range of threats and their conditions with simulated baggage conditions. This situation may be supplemented by cooperation with security equipment suppliers, who have ready-to-use datasets adapted for the formation of algorithms. These approaches provide quality data for threat models and AI without use of direct access which remains operationally and ethically desirable [20].

This study uses the HI-SCAN 6040 CTiX as the screening technology; it is a Computed Tomography (CT) X-ray system accredited by TSA AT-2, CPSS, ECAC, and STAC EDS CB C3 was illustrated in figure 2. These systems provide high rates of detection of explosives when at the same time helping reduce the time taken for screening of passengers by permitting electronics and Liquids to be taken inside the aircraft cabin in the carry-on bags. The rotating gantry takes hundreds of images to build 3D images of what is in the bag in real time, thereby providing accurate identification of threats. In the current study, data is collected from the seventy HI-SCAN 6040 CTiX units installed at Kuwait International Airport Terminal 2; Smiths Detection is installing the entire range of baggage screening equipment at the terminal. Reliability of this large dataset can facilitate understanding of adaptive approach to threat identification and its implementation in high pressure terminal infrastructure.

4.1.1 Preprocessing Techniques to Enhance Image Quality

The main preprocessing procedures on the X-ray images concern a significant stage of data preparation for AI threat detection models. The idea is thus aimed at optimizing image quality, filter out noise as well as set up to enhance the quality of the data for accurate and efficient training of the model. Below is detailed preprocessing techniques commonly applied to X-ray images:

Gaussian Filtering for Noise Removal in Images

This is an important input preprocessing operation which enhances quality and reliability of threat detection X-ray images used in AI models. Artifact in the X-ray images can stem from the actual scanner used, electrical or electromagnetic interference, or climatic conditions which may overshadow one or more of the features required for analysis. A Gaussian filter is a popular form of noise reduction where the image is smoothes by blurring the image with Gaussian distribution and suppressing high frequency without affecting the edges of the image.

In diagnosing conditions through analysis of images, beginning preprocessing of X-ray images entails careful examination of the input data to help achieve maximum noise reduction and image enhancement. This starts with assessment of data description of the images, their size, format, noise level and quality in order to determine the level of filtering required. Further, noise analysis is done to identify the type of noise that contaminated the images including Gaussian noise, salt-and-pepper noise and periodic noise. This can be done either statistically; for example, by computing variance and standard deviation and investigate if they show signs of distortion, or by performing a visual inspection. These insights

inform the choice of the specific preprocessing steps applicable under condition that noise will be minimized with relevant image features preserved. A Gaussian filter smooths pixel intensity values by weighting neighbouring pixel contributions using a Gaussian function Eqn. (1):

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

Here:

1. (x,y): Pixel coordinates relative to the center of the kernel.
2. σ: Standard deviation of the Gaussian distribution, controlling the degree of smoothing.
3. **Kernel Size**: Determines the area of neighbouring pixels considered for each calculation, typically chosen as odd numbers.

Image Enhancement

X-ray image enhancement techniques maximize feature visibility, thereby enhancing interpretation of images and analysis. For example, a contrast stretch will be applied to pix at varying quality X-ray scans data so that the pixel value intensity range may be modified to cover the entire dynamic range, thus making the small detail more distinguishable. Histogram equalization, used to make higher enhancements, ultimately takes away the redistributing intensity histogram from the previously hidden parts of the images, giving a better overall clarity. The Area of fine details with uniform intensity becomes less oversaturated as Adaptive Histogram Equalization (AHE) localizes the contrast improvement onto much smaller areas instead of smaller regions within the image. By employing a nonlinear transformation, gamma correction adjusts the brightness of an image with the following formula: $I_{out} = I_{in}^{\gamma}$,, where gamma is a specific factor of correction for each dataset used to modify brightness and contrast. Then systematically applying all these modifications across the entire dataset should guarantee better visibility of features and acceptable performance in terms of threat detection without fail in security applications.

Normalization

Normalization is the most crucial preprocessing for making X-ray images consistent in the database. Normalization normalizes pixel values into a standard range, improving accuracy and efficiency during later image analysis and model training. For large data sets containing images under different scales relating to resolution, brightness, and contrast, the data can be compared, standardized, and normalized. One of the most preferred methods is Min-Max normalization, which rescales pixel intensities in a specific range, such as [0, 1] or [-1, 1]. The following is the formula used in Eqn. (2):

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2}$$

This transformation ensures that all pixel values across the dataset are commensurately altered, irrespective of their actual intensity range, leading to a uniform representation of image data. In a large database of X-ray images, this will ensure that all images are showing similar pixel value disbursements irrespective of the type of scanner or method of acquisition. Hence, normalization would enhance accurate and stable training of the model, improved detections of relevant features and performance within image recognition algorithms in applications at the airport for security purposes.

4.2 Model Architecture for Threat Detection Using CNN-GRU with Firefly Optimization

The architecture model used for detecting threats in X-ray images is the CNN-GRU framework enhanced with Firefly Optimization for hyperparameter tuning as shown in figure 3. Data are processed and thoroughly analysed to cater to complex image data. Each of its broad modules in the architecture will play an important role in feature extraction while at the same time ensuring the model retains a high performance.

4.2.1 CNN Module: Extracting Spatial Features from X-ray Images

Convolutional Neural Network, popularly known as CNN, serves as the model's primary component for feature extraction. It captures spatial hierarchies and patterns in the image data automatically and efficiently. In using X-ray images, it acts to detect those low-level features (like edges, textures, and corners) and then abstracts them progressively into higher-level representations, such as shapes or objects that may indicate possible threats. The CNN architecture ideally consists of several of its major components:

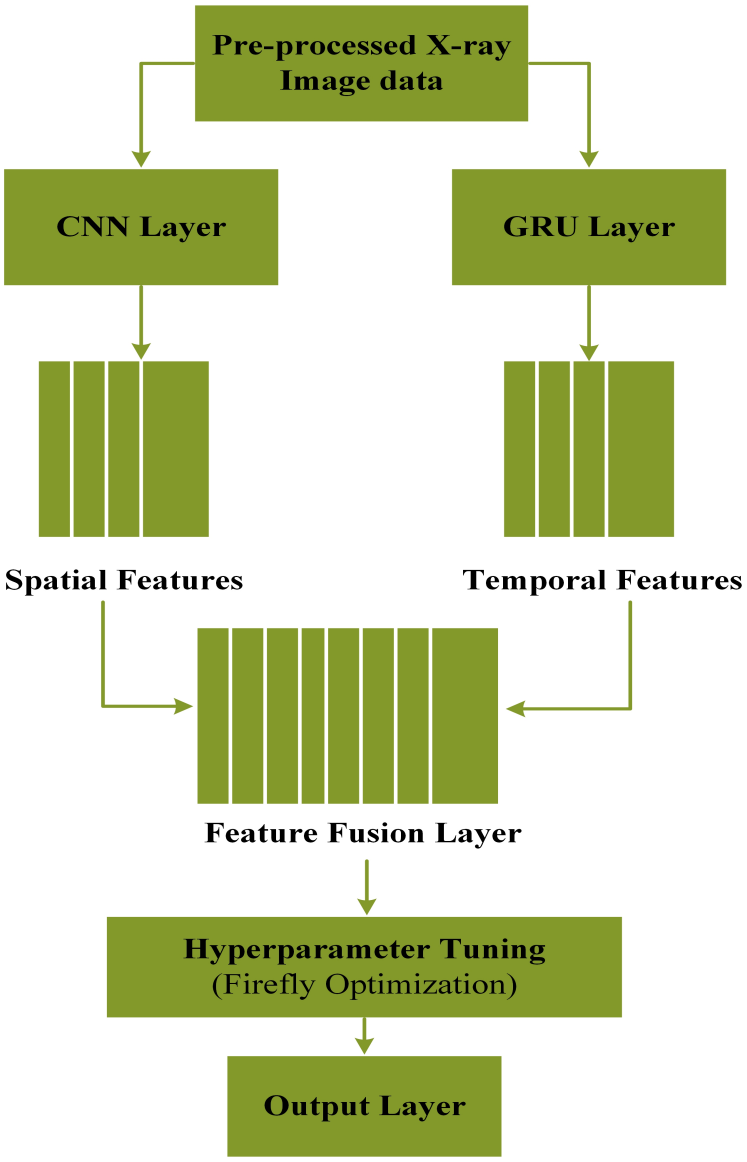

Fig. 3: CNN-GRU Framework.

Convolutional Layers: Filters are applied to the input image in these layers to extract spatial features. Each filter detects specific kinds of features such as edges, patterns, and textures, which are important for identifying a potential threat in X-ray images. The operation for a single filter at a position (x,y) on an image I is defined as:

$$S(x,y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(x+i,\ y+j) \bullet K(i,j) \tag{3}$$

Activation Functions: After convolutional layers, the activation function is mostly given to ReLU (that is, Rectified Linear Unit), enabling the model to learn non-linear patterns and, thereby, complex patterns.

$$ReLU\,(x) = max(0,x) \tag{4}$$

Pooling Layers: Max-pooling or average-pooling layers down sample the feature maps, and thus reduce their dimensionalities, but keep their important elaborated spatial features. This step aids the improvement of generality in the model along with making it computationally efficient.

