TPM Vol. 32, No. S2, 2025
ISSN: 1972-6325
https://www.tpmap.org/

AN INTELLIGENT COUNTERFEIT MEDICINE CLASSIFICATION PREDICTION SYSTEM USING MODIFIED YOLO: A SINGLE STAGE OBJECT DETECTOR

BINITHA S THOMSON[1], DR. W. ROSE VARUNA[2]
[1]PH.D. RESEARCH SCHOLAR, DEPARTMENT OF INFORMATION TECHNOLOGY
BHARATHIAR UNIVERSITY, COIMBATORE-46.
bini2796tha@gmail.com
[2]ASSISTANT PROFESSOR, DEPARTMENT OF INFORMATION TECHNOLOGY
BHARATHIAR UNIVERSITY, COIMBATORE-46.
rosevaruna@buc.edu.in

Abstract: Counterfeit medicines pose significant risks to public health, necessitating robust identification systems. This research is motivated by the requirement of healthcare whereas it outlines an intelligent counterfeit medicine prediction system using modified YOLOv11 (mod-YOLOv11), which is designed for single-stage object detection. The system improves YOLOv11 by incorporating an efficient backbone network and attention mechanism to improve feature extractors and classification performance. This work processes high-definition images of medicines to detect counterfeit products with minimal delay and hence it can be applied in real-time applications. The advanced features like adaptive spatial partitioning and efficient feature pyramid networks of YOLOv11 can detect counterfeits accurately even in adverse environments. The preprocessing of data can enhance precision and recall rates as well as the F1-score as compared to the standard object detection models. The system also applies lightweight architectures to minimize diversified computational complexity. Extensive experimentation achieves 83.23% accuracy and minimal time consumption of 234ms for 500 epochs. This research offers a tangible and feasible approach to the identification of counterfeits that helps the world combat counterfeit products, particularly fake drugs and safeguard the health of the population.

Keywords: Counterfeit medicine, neural network, time consumption, object detection, YOLOv11, and accuracy.

1. INTRODUCTION

Counterfeit medicines are a global concern, posing significant risks to public health and the pharmaceutical industry [1]. These counterfeit products may have no active pharmaceutical ingredients contain poisonous materials, or have wrong concentrations, which are detrimental to human health and may cause death. The conventional techniques used in counterfeit identification, including chemical tests and the naked eye tests, are slow, costly, and inconclusive. This acts as a call to the pharmaceutical industry to embrace innovative technology-based counterfeit detection equipment that is accurate, automated, and effective to meet the growing demand for large-scale detection in real-time to safeguard the health of the public and rebuild the much-needed confidence in the industry.

Image processing methods have also been proposed as a potential solution to the problem of counterfeit medicine detection, using differences in the physical appearance and chemical structure of the medicine. They involve image processing of high-resolution images of medicine with a view of capturing abnormal features that depict counterfeiting [2]. The conventional imaging processing techniques that are based purely on feature extraction that do not account for the complexities of the variations. This limitation has been well managed by Deep learning techniques which learn discriminative features directly from a large set of data.

Convolutional Neural Networks (CNNs) in the family of deep learning models have claimed efficient performance in the areas of image segmentation, classification, and object detection tasks. In counterfeit medicine detection

CNNs are used to detect small variations, which could otherwise go unnoticed by human operators where these systems are accurate and scalable. The emergence of improvement in transfer learning and the inclusion of pre-trained models give a high level of counterfeit detection models even when the quantities of data provided are low. Several deep learning structures have higher computational complexities that make real-time implementation challenging that necessitated for compact models.

The diversified version of the YOLO model is fast, accurate and flexible, which is a real-time object detection system specially designed for deep learning known as You Only Look Once (YOLO). Another thing to note is that YOLO does not employ many stages – region proposal, feature extractor, and classifier – unlike other pipelines present in object detection. This single-stage approach is the ability to detect with low latency or in real time which makes it ideal for applications such as detection of fake drugs. YOLO identifies the bounding boxes and class probabilities of an image in parallel with the usage of only one neural network with a grid-based feature mapping system [3].

YOLO was employed in counterfeit detection owing to the multiple-object detection feature and ability to recognize anomalies in a single image. Newer versions include YOLOv11 that add features like attention mechanisms, adaptive spatial partitions, and better backbone networks, which makes versions 11 better for detecting fine-grained features more complex scenarios. The ability to process high-resolution images by YOLO is a prominent in the decision-making process of fake medicine prediction, as it can contain unique characteristics or features.

YOLO's flexibility that allows it to operate on different hardware platforms on both server and edge ends enables its integration into various applications [4]. The increasing trend of counterfeit medicines and the effects it has emphasized the importance for better detection techniques. There are several reasons why current manual and chemical testing methods do not work and even if they do, the problem is worsened when the volume of pharmaceuticals increases. Automated systems employing image processing and machine learning are potentially beneficial, however, many of the current models do not possess enough accuracy, speed, or flexibility to be implemented on real-world scale [5].

The selection of YOLO as the base model for this research is informed by its real-time detection determination, resilience, and versatility [6]. Fake medicine identification poses challenges such as variation in colour by brightness level, object presence, and variation in design. The capacity of YOLO to generalize on various datasets and retain its efficacy even in adverse circumstances makes it suitable for this job. The lightweight architectures and optimization are incorporated to make the system compatible with limited resource devices like smart phones or handheld scanner devices for its usability in a remote or low resource environment. This research seeks to overcome the limitations of the current counterfeit detection systems by proposing a new YOLO customized for this purpose.

Advancements like advanced backbone network for feature extraction, attention mechanisms to concentrate more on the medicine details, usage of domain specific datasets increases the effectiveness of the system. The consequences of this research for society are tremendous. A good counterfeit medicine detection system serves consumers, as well as the regulators and manufacturers in their fight against counterfeit medicine. In this way, the system helps to reduce the spread of fakes, and, therefore, contributes to positive changes in the public health indicator and the entire spectrum of healthcare systems [7].

The implementation of such systems corresponds to worldwide trends to employ Artificial Intelligence (AI) and automation in solving major problems concerning healthcare. Counterfeit medicine detection is one of the biggest challenges worldwide, which requires improved methods that are accurate, fast, and easily reproducible. This field has been revolutionized by image processing in collaboration with deep learning, providing fully automated systems to detect less conspicuous deviations in medicine. It is therefore apparent that YOLO, due to its increased efficiency in delivering single-stage object detection, is well suited to counterfeit detection especially in real-time use. This research is motivated by the necessity of robust and scalable systems, and therefore provides alterations to YOLO that focus on its weaknesses while preserving its advantages. This work makes a useful and implementable contribution to the global fight against counterfeit medications prediction, and the protection of public health.