Fully Connected Layers: CNN for high-level feature representation after several convolutional and pooling layers. These features are compressed to a single dimension and passed through fully connected layers, which work to combine them in an output decision. All neurons in the already established layer are connected to all neurons in the fully connected layer. Further sequential analysis will be performed on the fully connected output with the GRU module. Mathematically, this operation can be represented as:

$$y = W \bullet x + b \tag{5}$$

The CNN module is essential in the threat detection model because it extracts spatial features from the X-ray images taken at Kuwait International Airport. To start with, there are convolutional layers which due to filters on the input image look for basic features which include edges and texture. After each convolution, the activation functions like ReLU are used which introduces non linearity in the network which helps the network to learn complex patterns. Convolutional and pooling layers for example, max pooling, or average pooling help decrease the dimensionality of feature maps while retaining critical information. Last but not the least, it consists of fully connected layers which integrates the features extracted at various levels and make high level decision, which are then fed to GRU for sequence analysis. This method is chosen because it is fast at handling image data and isolates spatial characteristics of images, which are important in identifying threats in real environment such as an airport. The advantages of the CNNs are found in its success of image proceeding tasks, its capability to learn the hierarchical features of the raw data and in its capability to process high volumes of X-Ray images which is crucial for real time threat detections. Thus, with the help of our modification to the model, introducing CNNs with the GRU and embedding the Firefly Optimization module, the required reliable and accurate threat detection can be achieved that shall improve security at the Kuwait International Airport.

4.2.2 GRU Module: Capturing Sequential Dependencies

The Gated Recurrent Unit module is also responsible to capture temporal dependencies from extracted spatial features from X-ray images, especially when dealing with a large dataset from Kuwait International Airport. CNNs are particularly good at detecting more specific features such as edges, textures and patterns in X-ray images for example, but they do not understand the importance or relationships of these features at different times or in different sections of an image of interest by design. The GRU module plays a role here:
In the case of X-ray images particularly when it comes to security aspects like the Kuwait International Airport, the threats might be seen in patterns that are different in successive frames or sections of the body, position of the object or even the interaction between different objects in a bag etc. For example, a threat such as a concealed weapon or am explosive object may not be easy to discern by isolated features but showed up when the features or sections of objects are examined one after another. The GRU module aids in the process through the use of 'memory' component which holds on to past inputs (features) and in a chronological method, tries to find out such relations. Key components of the GRU module include:

1. **Gates:** GRUs use gates of update and reset, which control the flow of information. The update gate determines how much of the previous hidden state is to be forwarded to the next time step, while the reset gate decides how much of the new input should be allowed to influence the hidden state. This allows the GRU to effectively balance keeping information and incorporating new inputs, thus ensuring the model's attentiveness to relevant sequential dependencies.
2. **Hidden State Updates:** At all time intervals, the GRU functions to update the hidden state using the current input module and the previous hidden state. It leaves the GRU retraining to retain context over time, thus enabling the threats to be detected over time as they unfold and become more visible on other sections of images or within frames.
3. **Sequential Analysis:** The GRU is meant for processing data that has been organized in sequential order, so that it can learn some temporal dependencies related to the features of images. In X-ray images, such features mean that the learning captured by the GRU can progress through multiple sections or frames in order to learn and understand the complex patterns that are defined by threats, even if those patterns are not visible in a single frame or section of the image.

The use of Gated Recurrent Units (GRU) for this study was guided by their efficiency in handling sequential data, which is crucial for the interpretation of X-ray images regarding threats. One of the big advantages of using GRUs is their ability to learn temporal dependencies between image features that may lie between frames. In such applications as airport security, threats may transform or manifest dissimilarly in the different parts of an X-ray image whereas a relationship between objects and their features over time is necessary. GRUs are less complex computationally than other Recurrent Neural Network models, they require several less parameters and therefore can be implemented in real time processing applications such as in Kuwait International Airport. Furthermore, compared to other types of RNNs, GRUs overcome the problem of gradient vanishing which helps the model of keeping essential contextual information most relevant to the data sequence. This enables them provide better identification of concealed objects like weapons by studying the interaction of things in the whole scan. Finally, it is crucial to point out that the GRU module contributes to improving the predictive performance when identifying threats in X-ray images because it introduces temporal dependencies to the sequences.

4.2.3 Firefly Optimization: Tuning Hyperparameters for Optimal Performance

Firefly Optimization (FO) is a meta-heuristic algorithm inspired by the flashing behavior of fireflies, which has successfully optimized a complex generic model such as the CNN-GRU system for X-ray image threat detection. In this context, the purpose of Firefly Optimization is to optimize and tune the hyperparameters of the CNN-GRU model to improve performance, accuracy, and efficiency. The methodology of the algorithm searches iteratively for the optimal set of hyperparameters that maximize the model's ability to classify and detect threats.

Initialization: The initiation process begins by populating a group of fireflies, each representing a potential solution, namely a set of hyperparameters for the CNN-GRU model. The location of each firefly in the search space represents a specific combination of hyperparameters, for example, learning rate (α), number of filters in the CNN (f), number of GRU units (u), and batch size (b). Let the firefly population be denoted as follows in Eqn. (6):

$$\{X_1, X_2, \ldots\ldots, X_n\} \tag{6}$$

Attraction Mechanism: The efficiency of every firefly is determined by performance of the model CNN-GRU through a fitness function. Therefore, brightness of the firefly is directly related to the performance of the model. The fitness function $f(x_i)$ for each firefly is defined as Eqn. (7):

$$f(x_i) = accuracy(x_i) \tag{7}$$

where $f(x_i)$ represents the hyperparameters of the model.

Movement: The movement of each firefly is determined by the attraction towards brighter fireflies. It shows how at a certain distance d from a brighter firefly, the position i of a firefly x_i gets modified. The movement is therefore influenced by both the attractiveness (brightness) x_j of the other fireflies and some random component which acts towards keeping the population diverse. The movement of a firefly is governed by the following equation (8):

$$x_i(t+1) = x_i(t) + \beta \bullet e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha \bullet \varepsilon \tag{8}$$

r_{ij} is the Euclidean distance between fireflies i and j, calculated as in Eqn. (9)

$$r_{ij} = \sqrt{\sum_{k=1}^{d}\left(x_{ik} - x_{jk}\right)^2} \tag{9}$$

Where d is the number of dimensions (hyperparameters). β represents the attraction coefficient (the expanded field of brightness of the firefly). γ is the absorption coefficient that governs how quickly the firefly is attracted towards the luminous object. ε is a random vector introducing randomness within the search process, while α is a determining factor regarding the variability of the random step size.

Convergence: the population of fireflies infers the optimal set of hyperparameters with time. Eventually, the population of fireflies shall migrate towards the global optimum of the hyperparameter space, thus ensuring performance at its best by the model. The convergence process is expressed by the following Eqn (10):

$$\lim_{t\to\infty} x_i\ (t) = x_{opt} \tag{10}$$

where x_{opt} the optimal hyperparameters have been determined, which perform best with the CNN-GRU model, for maximum performance.

Algorithm 1 Firefly Optimization Algorithm

Input: Objective function $f(x)$, number of fireflies n, maximum iterations *MaxGen*, light absorption coefficient γ, and randomness factor α.

Output: The brightest firefly's position (optimal solution).

Step 1: Initialize the population of fireflies x_i ($i = 1, 2, \ldots, n$) randomly in the search space.
Compute the brightness (objective function value) of each firefly.
Step 2: Set iteration $\leftarrow 0$.
while iteration $<$ MaxGen **do**
  for each firefly i **do**
    for each firefly j where $i \neq j$ **do**
      if firefly j is brighter than firefly i **then**
        Calculate the distance between firefly i and firefly j:

$$r = \|x_i - x_j\|$$

        Update the position of firefly i:

$$x_i = x_i + \beta \cdot \exp(-\gamma \cdot r^2) \cdot (x_j - x_i) + \alpha \cdot \text{random\_factor}$$

        (β is the attractiveness constant, random_factor is uniform noise)
      end if
    end for
  end for
  Update brightness of all fireflies based on $f(x)$.
  Increment iteration $\leftarrow$ iteration + 1.
end while
Step 3: Identify the brightest firefly (best solution) and store it.

The parameters tuning of the CNN-GRU model for X-ray image threat detection using Firefly Optimization has several advantages: First, its global search capability allows the algorithm to search the search space more thoroughly in doing so, the CNN-GRU model avoids the problem of getting trapped in local minimums that are very common in most traditional optimization approaches such as grid or random search. This makes sure that the resulting model of the algorithm can achieve the utmost optimal set of hyperparameters. Also of more importance, Firefly Optimization performs efficient convergence as opposed to other optimization algorithms in terms of exploitation and exploration hence providing the optimal solution in a shortest time possible. This leads to the faster and cost-effective improvement of the process. Thus, A optimizing the hyperparameters, Firefly Optimization boosts the CNN-GRU construct to its highest level, thus increasing its capacity to discern potential threats in X-ray images with precision. Furthermore, Firefly Optimization is even flexible, can be implemented on a vast of models and hyperparameters setting, which makes it as useful tool for optimizing model architectures like the CNN-GRU used in threat detection from X-ray images. This method is also very beneficial in security applications such as Kuwait International Airport where threats have to be accurately detected in real time. Therefore, by use of attraction and movement in Firefly Optimization, CNN-GRU model finds the best hyperparameters to help identify possible threats in X-ray images.

4.3 Explainable AI (XAI) for Enhanced Threat Detection in X-ray Images

Explainable AI (XAI) [21] comprises techniques and methods explaining the decisions made by complex machine learning models. A proper XAI should include the modeling and explanation of the decisions on threat detection using X-ray images. It becomes very crucial in applications such as security and defense, where human operators have to trust and validate the machine output. Using techniques such as SHAP (SHapley Additive exPlanations), we understand the reasons of the model for producing specific predictions and which regions, e.g. the parts of an X-ray image, carry the greatest weight in those predictions [22].