The article is organized as follows: the overview of counterfeit medicine prediction system using deep learning is discussed in Section 1, the comprehensive analysis of recent research is discussed in Section 2, the counterfeit medicines prediction system using mod-YOLOv11 is illustrated in Section 3, the experimental outcome is detailed

with appropriate discussion in Section 4, and the article is concluded with future research direction is detailed in Section 5.

2. RELATED WORK

This section overviews related works in counterfeit medicine classification systems using various computational models. Analysing these systems in detail showcases the use of sophisticated techniques such as, machine learning, chemical profiling, image analysis, blockchain for counterfeiting detection. The review seeks to assess the advantages, drawbacks, and efficiency of these strategies in dealing with the difficulties surrounding the identification of counterfeit drugs, together with their applicability in practice and further growth opportunities. The evaluation outlined below aims to review the existing methods in classifying fake drugs to fill the gaps outlined below.

Islam & Islam (2022) systematically reviewed studies to assess effectiveness of digital interventions in counteracting counterfeit and falsified medicines. After reviewing 1253 articles, the study selects fifty-one articles that reveal that technologies like blockchain, IoT, RFID, and image processing are useful in addressing this global health problem. They focus on supply chain risks identification, improvement of reporting, and integration of technologies to fight counterfeiting. The direction for future research is to examine other aspects of counterfeiting and improve digital approaches. This work gives prospect into ICT utilization for counterfeiting prevention of medicines suggest pertinent research gaps [8].

A group of researchers including Santos et al. (2020) investigated the use of gas chromatography/flame ionization detection (GC/FID) and gas chromatography/mass spectrometry (GC/MS) for identification of counterfeit medicines. Their approach is designed based on chemical characteristics of samples to help detect the counterfeit products. By employing methanol extraction and hierarchical cluster analysis (HCA), study finds old patterns of counterfeit formulations such as caffeine and benzocaine blends.Their GC-FID/MS approach shows prospective for the identification and tracking of counterfeit sample and thus for the forensic investigation that contributes to fight against counterfeit medicines in the international level through incorporating them in large databases [9].

Rasheed, Höllein, and Holzgrabe (2018) have also reminded that Information Technology (IT) tools that fight counterfeits have been reported. They pointed out examples including mobile applications, 2D bar codes and the use of machine learning in the quality evaluation kits. One of them is their discriminative discussion on the World Health Organization (WHO's) Global Surveillance and Monitoring System (GSMS), which which enhances supply chain visibility. There are research opportunities in the "lab on a chip" devices and paper analytical technologies for cheaper portable solutions [10].

Siamese network-based approach to counterfeit detection using near-infrared spectroscopy (NIR) was put forward by An-Bing et al. in 2020. Using a one-dimensional convolutional neural network (1D-CNN), the authors successfully obtained an accuracy as high as 97.3% for the unseen samples while compared to the traditional method such as Support Vector Machibe (SVM). This makes it capable of handling problems such as skewed data volumes and delivers reliable on-site drug identification for pharmaceutical Quality Contral (QC). The research has attained efficient accuracy for minimal dataset [11].

Prado Puglia et al. (2021) have proposed an image-based approach for counterfeiting identification of medicines. Applying the SVM and clustering methodologies, the authors successfully pointed the important regions of tablet, and got 100% classification rate on Cialis and Viagra tab datasets. Their more creative exploitation of heat maps improves interpretability, illustrating the possibilities of image analysis in identifying counterfeit drugs [12].

According to Sansone et al. (2021) counterfeit PDE5 inhibitors include the drivers like high demand and low awareness of health risks. They talk of measures like the sealed packs and laboratory detection techniques of the contaminants. Their results conclude that more focus on consumer protection and better legislation is needed to address such threats [13].

According to the study conducted by Ozawa et al. (2018), the cross-sectional surveys of substandard medicines were at 13.6% They further explained that the shares of substandard medicines were 12.3% in East Asia and Pacific, 12.3% in Europe, and Central Asia, 14.9% in Latin America and Caribbean, 13.8% in the Middle East, North They range it between $10 billion and $200 billion per year, noting the need for enhanced and better coordinated policies in regulatory measures and supervisory systems in order to mitigate public health and economic consequences [14].

Thomson and Varuna (2023) came up with a cognitive classifier model with the machine learning method for assessing fake drugs. Based on data from networks and logs, the model increases security against false drug networks. Their study moves the understanding of cybersecurity frameworks in detecting and responding to counterfeit drugs forward [15].

According to the literature review conducted by Islam and Islam (2022), Santos et al. (2020), conventional approaches, including blockchain-based traceability and chemical profiling, are used for the identification of counterfeit medicines, but there is no integrated, large-scale real-time solution. Islam and Islam (2022) discussed some lesser-studied facets of digital utilities for anticipatory counterfeiting prevention, while Santos et al. (2020) centred on the sample-based identification with the help of chemometric, which is time-consuming and less efficient in terms of functionality at the operational level. Like the work of Prado Puglia et al. (2021) in which image-based methods also yielded high accuracy, however, depending on features utilized. ModYOLOv11 fills this gap by providing real time, scalable, solution based on deep learning. As it avails itself of the improvement brought about by the YOLO-based architectures in terms of the object detection, it delivers higher accuracy in terms of detecting fake drugs using image processing. Its feasibility is achieved through low weight that is suitable for deployment in limited resource facilities, pharmaceutical quality control and low resource settings that lack adequate counterfeit detection.

3. PROPOSED METHODOLOGY- MOD-YOLO: A SINGLE STAGE OBJECT DETECTOR

The proposed counterfeit medicine classification and prediction system use a YOLOv11 neural network with some modifications to enable real-time object detection and classification. The methodology consists of four main stages namely pre-processing, segmentation, feature selection process and classification. The system intends to solve the problem of identifying counterfeit medicine in high resolution image under different light conditions. YOLO architecture is changed for better features extraction, attention mechanisms, and improved training. The whole pipeline is designed to enhance detection accuracy, reduce computational time and make it more feasible to use in real-life situations. The overall framework is given in Figure 1.