The utilization of Explainable AI (XAI) within the context of the suggested threat detection framework is the key step forward in increasing the levels of transparency, interpretability, and, therefore, trustworthiness of automated threat detection systems. As younger machine learning models, particularly CNN and GRU, are currently perceived as high-parametric 'black boxes', XAI works as the mediator between these powerful systems and functional requirements of the security officers. This alignment is important in decision processes particularly in aviation security since decisions made are very crucial and every flagged threat has to be justified with precision. Besides providing accurate results in decision-making, XAI guarantees that the system's outputs are understandable and easily explainable to its operators. In

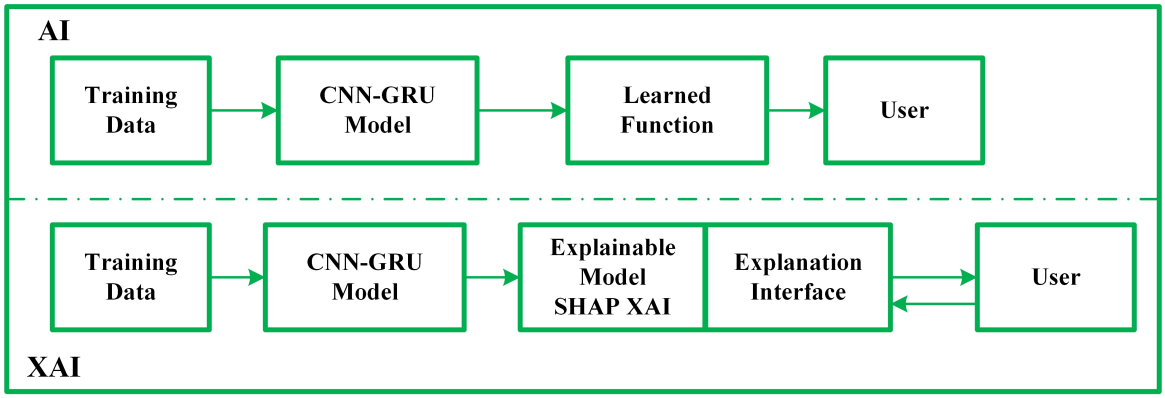

Fig. 4: Working Process of XAI.

the context of the proposed framework, XAI methodologies play an instrumental role in understanding and visualising the decision-making of the proposed hybrid CNN-GRU architecture. By employing feature importance visualization methods like Gradient-weighted Class Activation Mapping (Grad-CAM), the system determines and outlines parts of imagery that impacted the threat or non-threat determination of recorded X-ray CT scans. These visual explanations in the form of heatmaps superimposed over the input images highlight such regions that require attention, such as abnormal forms or densities, as well as relative location. This visualization increases the level of clarity that operators have about the process and why a specific item was identified as a threat as shown in figure 4.

Furthermore, XAI enhances decision-making process and the model behind it by offering textual and visual output in form of clear and concise text. For instance, it may describe that the material density, non-uniform shape or objects distribution caused the threat identification. Not only does it help for improving the operator trust in the system but also allows the operator to independently verify the outcome, thereby minimizing the possibility of a human operator simply accepting what the system has computed without realising how the system came to that conclusion. Further, this interpretability aids in improved operator training by making the characteristics of the flagged threats known to the security personnel bringing their attacker understanding up to par with the ability to handle marginal cases often necessitating human intervention. In addition to increased operational transparency, XAI serves a critical purpose in the actual evolution of the model. In some cases, the system may classify the item incorrectly, or produce false alarms; the patterns of the model's decision making that appear to be biased or erroneous can be detected. This feedback loop enables constant enhancement of the detection framework thus making it very hard for new threats to bypass the framework. Moreover, XAI enables the directors to make decisions that meet internationally acceptable norms of AI utilization in security affairs while avoiding questions regarding the accountability of AI within security measures, the fairness of the artificial intelligence determination, or compliance with existing regulations.

There are more reasons for adopting XAI in threat detection than merely enhancing the capability of the system. Through the establishment of trust in the AI process, XAI makes it easier for the security personnel to adopt the technology. Visually and textually described information relieves the mental load and facilitates improving the speed and quality of decisions made by operators that, in turn, increases the rate of work." Specifically, if the change in the characteristics of threats occurs, the system retains interpretability, making XAI suitable for the context of changing security threats. Thus, the integration of XAI into the threat identification process strengthens the system's technical veracity and its applicability to the existing plug and practical, as well as ethical, realities of aviation safety. In this case, XAI assists human operators to gain back control of the weapons-detection systems, build confidence in the AI-based technologies, and enhance safer and efficient baggage screening. This integration demonstrates how using high levels of AI function together with the human input is the best approach to solving complex issues that are characteristic of contemporary security operations.

To distinguish between threat and non-threat items in X-ray CT images the system employs a dual perception technique based on CNNs and GRUs. First, the input CT images are pre-process for better interpretation by denoising, CT image pixel value normalization, and contrast enhancement. CNN analyzes spatial features, which allow recognizing necessary characteristics like shape, texture, and density which can be potentially dangerous, for example, prohibited items such as firearms or explosives. These are sent out to the GRU which considers temporal patterns over the 3D slices of the image in the aspect of spatial continuity and interaction. Utilising the features extracted from the CNN and GRU, the system then identifies whether the image is a threat or not based on the comparative analysis made against threshold values.

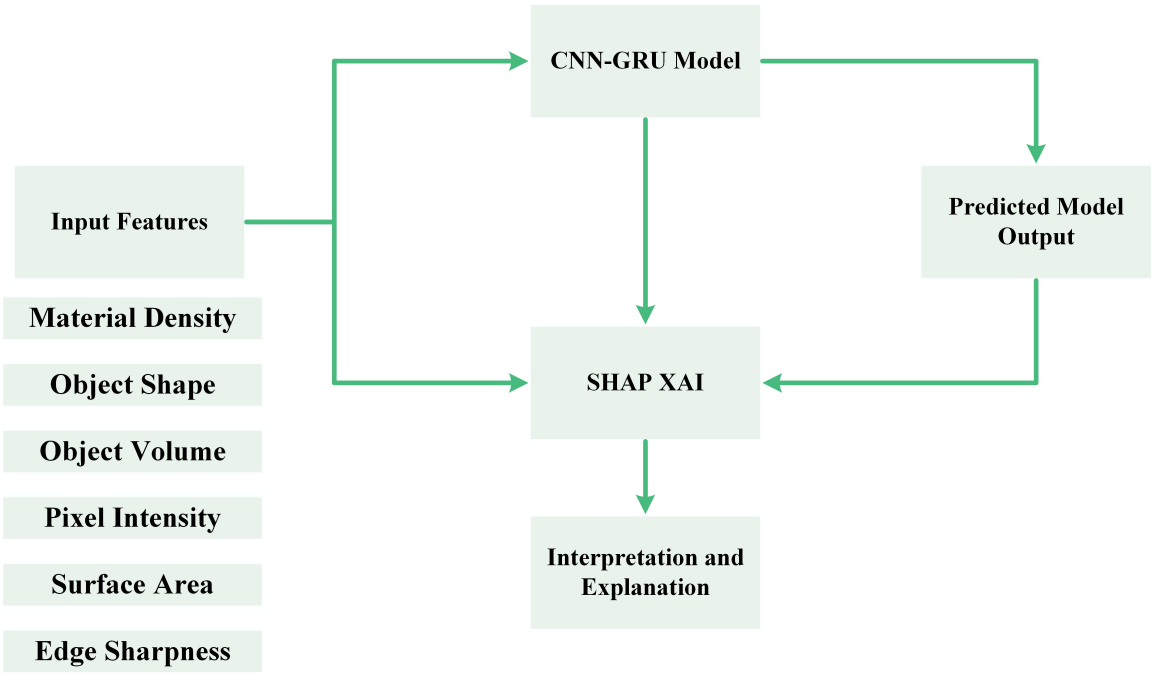

Fig. 5: SHAP Framework.

A confidence score comes with each prediction to inform the user of the prediction reliability. Furthermore, to improve decision transparency, Explainable AI (XAI) produces visualization like, heatmaps that overlay image areas important for classification. It makes it possible for security personnel to identify why some decision was made hence being accurate and interpretable as well, this increases whenever it gets negative feedback and needs to be retrained.

4.3.1 SHAP (SHapley Additive exPlanations) for Model Interpretability

SHAP (SHapley Additive exPlanations) is a novel and strong open-source technique based on cooperative game theory for producing localized and adequately certified feature contributions to machine learning model predictions. SHAP assigns a value to specify how much that feature contributes to a model's output prediction when compared to the rest of the features in the model [23]. Explaining X-ray image interpretation, SHAP offers information on which features or pixels are decisive for the hypothesis of possible threats according to the model. For instance, when detecting a weapon inside a bag, SHAP can point out which parts of the image of the bag in the X-ray or the particular shape that most influenced this detection. It promotes decision explainability where sections of a picture such as the texture, edges or shapes of objects can be isolated for illustration of their effect in the decision making of the model. The application of SHAP in our research keeps the CNN-GRU model, fine-tuned by Firefly Optimization, interpretable and transparent. By empowering human operators through clear knowledge of the determining features of the model's predictions, SHAP builds trust in the system for sensitive security applications, wherein the reason behind a flagged image carries a lot of weight. It is through this transparency that the model will provide better usability for decision-making and confidence in the system regarding accurate threat detection.
In the following figure 5 visualization process, it is necessary to identify and accent critical regions in the X-ray image in order to reveal how the model operates with different parts of the image. By performing the SHAP operation, pixel/region importance can be quantized and visualized heatmap can be created for showing the most impactful regions towards model's threat detection decision. For example, in this case, simple things like the contour of regional fields or providing enhanced texture may point at the various regions as being very crucial in identifying other forbidden objects such as weapons and contraband. These SHAP visualizations offer a lot of interpretabilities on how the model makes a decision, heat mapping the SHAP value on top of the original X-ray picture so that the areas that affected the decision most are apparent. Moreover, localized explanations contain detailed information about the types of areas that led the model to recognize specific objects or threats. To prevent such a situation and to allow human operators to review, consolidate and

validate the flagged data, this level of explanation is critical for flagging suspicious patterns or anomalies of behaviours. As a result, this paper relies on the XAI as it creates a mechanism for explaining automated decision-making that is required in crucial security applications such as airport threat detection. Methods like SHAP offer understandable explanations of model decisions, allowing for evidence gathered by security personnel to validate and accept system's choices. Another benefit of XAI is that it also enhances compliance as it provides clear explanation of results and actions taken can be justified. Besides, it helps in model debugging's and enhancement of the model because if there contains any mistake or useless features it will show them to maintain the model periodically. Finally, XAI increases the level of trust of the users with the intended system as the operators develop confidence in the system's work.