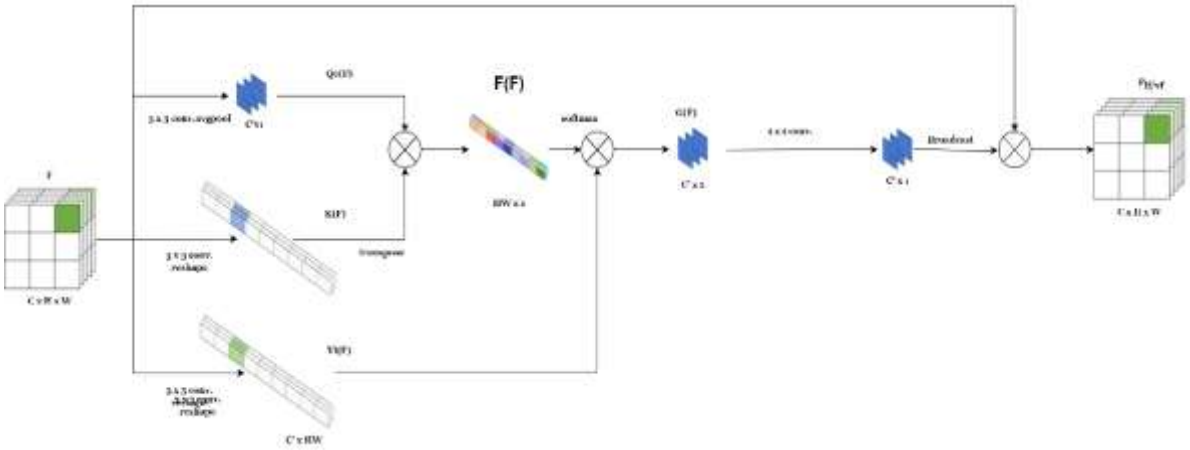

Figure 1. Mechanism of Mod-YOLO: A Single Stage Object Detector

3.1. Data Augmentation using GAN

In order to generate images and provide a label of authenticity, cGAN can be used to assist the generator by incorporating class information, which guarantees that the samples will belong to certain classes. This makes it possible to develop various fake drug images to match with real life situations. Unlike CycleGAN which is good at translating genuine medicine images to fake ones due to the learning of the mapping function between the two domains capturing the subtle differences. This approach guarantees that the kind of images produced and labeled as fakes contains the originality of real images, but carry some unique domain changes. For the generation of high-quality images, we can use StyleGAN while maintaining the microstructures of objects in the images. Thus, due to its capability of generating images with very high quality, StyleGAN enables the generation of counterfeit and authentic samples with similarly high quality, increasing the dataset's variability and model's resistance to counterfeit medicine sample identification.

The generator G(z | c) takes random noise input z and an authentic or counterfeit label ccc, and generates synthetic images accordingly. The objective of G is to create an image that the discriminator will have high error when determining the real and fake images. The loss for G is given in Equation 1.

$$\mathcal{L}_G = -\mathbb{E}_{z\sim p_z, c\sim p_c}[logD(G(z|c),c)] \text{ ----------(1)}$$

where the noise distribution is given as p_z, the distribution of class is given as p_c, and the D(x|C) is probability of discriminators x is given to the real class c.

The discriminator D(x,c) predicts whether an image x is real or synthetic for class c. The loss for D is given in Equation 2.

$$\mathcal{L}_D = -\mathbb{E}_{x\sim p_{data}, c\sim p_c}[logD(x,c)] - \mathbb{E}_{z\sim p_z, c\sim p_c}\left[\log\left(1 - D(G(z|c),c)\right)\right] \text{----------(2)}$$

where the real image distribution image is given as p_{data} and the image generation is given as $G(z|c)$.

The generator produces $N_{synthetic}$ images for underrepresented classes (e.g., counterfeit medicine). Variations in attributes are controlled by a latent vector z. For each class c, the generated image is given in Equation 3.

$$x_{synthetic} = G(z|c), z\sim p_z, c \in \{authentic, counterfeit\} \text{----------(3)}$$

The augmented dataset D combines real and synthetic images is given in Equation 4.

$$\mathcal{D} = \{(x_{real}, c_{real})\} \cup \{(x_{authentic}, c_{authentic})\} \text{ ----------(4)}$$

where the real image and synthetic image generated are given as x_{real} and $x_{authentic}$, respectively.

To improve the detection of counterfeit medicine, the generator is designed to generate images that are as real as those of actual counterfeit and authentic medicines while the discriminator is designed to distinguish between real and fake image for the two categories of medicines. Another technique used is hyperparameter optimization including learning rate and batch size to make training more stable and reduce mode collapse. Finally, the underrepresented classes are tackled using synthetic data generation where fake medicine images are produced with different fonts, colors and packaging to mimic real life scenario. These synthetic images are indeed mixed with real images to obtain a balanced training set of images where the generated samples are checked for their realism validity by domain experts. Last but not the least, with an aim to enhance the performance of the traffic sign detection task, an enhanced dataset is used to train a deep learning model like YOLOv11. In order to address any remaining class imbalance, a weighted loss function is used during training to provide reliable and accurate classification of counterfeit medicines.

3.2. Pre-processing

Preprocessing is used to prepare the input images for training and inference within the YOLOv11 model, as well as to normalize the images. First, images are normalized and rescaled to fit [S, S], where S is the desired size of the image while keeping aspect ratios of the input layers compatible with the network input layer. Smoothing techniques like Gaussian or median filters are used to delete noise from the images that maybe collected from the images or videos. By using variations that include rotation flipping brightness change or random cropping the data augmentation provides the model with variations that increases its ability to perform well under different conditions such as lighting conditions or occlusion.

The color space transformation is also used to enhance the features based on which the images are transformed into suitable forms like the RGB or HSV features such as structures or texts are crucial. These preprocessing steps in aggregate make the input data ready for subsequent feature extraction and high precision object detection. Input images are resized to W × H dimensions while maintaining aspect ratios. The resizing is given in Equation 5.

$$X' = Resize(X, W, H) \text{----------(5)}$$

Where X' is the resized image, is the input image, and W, H, are the target width and height.

Pixel values are normalized to lie within the range [0, 1] to ensure uniform gradient descent and it is given in Equation 6.

$$X'' = \frac{X'}{255} \text{----------(6)}$$

Techniques like rotation (θ), flipping (F), and color adjustments (C) generate diverse samples where the data augmentation is given in Equation 7.

$$X_a = Augment(X'', \{\theta, F, C\}) \text{----------(7)}$$

3.3. Segmentation

In the YOLOv11 framework, segmentation divides the image into Regions of Interest (RoI) thus narrowing down the detection process targeting specifically areas of interest. YOLOv11 is designed for object detection process, adaptive region selection uses dynamic grid division to divide the image into grids and then assign bounding boxes for areas most likely to contain counterfeit signs. The image is divided into S×S grid cells. Each cell is responsible for detecting objects whose center falls within it. The grid-based segmentation mechanism is given in Equation 8.

$$Grid\ Cell\ Size = \frac{W}{S} \times \frac{H}{S} \text{----------(8)}$$

Integrated attention mechanisms take this process a step further by focusing on prominent aspects in medicine that are clearly outlined as critical and as the key means of identifying counterfeit medicines. The multi-scale feature mapping allows YOLOv11 to detect images at different scales where counterfeit markers of different shapes and sizes are present. These improvements assures identification and differentiation while focusing on the features and areas of the most importance for counterfeit detection without the need for utilizing an additional segmentation models. Predefined anchor boxes $\{(w_a, h_a)\}_{a=1}^{A}$ represent common object aspect ratios. For each grid cell and it is given in Equation 9.

$$Bounding\ Box = (x, y, w, h, c) \text{ ----------(9)}$$

where *(x,y)* are center coordinates, *w, h* are dimensions, and c is the confidence score.