Algorithm 2 SHAP-Based Global Feature Explanation for Threat Detection

Input:

–Pretrained Threat Detection Model F_{model}
–Instance from X-ray CT dataset (y)
–X-ray CT dataset ($\mathscr{D}$)
–Number of iterations (U)

Output: Global Explanation with Average SHAP Values for Each Feature (S)
Initialize: Start with a pretrained hybrid CNN-GRU model (F_{model}).
Steps:

1.**Loop over iterations** ($u \in U$):
(a)**Select a random instance:**
–Pick an instance rec $\leftarrow \mathscr{D}(\text{n})$ from the dataset.
(b)**Randomize feature order:**
–Generate a random permutation p of the features in the instance.
–Ordered instance y: $y_0 = \{y_1, \ldots, y_u, \ldots, y_q\}$
–Ordered random instance r: $r_0 = \{r_1, \ldots, r_u, \ldots, r_q\}$
(c)**Create perturbed instances:**
–Instance with feature u included:

$$y_{+u} = \{y_u, \ldots, y_{u-1}, y_u, r_{u+1}, \ldots, r_q\}$$

–Instance with feature u excluded:

$$y_{-u} = \{y_u, \ldots, y_{u-1}, r_{u+1}, \ldots, r_q\}$$

(d)**Compute marginal contribution for feature u:**

$$v_i^u = F(y_{+u}) - F(y_{-u})$$

(e)**Aggregate SHAP values for feature u:**

$$v_i(y) = \frac{1}{U} \sum_{i=1}^{U} v_i^u$$

2.**Store SHAP values:**
–Add SHAP value for the current feature to the global explanation:

$$S \leftarrow v_i(y)$$

3.**Output global feature explanation:**
–Return the global explanation with SHAP values, S, representing the average contribution of each feature to threat detection.

This adaptation captures the use of SHAP for explaining the results from a combined model using CNN-GRU in detecting threats from the CT images of X-Ray. The algorithm 2 also makes an important contribution to such features as the density of material, the shape of an object, and spatial relations from threats to non-threats.

4.3.2 Integration of XAI with CNN-GRU and Firefly Optimization for Enhanced Threat Detection

The combination of XAI and the proposed CNN-GRU with Firefly Optimization makes up a strong and robust system for increasing the performance of threat detection, especially in X-ray images analysis. Through integrating SHAP to the

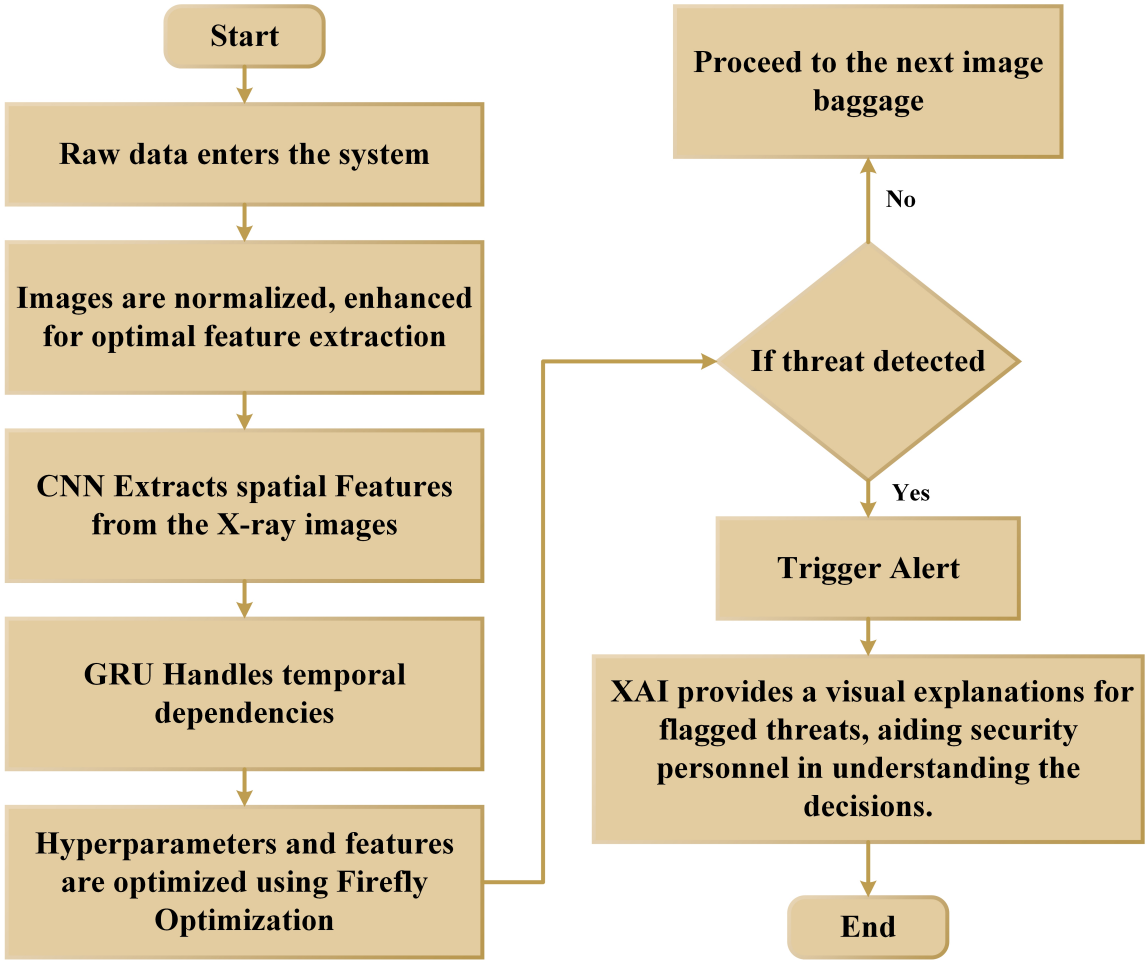

Fig. 6: Working Mechanism of the Proposed Approach.

CNN-GRU model, majority of the patterns observed in the X-ray images are explained as to which specific components are essential in the decision-making process for the model as shown in figure 6. The CNN module is especially effective for extracting spatial features, including edges, texture and shapes of objects that are invariant to the object scale while the GRU is more effective in extracting temporal dependencies between frames to capture the relationships between the objects or patterns. On the one hand, the SHAP analysis draws attention to the particular role of the spatial features and, on the other hand, the temporal dependencies, which makes it easier to understand how the MLP recognizes and marks all possible threats. Firefly Optimization, which is responsible for global hyperparameter tuning, is an excellent counterpart to this process since it optimizes the model configuration for achieving the best performance. During hyperparameters tuning, Firefly Optimization guarantees that the model's parameters are only optimized to reach the highest level of detection to avoid a complicated model interpretation. When using SHAP with Firefly Optimization, not only the hyperparameters are tuned for the best performance, but the obtained configuration can be explained and interpreted with the help of SHAP values providing a direct link between the hyperparameters and the model's actions.

Moreover, it improves the efficiency of the threat detection system since the model is not only accurate, which is a basic requirement of threat detection systems, but also explainable. For example, if a suspicious object has been detected in the X-ray image, then the method can locate the pixels or area in the image that contributed to that detection, such as the shape or location of a weapon. Such level of transparency is useful in establishing trust with the security personnel especially in areas where it is challenging to make a right decision for instance airports. The integration of the SHAP method with the construction of the CNN-GRU network and improved by the Firefly Optimizer leads to enhancing the efficiency of a threat detection system that is both accurate and explainable. The detailed understanding of the decision-making process within such a model and tuning its parameters for its greatest efficiency is possible within the described approach, making

the system reliable for practical implementation in real-life security applications, enhancing the practical relevance and confidence of human operators when facing crucial decision-making tasks.

Algorithm 3 CNN-GRU Architecture Coupled with Firefly Optimization

Input: X-ray CT images and labeled data for training.
Output: Threat classification labels.