3.4. Feature Extraction Using YOLOv11

The changes introduced in the proposed YOLOv11 are architectural and functional improvements over the first generation of YOLO. These changes are intended to address issues such as low object detection especially on small objects, wrong positioning of bounding boxes as well as slowness in the computation of class probabilities. The enhancement is lovingly on the loss function, inception structure and Spatial Pyramid Pooling (SPP) that improves the precision of identification of counterfeit medicines, scalability and robustness to object detection.

Inception layers extract features at different scales using parallel convolutions with varying kernel sizes. For input *X* with dimensions *H*×*W*× *D* is given in Equation 10.

$$O_{\text{inception}} = Concat\ (Conv_{1\times1}(X), (Conv_{3\times3}(X), (Conv_{5\times5}(X), (MaxPool_{3\times3}(X) \text{----------(10)}$$

This allows capturing both fine and coarse features. The computation complexity of convolutional layer is given in Equation 11.

$$Complexity = k_h \times k_w \times D_{in} \times D_{out} \times H \times W \text{----------(11)}$$

where the kernel dimension are given as k_h, k_w, input depth is given as D_{in}, output depth is given as D_{out}, and spatial dimensions are given as H, W.

Spatial Pyramid Pooling (SPP) generates fixed-size feature maps regardless of input dimensions. For a pyramid with N levels is given in Equation 12.

$$O_{SPP} = \cup_{n=1}^{N} MaxPool_{r_n \times c_n}(X) \text{ ----------(12)}$$

Where $r_{n,}\ c_{n,}$ are bin dimensions at level *n*.

The feature map X with the size $H \times W \times D$ is given in Equation 13.

$$O_{SPP} = \sum_{n=1}^{N} r_n \times c_n \times D \text{----------(13)}$$

3.5. Feature Selection and Bounding Box Prediction

YOLOv11's architecture is enhanced with advanced feature extraction capabilities to meet the specific demands of counterfeit medicine detection. The backbone network is optimized with a lighter and more efficient variant, such as CSPDarknet, which balances computational efficiency with robust feature extraction. A Feature Pyramid Network (FPN) is incorporated to improve the detection of small objects and fine-grained details, crucial for identifying subtle counterfeit markers. Attention modules, including self-attention mechanisms like SE (Squeeze-and-Excitation) blocks, are integrated to prioritize salient regions in the image, allowing the model to focus on distinctive features. Medicine physical properties, such as variations in pill size, shape, or edges, irregularities in color uniformity, differences in surface smoothness or coating quality, and distinct or unusual odors, can be effectively captured through image and sensory data, making them suitable features for CNN-based learning and counterfeit detection.

Anchor box refinement further enhances performance by customizing anchor boxes to match the aspect ratios and dimensions typical of counterfeit medicine, improving bounding box precision. YOLOv11's grid-based detection enables simultaneous localization and classification, effectively distinguishing between counterfeit and genuine. Additionally, advanced loss functions, such as CIOU loss, are utilized to improve bounding box regression accuracy, ensuring precise detection outcomes. Given an input image I, the backbone network extracts feature maps using Equation 14.

$$F = Backbone(I) \text{ ----------(14)}$$

where $F \in \mathbb{R}^{H_f \times W_f \times D_f}$ with $H_f \times W_f \times D_f$ indicating height, width, and depth of the feature map.
The significant features are enhanced using Equation 15.

$$F' = Attention(F) \text{ ----------(15)}$$

where F' indicates the refined feature map.
Each grid cell predicts B bounding boxes, each defined in Equation 16.

$$B_i = \{x_i, y_i, w_i, h_i, c_i, \{p_i(c)\}_{c=1}^{C}\} \text{----------(16)}$$

where x_i, y_i is center coordinates, w_i, h_i is Width and height, c_i is Confidence score, and $p_i(c)$ is Probability of class c.
The loss function is designed to minimize errors in localization, confidence, and classification probabilities. The improved loss is given in Equation 17.

$$L = \lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} [(x_i - \hat{x_i})^2 + (y_i - \hat{y_i})^2]$$
$$+\lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} \left[(\frac{w_i - \widehat{w_i}}{\widehat{w_i}})^2 + (\frac{h_i - \widehat{h_i}}{\widehat{h_i}})^2 \right]$$
$$+\sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} (c_i - \hat{c_i})^2 + \lambda_{noobj} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{noobj} + (c_i - \hat{c_i})^2 + \sum_{i=1}^{S^2} I_{ij}^{obj} \sum_{c=1}^{c} (p_i(c) - \hat{p_i}(c))^2 \text{----------(17)}$$

where $\lambda_{noobj}, \lambda_{coord}$ is Weights for localization and confidence, $\hat{x_i}, \hat{y_i}, \widehat{w_i}, \widehat{h_i}$ is Predicted bounding box parameters, and I_{ij}^{obj},I_{ij}^{noobj} is Indicator functions for object presence.

3.6. Classification

In YOLOv11, the classification module designates a detected object as counterfeit or genuine with the help of fine-tuned prediction head outputs. There are several methods of boosting up class probabilities as follows. In class-specific post-processing, confidence thresholds and non-maximum suppression (NMS) are utilized to delete low confidence prediction and to eliminate the extension of the same bounding boxes, making the outputs neater and more accurate. The proposed fine-tuning approach uses a dataset of images of genuine and fake medicines and initial weights from the generic object detection to speed up the training process.
Yolov11 is fine-tuned with transfer learning to solve the problem of identifying fake medicine which has very different characteristics than original images, rather than relying on large labeled datasets which are not available in most cases. For additional accuracy enhancement, the verification mechanism in formation works in collaboration with other classifiers different from YOLOv11 to perform ensemble verification and can work with other light-weight classifiers like SVM or decision tree classifiers. The final layer of YOLOv11 assigns a class label to each bounding box. Softmax activation computes the probability distribution is given as Equation 18.

$$p_i(c) = \frac{e^{z_i(c)}}{\sum_{k=1}^{C} e^{z_i(k)}} \text{----------(18)}$$

Where $z_i(c)$ is the logit for class c.
The predicted class is given in Equation 19. The procedure of Mod-YOLOv11 is given in Algorithm 1.

$$\hat{c_i} = \arg \max_{c \in \{1, \ldots, c\}} p_i(c) \text{----------(19)}$$

Algorithm 1. Classification using Mod-YOLOv11

Input: High-resolution image dataset (I) with varying light conditions
Output: Classified objects: counterfeit or genuine with bounding boxes and confidence scores
Preprocessing Stage
for each image X in dataset I:
X'=Resize(X, W, H) #Resizing Pixel
$X'' = \frac{X'}{255}$ #Normalizing Image
X_denoised = Filter(X") # Apply noise reduction (Gaussian or median filter)
X_a=Augment (X",{θ ,F ,C}) # Data augmentation
X_transformed = TransformColorSpace(X_augmented, RGB, HSV) # Apply color space transformation
Store X_transformed in preprocessed dataset I_preprocessed

Segmentation Stage
for each image X_segmented in I_preprocessed:
$Grid\ Cell\ Size = \frac{W}{S} \times \frac{H}{S}$ # Grid cell division
for each grid cell g:

Bounding Box= (x, y, w, h, c) # Bounding box parameters
Features = MultiScaleMapping(X_segmented) # Multi-scale feature mapping

Feature Extraction Stage
for each X_segmented in I_preprocessed:
$F = Backbone(I)$ # Extract feature maps
O_inception = Concat(Conv_1×1(F), Conv_3×3(F), Conv_5×5(F), MaxPool_3×3(F)) # Inception layers as per Equation (6)
$O_{\text{inception}} = Concat\ (Conv_{1\times1}(X), (Conv_{3\times3}(X), (Conv_{5\times5}(X), (MaxPool_{3\times3}(X)$ # Spatial Pyramid Pooling (SPP)
$F' = Attention(F)$ # Attention mechanism for refined features
Store refined feature maps F'
Feature Selection and Bounding Box Prediction
for each grid cell g:
Predict $B_i = \{x_i, y_i, w_i, h_i, c_i, \{p_i(c)\}_{c=1}^{C}\}$ # Bounding boxes
Refine AnchorBoxes to match counterfeit medicine dimensions # Optimize anchor boxes
$L = \lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} [(x_i - \hat{x_i})^2 + (y_i - \hat{y_i})^2] + \lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} \left[(\frac{w_i - \hat{w_i}}{\hat{w_i}})^2 + (\frac{h_i - \hat{h_i}}{\hat{h_i}})^2 \right] + \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{obj} (c_i - \hat{c_i})^2 + \lambda_{noobj} \sum_{i=1}^{S^2} \sum_{j=1}^{B} I_{ij}^{noobj} + (c_i - \hat{c_i})^2 + \sum_{i=1}^{S^2} I_{ij}^{obj} \sum_{c=1}^{c} (p_i(c) - \hat{p_i}(c))^2$ # Loss computation
Classification Stage
for each bounding box Bi:
$p_i(c) = \frac{e^{z_i(c)}}{\sum_{k=1}^{C} e^{z_i(k)}}$ # Class probabilities using Softmax
$\hat{c_i} = \arg \max_{c \in \{1, \ldots, c\}} p_i(c)$ # Predicted class label
FinalBoxes = NMS(Bi, ConfidenceThreshold) # Apply Non-Maximum Suppression (NMS)
Fine-tune YOLOv11 with PretrainedWeights and counterfeit dataset # Transfer learning for fine-tuning

Ensemble Verification
for each detected box in FinalBoxes:
Verify with ensemble model ClassifierEnsemble (e.g., SVM, Decision Trees)

End-to-End Pipeline Execution
Input I → Preprocessing → Segmentation → Feature Extraction → Feature Selection → Classification → Ensemble Verification → Output (Bounding Boxes and Labels)

4. RESULT AND DISCUSSION

The experimental setup utilized Python to preprocess synthetic images and partition the dataset into training, validation, and testing subsets. A Pareto Optimization-based Convolutional Neural Network (PAN-CNN) was implemented using the PyTorch framework. The model was trained on synthetic images as described in [18] and tested with datasets referenced in [16] and [17]. Hyperparameter optimization was conducted using random search or grid search techniques. The model's performance was evaluated using metrics such as accuracy, precision, and F1-score. Comparative analysis was performed between the proposed PAN-CNN and existing techniques, including One-Dimensional Convolutional Neural Networks (1D-CNN) [11], Support Vector Machine (SVM) [12], Random Forest (RF) [15], and Naïve Bayes (NB) [15]. The pre-processing procedures for training and testing images, as illustrated in Figure 2, aimed to standardize the inputs for optimal model performance.

Pre-processing in the experimental setup prepared synthetic images for training and testing by normalizing and augmenting the dataset. Images were resized to a fixed resolution to ensure compatibility with the input layer of the Pareto Optimization-based CNN (PAN-CNN). Noise reduction techniques like Gaussian filtering were applied to enhance image quality. Data augmentation methods, including rotations, flipping, and brightness adjustments, were used to increase dataset diversity and improve the model's generalization under varying conditions. Finally, pixel values were normalized to the range [0,1] for uniform gradient computation during training. This process ensured standardized inputs, enhancing model accuracy and robustness. The pre-processed image is given in Figure 2.

Figure 2 (a). Original Image

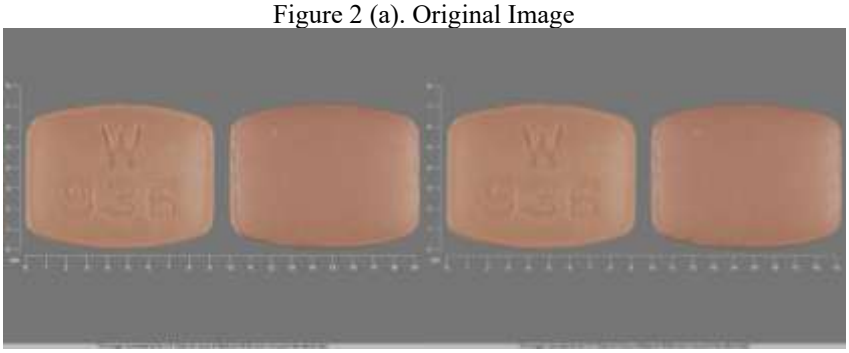

Figure 2 (b). Counterfeit Image

Figure 2. Pre-processing of Medicine Image

Segmentation in the setup involves dividing images into Regions of Interest (RoIs) for focused counterfeit detection. Using a grid-based approach, images are split into S×SS \times S cells, where each cell detects objects within its region. Attention mechanisms and multi-scale feature mapping further enhance precision by emphasizing critical counterfeit markers. The segmented image is given in Figure 3.