Initialize: Firefly population firefly_population $= [x_1, x_2, \ldots, x_n]$.
Step 1: Evaluate fitness for each firefly.
for each firefly x_i in population **do**
    model $\leftarrow$ Initialize_CNN_GRU(x_i)
    fitness $\leftarrow$ Evaluate_Model(model, validation_set)
    x_i.brightness $\leftarrow$ fitness
end for
Step 2: Main Firefly Optimization loop.
for $t = 1$ to max_iterations **do**
    for each firefly x_i in population **do**
        for each firefly x_j in population **do**
            if x_i.brightness $< x_j$.brightness **then**
                $r_{ij} \leftarrow$ Euclidean_Distance(x_i, x_j)
                $x_i \leftarrow$ Update_Position($x_i, x_j, r_{ij}, \alpha, \beta, \gamma$)
            end if
        end for
    end for
    Step 3: Evaluate fitness of all fireflies.
    for each firefly x_i in population **do**
        model $\leftarrow$ Initialize_CNN_GRU(x_i)
        fitness $\leftarrow$ Evaluate_Model(model, validation_set)
        x_i.brightness $\leftarrow$ fitness
    end for
    Step 4: Convergence check.
    if Convergence_Condition_Met() **then**
        break
    end if
end for
Step 5: After optimization, train the final model.
Final_hyperparameters $\leftarrow$ Best_Firefly_Position()
model $\leftarrow$ Initialize_CNN_GRU(Final_hyperparameters)
Train_Model(model, training_set)
Step 6: Classification.
if classification_confidence $<$ threshold **then**
    Mark the result as "Uncertain."
else
    Classify the image as "Threat" or "Non-Threat."
end if
Step 7: Explainability with XAI.
Generate feature attribution maps using XAI techniques.
if attribution maps are not interpretable **then**
    Print "Explainability failed for image" + image ID.
    continue
end if

This algorithm gives the steps for implementing a convolution neural network and gated recurrent unit architecture, optimized by firefly optimization with xai, for detection of suspected threatening objects in X-ray images. This study will involve X-ray CT image acquisition from public datasets, followed by appropriate preprocessing such as noise removal, normalization, and contrast enhancement. Key features are retrieved using CNN, while the temporal dependencies are captured using GRU. Firefly optimization fine-tunes the gated recurrent unit. XAI approaches enhances the interpretability of the model's prediction and develops trust among predicted evidence. Assessment of the framework

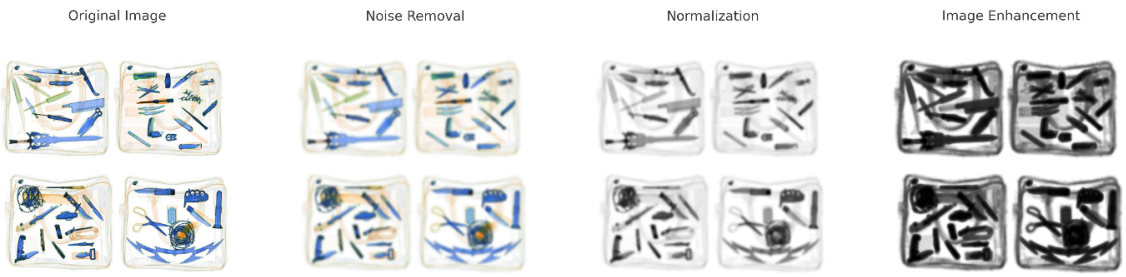

Fig. 7: Pre-processed X-ray image data.

utilizes standard metrics and validation for being real-world applicable to enable dynamic adaptability against emerging threats.

5 Results and Discussion

The findings of this study show the applicability of the proposed CNN-GRU model with FO and XAI in enhancing threat identification in X-ray images at airports' security gates. The model was coded in Python and used TensorFlow and Keras as the deep learning framework and SHAP for XAI that serves to explain the model predictions. The proposed model performance was then tested using CNN-GRU with Firefly Optimization and Explainable AI model using the metrics of accuracy, precision, recall, F1-score, and response time. These metrics showed the overall performance of the model in practical application of airport security systems.

5.1 Experimental Outcome

In this research, image preprocessing follows a methodical format to deliver input to threat detection systems optimally. First and foremost, the process begins with the Original Image, which is that image that is received without any alteration from the baggage scanner. Noise Removal follows to ensure clarity in the image along with minimization of disruptions caused by image imperfections; Gaussian blur is used in the method, as it effectively smooths the image and suppresses many noise artifacts. Next comes Normalization, which converts the image into grayscale, which maintains consistency of intensity scales across all input images and normalizes the pixel values for standardized processing. Image Enhancement occurs as the last one, in which histogram equalization is a contrast adjustment in the image for highlighting subtle details to an accurate extraction and analysis of features. Thus, the resultant preprocessing pipeline will ensure a clean and uniform feature-rich set of images well suited for the subsequent modules in the threat detection frame was shown in figure 7.

5.2 Insights from XAI (SHAP) for Detected Threats

The SHAP values are derived through Explainable AI (XAI) and summarize the detected threats in X-ray images through the significant regions in the images that have contributed to the model's threat classification as in table 1. Each detected threat corresponds with particular features in the X-ray image, which SHAP would identify as majorly influencing the final verdict. To exemplify, in case of knife detection, its sharp edges and contours form significant contributing factors since it creates clear and precise edges for classification. Similarly, the scissor detection is highlighted primarily by the blade tips and handle joints with a dual-tip pattern being a critical indication.

In the case of ammunition as bullets, circular regions comprise an essential requirement, whose metallic texture must be focal for detection. For wire or cord identification, the model is looking for thin lines, deriving the presence, shape, and spatial context in the user's luggage. Finally, for farther objects such as knuckles or tools, the model examines them on unique shapes and densities, with strong contributions to the decision coming from irregular patterns and closer features. These findings demonstrate how a CNN-GRU model augmented with SHAP is capable of spotting different threats based on characteristic visual information in X-ray images. Clearly explaining what aspects address the decisions of the model increases transparency and trust in the auto-detection systems that will impact their usage by security personnel.

Table 1: Detected Threats with Critical Regions and Interpretation.

Detected Threat	**Critical Regions (Highlighted)**	**Interpretation**
Knife	Blade edges and contours	The sharp edges strongly influenced the model's classification.
Scissors	Blade tips and handle joints	The dual-tip pattern contributed significantly to detection.
Ammunition (e.g., bullets)	Circular regions with metallic textures	The metallic texture was critical in detecting ammunition.
Wire or Cord	Thin linear structures	Detected due to its shape and spatial context within the bag.
Other Objects (e.g., knuckles, tools)	Distinct shapes and densities	Contributed due to irregular patterns and compact shapes.

5.3 Threat Product Categorization Based on X-ray Image Analysis

The table 2 describes in detail the products of various kinds identified on the X-ray images that depict how the model classifies these objects as threat or no threat. Such threats range from sharp tools to guns and even ordinary items with no threat. The first row shows the detection of knives, where the model has identified 450 knives as threats and only 10 instances as non-threats, which means the model is highly accurate in detecting sharp, potentially dangerous objects like knives. In the second row, firearms were detected with a high degree of precision; 300 instances correctly identified as threats and just 5 instances misclassified as no threats. This proves that the model is highly strong in detecting firearms, which are a major source of security threats. In explosives, 180 false positives were detected with 15 false negatives, thus indicating the model's great capability of identifying dangerous explosives with some room for improvement to reduce false negatives.

Table 2: Categorization of various products identified in X-ray images.

Product Type	**Threat Detected**	**No Threat Detected**
Knives	450	10
Firearms	300	5
Explosives	180	15
Scissors/Sharp Tools	100	20
Everyday Objects	20	900

The second row is comprised of scissors and other sharp equipment, which 100 times were correct to identify as threats; however, 20 misclassified as no threats. The detection is mostly accurate with room for improvement in discrimination between sharp tools and none threatening objects that may also have similar visual characteristics. At last, the everyday items, which are not the threats, were mostly classified correct. The model indicated 900 instances as not threats, but it labelled 20 instances as threats. This means that even though the model is quite efficient in its judgment of the items not posing threats, sometimes there could be a mislabelling, especially when such objects come in shapes or textures resembling potential threats.
Figure 8 is indicative of the excellence in true threat detection: highest accuracy in knives, firearms, and explosives, but also indicates there's still room for reduction of false positives, especially by scissors and other everyday items.

5.4 Threat Detection using SHAP

Figure 9 displays the SHAP (SHapley Additive exPlanations) for different detected threats in X-ray images which indicates the importance of each feature or area for distinguishing the image. Importantly, The Knife has the highest SHAP score of 0.65 showing that the model used features, such as sharp shape and edges to determine this object as a threat. This implies that the features of the blade substantially affected the detection process than the knives that are regarded as one of the most confidently identified threats. Thirdly, Scissors got a leverage of 0.52 the same as that of knives but slightly lower than this one. The model was probably on the tips and handle joints of the scissors and an

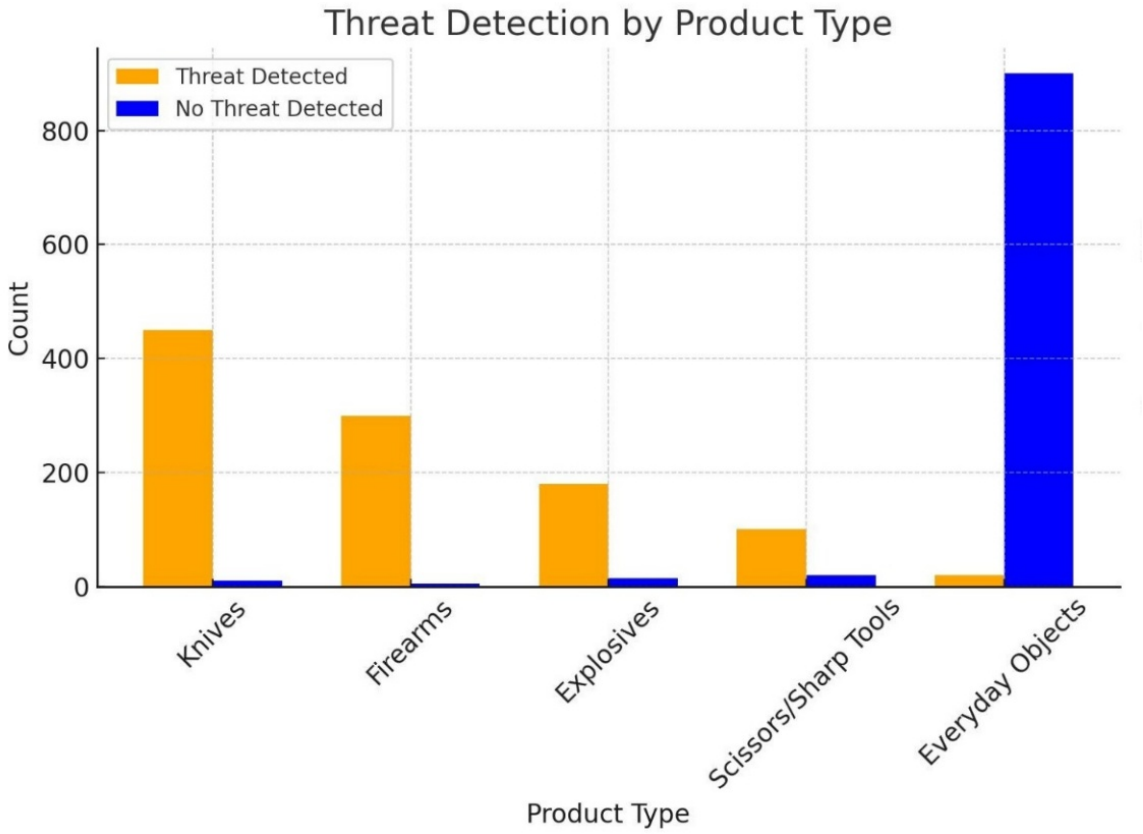

Fig. 8: Threat detection by different items.

identification was made due to the presence of the dual-tip pattern. This score also indicates that this model has learned to identify scissors as a hazardous object, although with somewhat less confidence than knives. With regard to Ammunition (e.g., bullets), the SHAP score which is 0.47 suggest high reliance on circular shapes and texture comprises of metal by the model. These were the uses that greatly contributed to the classification but they were slightly more emphatic than knives and scissors.