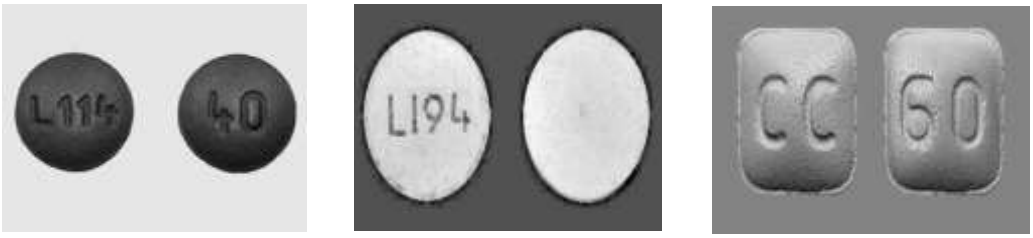

Figure 3 (a). Original Image

Figure 3 (b). Counterfeit Image

Figure 3. Segmentation of Medicine Image

Accuracy evaluates the overall correctness of the model by measuring the proportion of correctly predicted instances among all instances. It is quick to compute and suitable for balanced datasets but may not reflect performance for imbalanced classes. Precision assesses the quality of positive predictions, focusing on how many predicted positives are true. It is computationally light and critical in applications where false positives are costly. Recall measures the model's ability to identify all actual positives, emphasizing minimizing false negatives. It is time-efficient and vital in cases where missing positives has severe consequences, such as medical diagnoses.

F1 Score balances precision and recall by calculating their harmonic mean. It provides a more comprehensive metric, particularly for imbalanced datasets, but is slightly more time-consuming due to the harmonic mean calculation. Time Consumption: Metrics like precision, recall, and F1 score are computationally inexpensive for small datasets but scale with the size of predictions. Time consumption also depends on whether metrics are evaluated on the training or testing phase and the complexity of classifying instances. The performance is evaluated using Equation 20 – 24.

$$Accuracy = \frac{Number\ of\ Correct\ Prediction}{Total\ Number\ of\ Prediction} \text{-------(20)}$$

$$Precision = \frac{TP}{TP+FP} \text{---------(21)}$$

$$Recall = \frac{TP}{TP+FN} \text{---------(22)}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \text{---------(23)}$$

$$T = N \cdot t_{instance} + t_{aggregation} \text{---------(24)}$$

where TP, TN, FP, and FN denote True Positives, True Negatives, False Positives, and False Negatives, respectively. N is the total number of instances in the dataset (predictions), $t_{instance}$ is the average time to process a single instance (prediction comparison, confusion matrix updates, etc.), and $t_{aggregation}$ time to compute aggregated metrics such as accuracy, precision, recall, and F1-score from the confusion matrix.For real-time applications, optimizing computational efficiency is key.

Table 1. Comparison of Accuracy in %

Image Count	1D-CNN	SVM	RF	NB	YOLOv11	Mod-YOLOv11
100	71.12	72.33	70.03	72.11	76.23	81.23
200	72.23	72.53	71.56	72.99	76.89	81.91
300	72.99	73.33	73.11	73.14	77	82.09
400	73.17	74.63	74.57	74.77	77.23	82.9
500	73.23	74.65	75.31	75.71	77.34	83.23

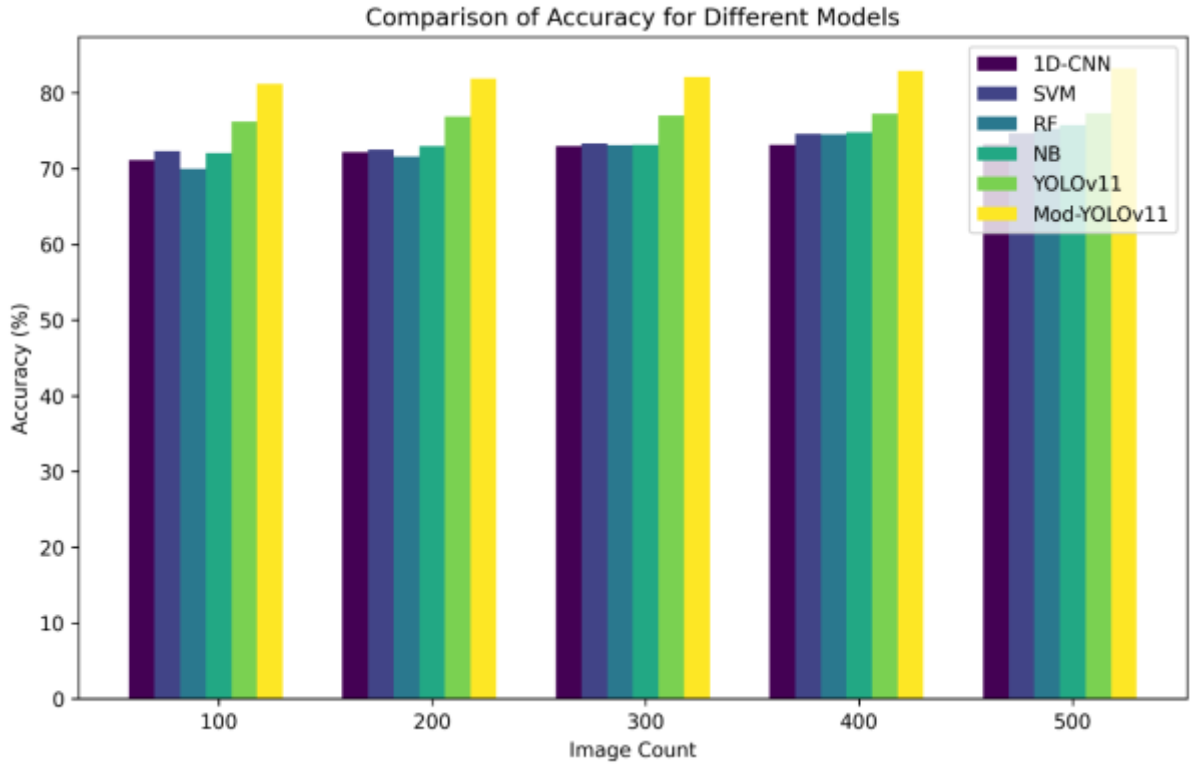

Figure 4. Comparison of Accuracy

Accuracy indicates the overall percentage of correctly classified instances. As seen in Table 1**,** Modified YOLOv11 consistently outperforms the other models across all image counts, with an accuracy increase as the image count rises. For instance, at 100 images, Mod-YOLOv11 achieves 81.23%, while 1D-CNN and SVM perform at 71.12% and 72.33%, respectively. At 500 images, Mod-YOLOv11 reaches an accuracy of 83.23%, while the best competitor, Random Forest (RF)**,** achieves 75.31%. Mod-YOLOv11 outperforms YOLOv11 in accuracy across all image counts, with a noticeable improvement. For example, at 100 images, YOLOv11 achieves 76.23% accuracy, while Mod-YOLOv11 reaches 81.23%. This trend continues, showing consistent enhancements in classification performance. This consistent superiority in accuracy shows that Mod-YOLOv11 is well-suited for large-scale image classification tasks, particularly in complex scenarios where the model needs to identify multiple objects in images with high precision. Figure 4 visually reinforces this, displaying a clear upward trend in Mod-YOLOv11's accuracy compared to the other models.