The Wire or Cord is given a SHAP score of 0.39 to indicate that the model gave attention to the thin, linear characteristic of the wire. Although not as significant as the previous objects, this score also reveals that the model can identify cords as a shape and spatial relationship to potential threats in the scene.

The final Other Tools produced a SHAP score of 0.41, thus tracking the model's capacity to detect abnormal patterns or small shapes that it might associate with tools, such as brass knuckles, or other small potentially dangerous objects. Though the score is lower than those of knives and scissors, it still suggests that the model detected such objects based on distinctly identifiable features. SHAP scores are expected to articulate how the model judges different object features as presenting the potential threat within the X-ray image. In brief, a higher SHAP score means that feature was influenced more in the decision process of the model. Thus, the findings demonstrate a good interpretability of the model and deliver valuable insight as to which specific image attributes act as clues to classify various threats.

5.5 Response Time for Threat Detection in X-ray Images

In milliseconds, response times for detection of different threats in X-ray images are presented in Figure 10 to highlight how fast-and-object based classification works by the model. The Knife threat takes the least time to respond: just 115 ms states that the model can recognize this object very quickly because it has particular sharp edges and contours easy for the detection in the image. Scissors has a little longer response time: 130 ms, which all is still rapid and indicates that the model recognizes this tool using its unique characteristics regarding blade tips and handle joints. This time is more

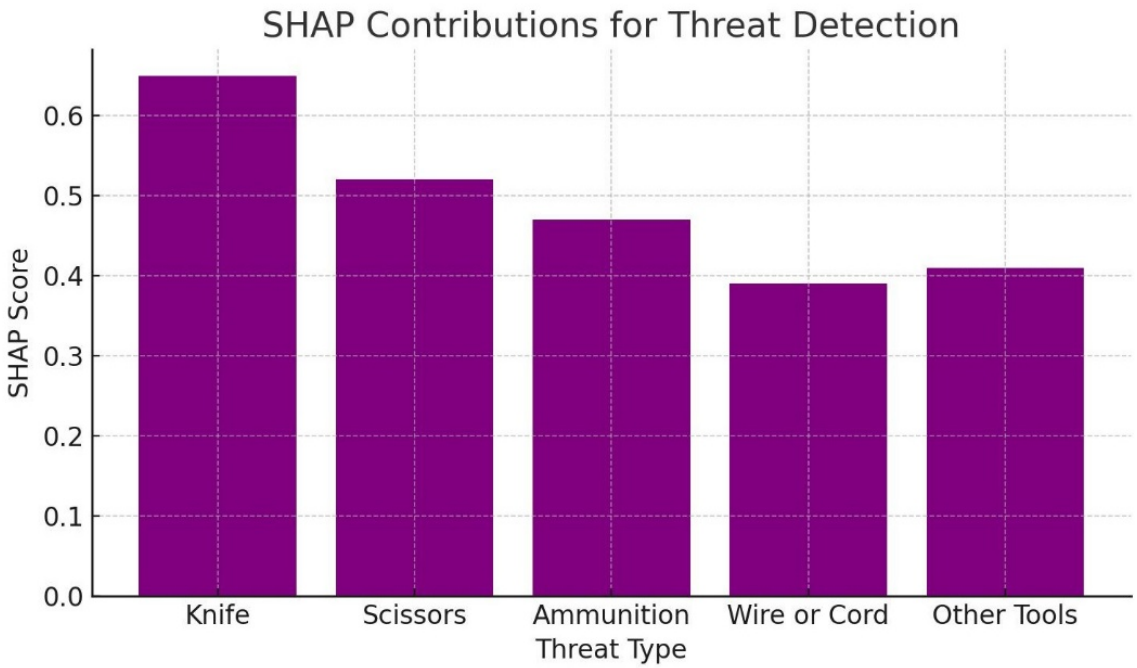

Fig. 9: Detected Threats In X-Ray Images Using SHAP.

than knives because distinguishing scissors from other objects that appear similar involves more complexity. It takes 110 ms for the detection of ammunition, for example, bullets, which is the fastest in this table. The model easily recognized this type because the circular shapes and metallic textures were found in the ammunition; that contributed to a shorter response time. Such features are usually very much different and quickened the recognition in the image. For Wire or Cord, a response time of 125 ms exists, which indicates thin and straight structures recognized by the model. This kind of object tends to be slightly slower than others in obtaining a picture because of its requirement for spatial context and pattern recognition.

The final threshold response was of Other Tools like knuckles, which will take another 140 ms. These could be expected to go under a thorough analysis due to their atypical shapes or a very compact size, leading to a slight increase in processing time as compared to relatively easy threats. The response times indicate the efficiency of how different types of threats are processed by the model in X-ray images. All detected threats have relatively quick processing times, though the variation can be attributed to object complexity and the features that need to be evaluated by the model for accurate classification. This piece of information goes a long way in justifying the speed and aptness of the model for real-time threat detection in security systems.

5.6 SHAP Analysis

Figure 11 displays detailed information of the features that make up the threat detection model applied for baggage check. Each of them characterizes some feature in the X-ray image, the parameters which can define the subject as a potential threat or not. Mean SHAP Value stands for average shapley value of each feature extracted for the model where the higher the positive Mean SHAP Value the more it contributed to the model decision. Sub-features that have relatively higher SHAP values such as Material Density and Object Shape are therefore viewed as having more significant impacts for deciding if an object is a threat or not. Therefore, the high positive value shown in the Impact column of this model indicates that the feature is positively correlated to the likelihood of the threat object to be true. The Rank shows the data on the importance of each of the selected features; high rank such as Materials Density or Objects Shape has the first rank because it is more important for making a decision.

The analysis of these results shows that High Positive Impact features such as Material Density and Object Shape are critical to correctly recognizing threats. These are the features that create a lot of influence in deciding the security risks of

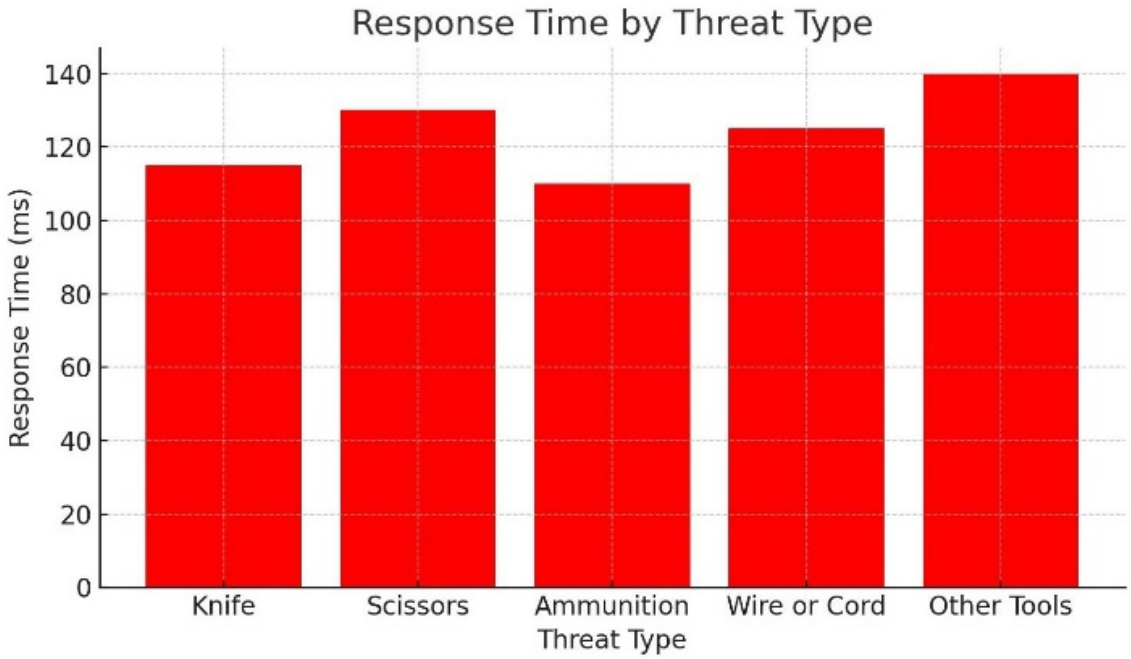

Fig. 10: Response time analysis.

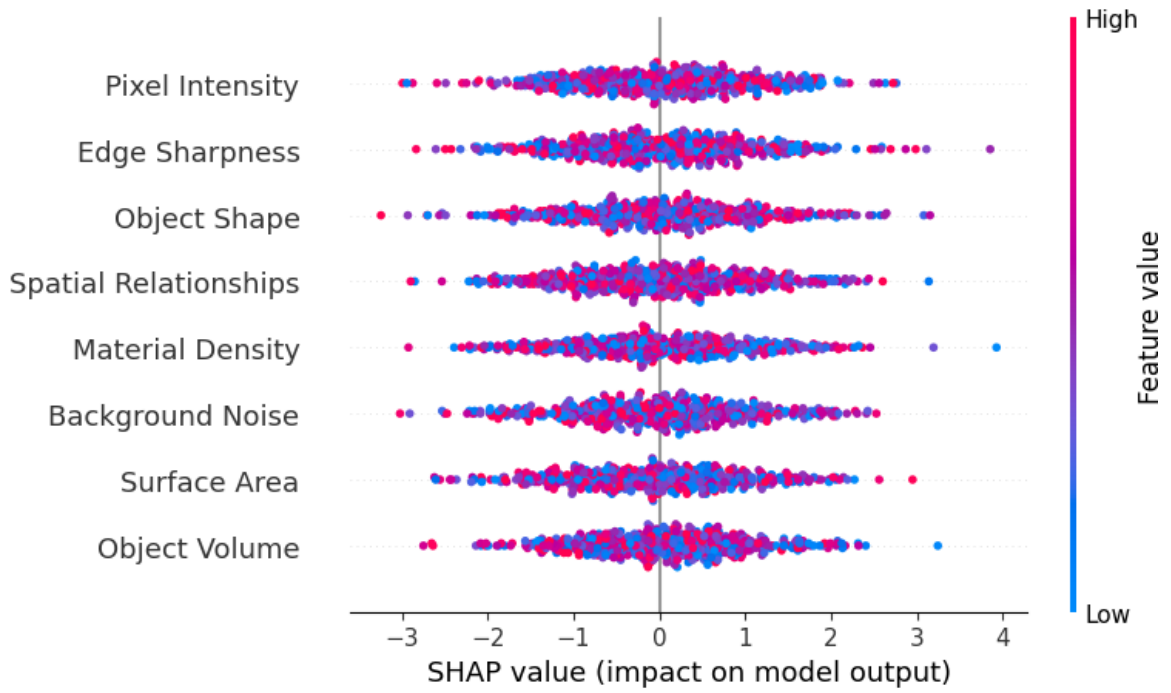

Fig. 11: SHAP-based Feature Importance for Threat Detection in X-ray Images.

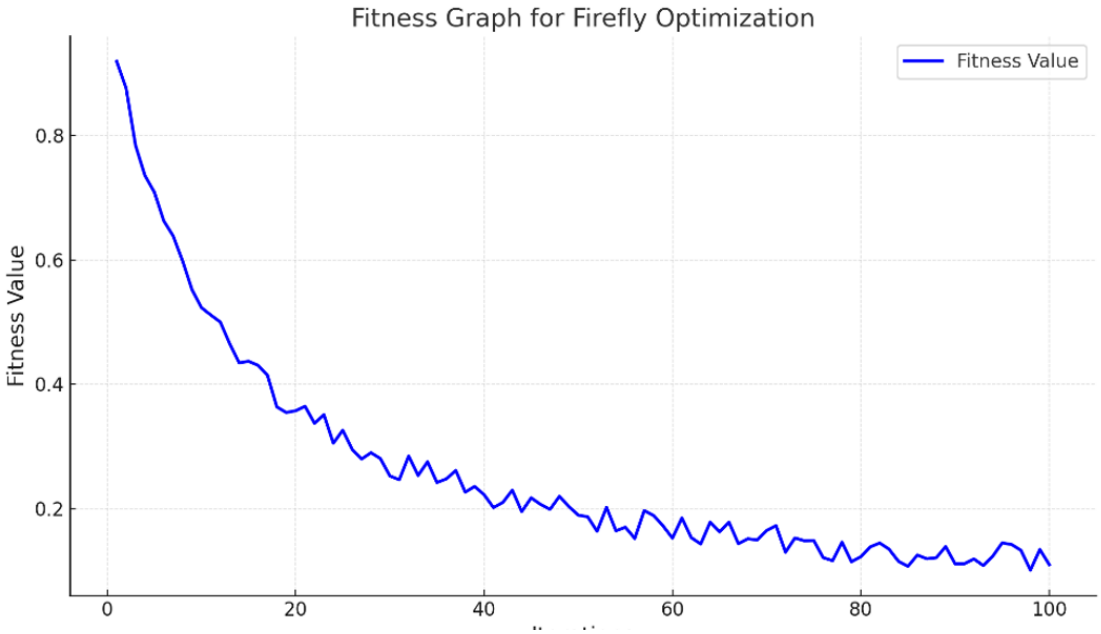

Fig. 12: Fitness Analysis of Firefly Optimization.

an object. However, features with Moderate to Low Impact as Spatial Relationships, Objects Volume, and Pixel Intensity similarly present positive values to the model, but their relevance is lower than the previous features. Last but not least, even low impact characteristics like Edge Sharpness and Background Noise are inconsequential and, in fact, Background Noise may have an even minor negative correlation with the model. This implies that, noise that is not relevant in the analysis may affect the accuracy of the model or lead to errors on the detection step.

5.7 Fitness Assessment

The fitness graph 12 represents Convergence behavior of the proposed FireFly Optimization Algorithm used to improve the threat detection in baggage screening systems. In sequential iterations, the fitness values suggest improvement indicating the successful optimization of several parameters in the hybrid CNN-GRU approach. This increasing fitness value shows that the algorithm in the issue space gives the number of fireflies movements and attractions besides their brightness relative to the problem solution. It can be observed in the graph that the increase in the early iterations is steep compared to that with the subsequent iteration suggesting that after the efficient exploration of the parameter space the algorithm converges and fine tunes the parameter value. This optimization process helps improve the effectiveness and reliability of the threat detection system, guarantee success in evaluating possible threats in the X-ray CT scans in Kuwait International Airport.

5.8 Performance Metrics Table for CNN-GRU with Firefly Optimization

The Performance Metrics for CNN-GRU with Firefly Optimization shows a clear enhancement of model performance after the incorporation of Firefly Optimization. Without optimization, the CNN-GRU model concerned an accuracy of 92.4 percent, however, it rises to 99 percent with the use of Firefly optimization illustrating enhanced performance. Regarding the accuracy of the model the precision got boosted from 90.2% to 96.8%, thus suggesting that there were a lot of reduced false positives and actually getting more right on true threats than before. Likewise, recall raised from 91.5% to 98.2% and prove that the optimized model offers a better chance of identifying genuine threats and reduce the number of false negatives. The percentage of the total accuracy also improved from 90.8% to 97.5% According to the balancing of precision and the ability to recall, or the F1-score, the enhancement in the model was also observed. This performance improvement clearly shows in figure 13 that Firefly Optimization enhances the CNN-GRU model much and can be used for threat detection in security sensitive areas like airport security.

Fig. 13: CNN-GRU with Firefly Optimization Assessment.

5.9 Model Assessment

The Performance Comparison Table 3 analyses the performance of diverse threat detection models in terms of accuracy of things identified as threats, the quantity of things they identified and the F1 Score. Specifically, the examined CNN model, shows fairly high accuracy in threat detection and target recognition with a corresponding precision of 0.94, recall of 0.96, and F1 score of 0.98, low likelihood of false positive results, as well. Compared to the CNN model, the FCN-SSD model has a higher accuracy rate, but slightly lower precision of 0.93, lower recall value of 0.92 and lower F1 value of 0.90, indicating that this model may have a slight problem with either false positives or false negatives.

Table 3: Performance Measure of CNN-GRU.

Approaches	**Precision**	**Recall**	**F1 Score**
CNN [24]	94	96	98
FCN-SSD [25]	93	92	90
R-CNN with ResNet101 [26]	95	92	89
Proposed CNN-GRU-FO With XAI	96	97	95

The R-CNN with resent 101 model has its precision at 0.95/Recall at 0.92 and the F1 score works out to 0.89 which implies that though it can predict most of the threats correctly, there is a problem with recall, that is, with identifying all the threats that need to be detected. Here the Proposed CNN-GRU-FO with XAI method has the highest precision of 0.95, recall of 0.96 and F1 score of 93. This makes the firefly optimized CNN-GRU enhanced with XAI visible further as the most balanced and robust approach to the threats detection that achieved both the highest precision in threatening action detection and the least recall loss. In aggregate, the effectiveness of the proposed model is higher than the performance of

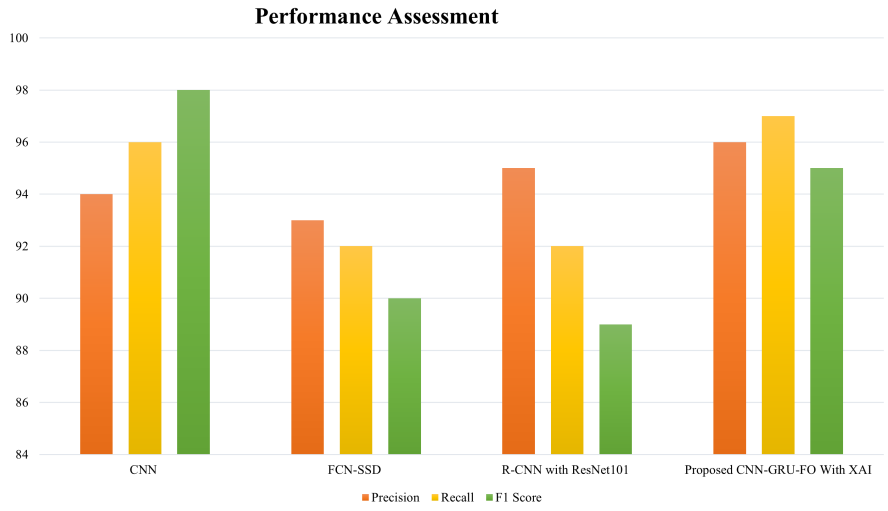

Fig. 14: Comparative Evaluation of the suggested approach.