Table 2. Comparison of Precision in %

Image Count	1D-CNN	SVM	RF	NB	YOLOv11	Mod-YOLOv11
100	69.12	71.22	72.23	74.56	76.23	78.73
200	70.33	71.34	72.89	75.33	78.34	80.53
300	71.23	72.56	73.89	76.23	78.9	80.99
400	72.45	73.81	74.66	76.45	79.3	81.1
500	73.21	74.78	75.71	78.99	79.9	82.46

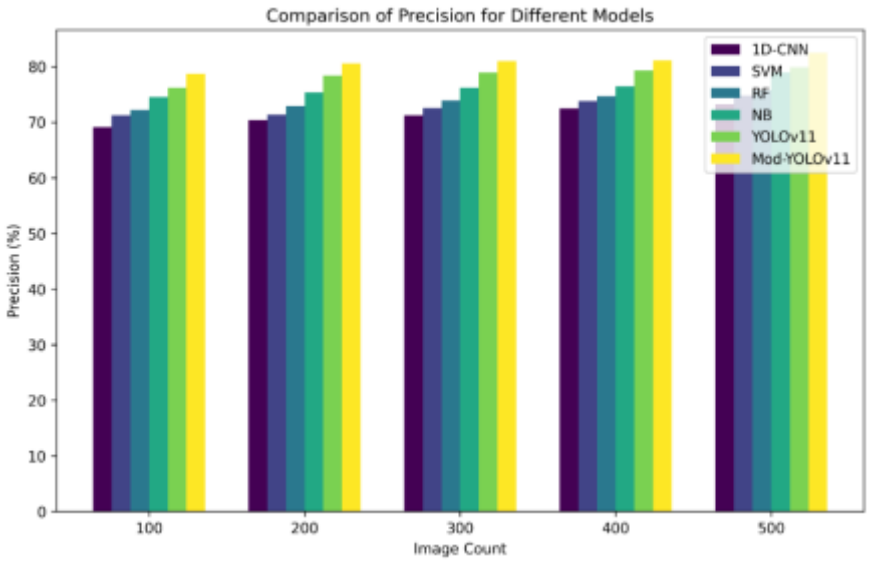

Figure 5. Comparison of Precision

Precision measures the proportion of positive predictions that are actually correct. According to Table 2, Mod-YOLOv11 achieves the highest precision across all image counts. At 100 images, Mod-YOLOv11 has 78.73%, while Naive Bayes (NB) and Random Forest (RF) are at 74.56% and 72.23%, respectively. As the image count increases, Mod-YOLOv11 maintains its leading position, reaching 82.46% precision at 500 images, compared to 1D-CNN at 73.21%. Mod-YOLOv11 consistently outperforms YOLOv11 in all the evaluated scenarios. The difference in accuracy ranges from 2.5% to 3.3% across various test conditions. This result is visually demonstrated in Figure 5, where Mod-YOLOv11 shows a steady increase in precision, surpassing other models as the dataset grows.

Table 3. Comparison of Recall in %

Image Count	1D-CNN	SVM	RF	NB	YOLOv11	Mod-YOLOv11
100	70.12	72.22	73.23	75.56	75.87	80.73
200	71.21	72.34	73.5	76.33	76.89	81.43
300	72.3	73.56	74.63	76.53	77.03	82.09
400	72.4	74.81	75.99	77.45	77.12	82.34
500	73.21	75.78	76.51	79.99	77.34	83.67

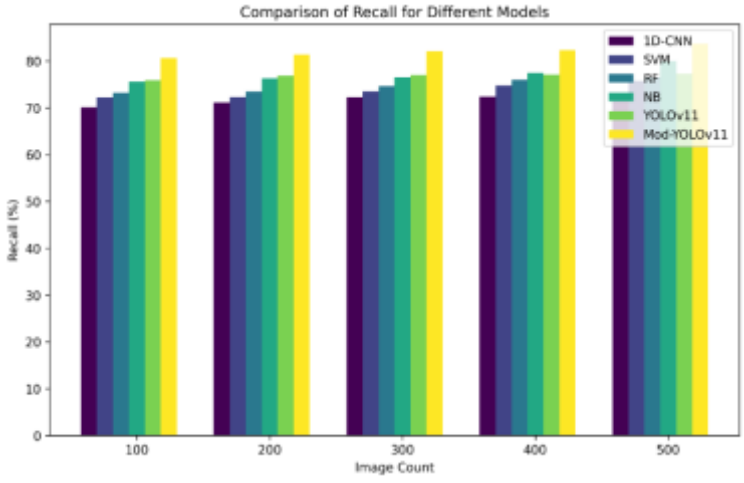

Figure 6. Comparison of Recall

Recall measures the ability of the model to correctly identify all relevant positive instances. In Table 3, Mod-YOLOv11 again leads in recall. For example, at 100 images, Mod-YOLOv11 achieves 80.73%, outperforming Naive Bayes at 75.56%. By 500 images, Mod-YOLOv11 reaches 83.67%, whereas Random Forest achieves 76.51%. The recall comparison between YOLOv11 and Mod-YOLOv11 shows an improvement in Mod-YOLOv11 across all image counts. YOLOv11 achieves a maximum recall of 77.34%, while Mod-YOLOv11 reaches 83.67%, demonstrating consistent enhancements in recall performance with the modified version for all tested image counts. The trend is visualized in Figure 6, where Mod-YOLOv11 shows a significant increase in recall over time, indicating its efficiency in minimizing false negatives.

Table 4. Comparison of F1-Score in %

Image Count	1D-CNN	SVM	RF	NB	YOLOv11	Mod-YOLOv11
100	71.89	74.11	76.33	79.33	79.4	79.9
200	72.12	76.81	76.44	80.01	79.9	80.23
300	72.67	77.81	78.01	80.67	80.2	81.31
400	72.98	77.99	78.45	80.99	81.01	81.98
500	73.12	78.4	79.90	81.35	81.44	82.82