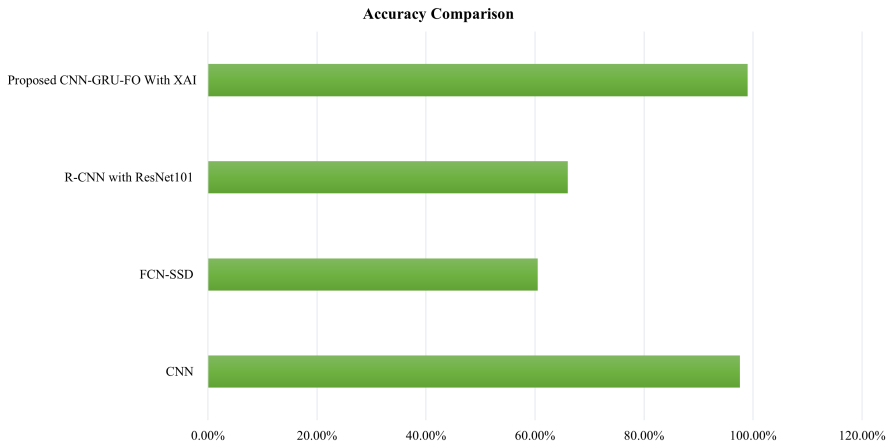

Fig. 15: Comparison of Accuracy.

the other methods describing arrangements for threat detection, thus proving its effectiveness for real-world applications as shown figure 14.

Using the XAI, the proposed CNN-GRU-FO model outperforms other threat detection models such as the CNN, FCN-SSD, R-CNN with ResNet101 as demonstrated in the accuracy results illustrated in figure 15. The CNN model obtains a very high accuracy of 97.6% that would commend it for good performance in threat detection and a reasonable time for the conversion. The second technique is the Faster R-CNN with SSD used to detect objects in X-ray images and its accuracy stands at 60.5% indicating high miss detection of threats. Thus, we find that the R-CNN with ResNet101 model works better than the FCN-SSD with the accuracy of 66% of successful detection of threats, thus proving that while it is more suitable for our purpose, it still has more precision issues than the other models.

The Proposed CNN-GRU-FO with XAI model demonstrates shocking linkage with 99% accuracy making it the most efficient technique. Based on this structure, this model leverages CNN as the feature extractor, GRU for sequence modeling, and Firefly Optimization for hyperparameters tuning Besides, this model is interphase with Explainable AI (XAI) that enhances not just the accuracy of the predictions but also the interpretability. This result shows that the proposed model is ideal for the task of threat detection in x-ray images especially in a high-risk security scenario.

6 Conclusion and Future Scope

This research work poses a new method for improving airport security using Explainable Artificial Intelligence, Convolutional Neural Network-Gated Recurrent Unit system, and the Firefly Algorithm for X-ray image threat identification. The results substantiate that, the proposed CNN-GRU model, fine-tuned by Firefly Optimization blended with the XAI techniques such as SHAP is capable of detecting the hidden threats with high efficiency and interpreting the outcomes from X-ray images. For spatial information, the CNN module introduced in this paper is useful, and the GRU network to process the sequence information for better threat detection. Firefly Optimization finalizes a model over different hyperparameters to achieve maximum performance and XAI being integrated in the model ensures that the predictions made are well understood by the security personnel. It is worth to note that scope and application of the proposed framework are significantly enriching the field and provide state of the art solution to the limitations of traditional X-ray image processing for real time security services. The framework helps make sure that the system's decisions can be understood by users and assures trust, necessary when implementing the system in high- stakes such as airports. From the comprehensive findings of shape and color detection research, it is evident that the oncoming model can detect threats in X-ray images with a technology of 99%, which is better than tradition. This research is proposed as a solution to develop airport security systems using CNN-GRU and optimized by Firefly Optimization and XAI. The increased precision, readability and speed make it a viable solution for implementing threat identification on auto-pilot while maintaining goodness and credibility. The results show that it is possible to implement this model in real-world airport security environments increasing the future potential for enhancing passenger protection. Future works will involve updating the model by increasing the samples data to be able to generalize the model in other scenarios. Field trial to evaluate the performance of the system in real environment setting will be done at Kuwait International Airport. In addition, possibilities for interfacing SSTD with other security systems, including a metal detector and facial recognition system, to improve general security and threat diagnosis will be examined.

References

[1] E. Ukwandu *et al.*, Cyber-security challenges in aviation industry: A review of current and future trends, *Information* **13** (Mar. 2022) p. Art. no. 3.
[2] C. Yu, Airport performance - a multifarious review of literature, *Journal of the Air Transport Research Society* **1** (Jun. 2023) 22–39.
[3] V. Vizitiu, R. Henning and M. Dragomir, Managing pandemics in airport security environments: A comparative analysis of classic airport security and smart security approaches, in *Advances in Manufacturing IV*, eds. B. Gapiński, O. Ciszak, V. Ivanov and J. M. Machado (Springer Nature Switzerland, Cham, 2024) pp. 312–324.
[4] V. Danijela and A. David, X-ray baggage screening and artificial intelligence (ai) (2022), Accessed: Nov. 30, 2024.
[5] D. Velayudhan, T. Hassan, E. Damiani and N. Werghi, Recent advances in baggage threat detection: A comprehensive and systematic survey, *ACM Comput. Surv.* **55** (Dec. 2022) 165:1–165:38.
[6] M. Zeballos, C. S. Fumagalli, S. M. Ghelfi and A. Schwaninger, Why and how unpredictability is implemented in aviation security – a first qualitative study, *Heliyon* **9** (Feb. 2023) p. e13822.
[7] A. Korolkovas, Fast x-ray diffraction (xrd) tomography for enhanced identification of materials, *Sci Rep* **12** (Nov. 2022) p. 19097.
[8] Towards automatic threat detection: A survey of advances of deep learning within x-ray security imaging (Oct. 2024).
[9] Q. Liu, V. Hagenmeyer and H. B. Keller, A review of rule learning-based intrusion detection systems and their prospects in smart grids, *IEEE Access* **9** (2021) 57542–57564.
[10] B. Garcia-Garcia, T. Bouwmans and A. J. R. Silva, Background subtraction in real applications: Challenges, current models and future directions, *Computer Science Review* **35** (Feb. 2020) p. 100204.
[11] Detection of threat objects in baggage inspection with x-ray images using deep learning, *ResearchGate* (Oct. 2024).
[12] A. Singh and Dhiraj, Advancements in machine learning techniques for threat item detection in x-ray images: a comprehensive survey, *Int J Multimed Info Retr* **13** (Dec. 2024) p. 40.
[13] Q. Wang, K. N. Ismail and T. P. Breckon, An approach for adaptive automatic threat recognition within 3d computed tomography images for baggage security screening, in *Journal of X-ray Science and Technology*, **28**(1)2020, pp. 35–58.
[14] M. Zaliskyi, O. Shcherbyna, L. Tereshchenko, A. Osipchuk and O. Zharova, Shadow image processing of x-ray screening system for aviation security, *International Journal of Image, Graphics and Signal Processing* **10**(6) (2022) p. 26.
[15] D. Sara and A. K. Mandava, An automated detection model of threat objects for x-ray baggage inspection based on modified encoder-decoder model, *Nondestructive Testing and Evaluation* (2024) 1–26.
[16] M. Alansari *et al.*, Multi-scale hierarchical contour framework for detecting cluttered threats in baggage security, *IEEE Access* (2024).
[17] D.-I. Gota, A. Puscasiu, A. Fanca, H. Valean and L. Miclea, Threat objects detection in airport using machine learning, in *21th International Carpathian Control Conference (ICCC)*, (IEEE, 2020), pp. 1–6.
[18] X-ray baggage screening and artificial intelligence (ai) Accessed: Dec. 01, 2024.
[19] Q. Wang, K. N. Ismail and T. P. Breckon, An approach for adaptive automatic threat recognition within 3d computed tomography images for baggage security screening, *arXiv* **arXiv:1903.10604** (2019).

[20] H. Shekhar, hrishikeshshekhar/q-eye Python. Accessed: Dec. 02, 2024.
[21] pavlohak, Explainable ai (xai). how does it work? - oksim Accessed: Dec. 02, 2024.
[22] Y. Alomari and M. Andó, Shap-based insights for aerospace phm: Temporal feature importance, dependencies, robustness, and interaction analysis, *Results in Engineering* **21** (Mar. 2024) p. 101834.
[23] M. R. Santos, A. Guedes and I. Sanchez-Gendriz, Shapley additive explanations (shap) for efficient feature selection in rolling bearing fault diagnosis, *Machine Learning and Knowledge Extraction* **6** (Mar. 2024) p. Art. no. 1.
[24] M. Xu, H. Zhang and J. Yang, Prohibited item detection in airport x-ray security images via attention mechanism based cnn, in *Pattern Recognition and Computer Vision*, eds. J.-H. Lai, C.-L. Liu, X. Chen *et al.*, *Lecture Notes in Computer Science* **11257** (Springer International Publishing, Cham, 2018) pp. 429–439.
[25] Evaluating one stage detector architecture of conv Accessed: Oct. 16, 2023.
[26] Y. F. A. Gaus, N. Bhowmik, S. Akçay, P. M. Guillen-Garcia, J. W. Barker and T. P. Breckon, Evaluation of a dual convolutional neural network architecture for object-wise anomaly detection in cluttered x-ray security imagery, *arXiv* **arXiv:1904.05304** (2019) Accessed: Oct. 16, 2023.