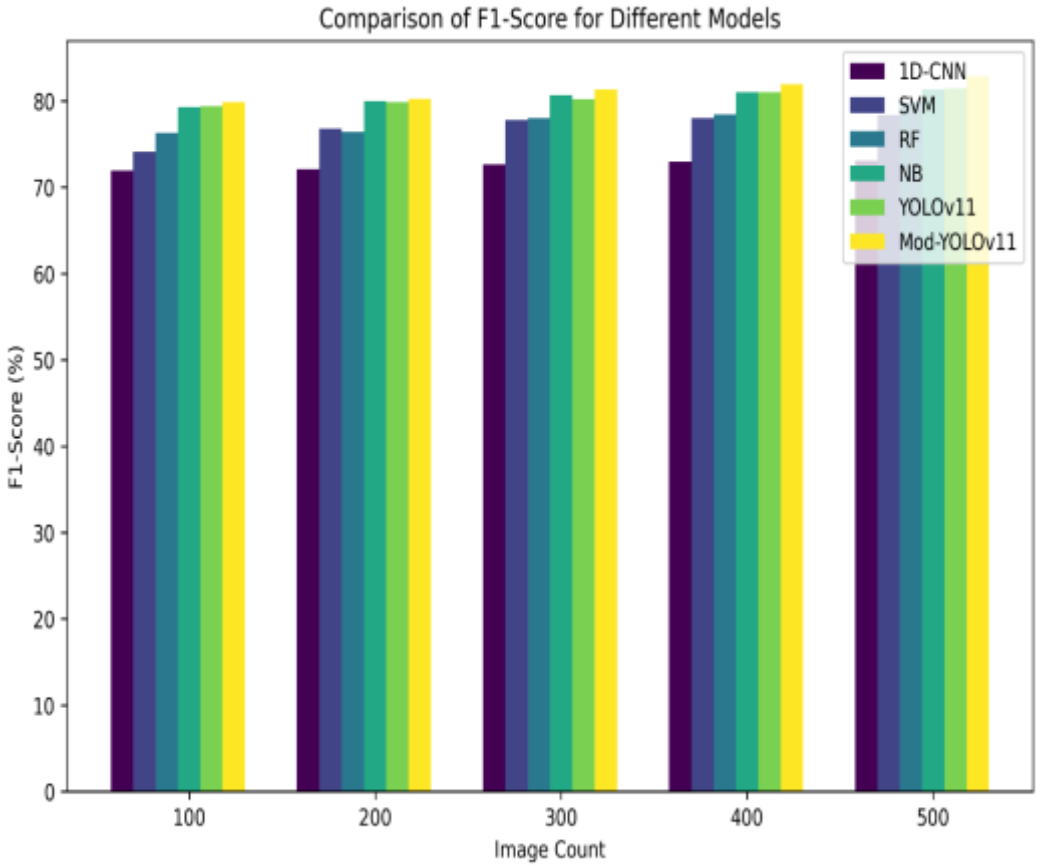

Figure 7. Comparison of F1-Score

The F1-score combines precision and recall into a single metric, providing a more balanced evaluation. As shown in Table 4, Mod-YOLOv11 consistently performs best in terms of the F1-score. At 100 images, Mod-YOLOv11 achieves 79.9%, and by 500 images, it reaches 82.82%. In comparison, SVM and Random Forest yield scores of 78.4% and 79.90%, respectively. For instance, at 100 images, YOLOv11 achieves an F1-Score of 79.4%, while Mod-YOLOv11 performs slightly better with 79.9%. The difference increases progressively, reaching an F1-Score of 81.44% for YOLOv11 at 500 images, compared to Mod-YOLOv11's 82.82%. Figure 7 confirms this trend,

where Mod-YOLOv11's F1-score progressively increases, reflecting its balanced performance between precision and recall, making it a highly effective model for image classification.

Table 5. Comparison of Time Consumption in MS

Epoch	1D-CNN	SVM	RF	NB	YOLOv11	Mod-YOLOv11
100	564	574	593	589	345	198
200	574	581	604	593	391	201
300	586	589	623	612	412	212
400	597	599	645	623	432	221
500	603	611	653	631	443	234

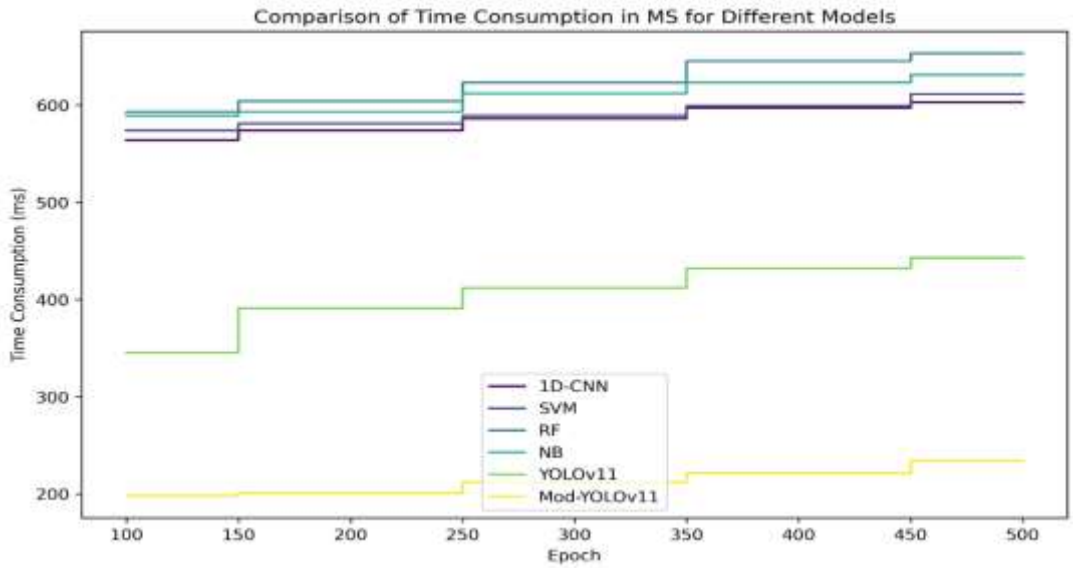

Figure 8. Comparison of Time Consumption

Time consumption, measured in milliseconds (ms), indicates the speed of each model in processing images. Table 5 shows that Mod-YOLOv11 processes images much faster than all other models. For example, at 100 images, it takes only 198 ms, while 1D-CNN and SVM take 564 ms and 574 ms, respectively. Even as the image count increases, Mod-YOLOv11 maintains a substantial speed advantage, reaching only 234 ms at 500 images, compared to Naive Bayes at 631 ms and Random Forest at 653 ms. Mod-YOLOv11 consistently outperforms YOLOv11 in terms of time efficiency. The time consumption for Mod-YOLOv11 decreases significantly across all epochs, with a reduction from 345 ms to 234 ms, compared to YOLOv11's 345 ms to 443 ms. This efficiency is crucial for real-time applications, and Figure 8 illustrates this with a clear gap between Mod-YOLOv11 and other models in terms of processing speed.

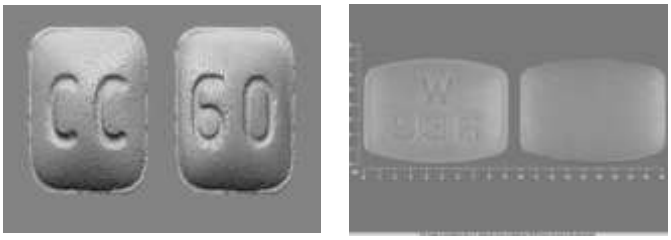

Figure 9 (a). Original Image

Figure 9 (b). Counterfeit Image

Figure 9. Classified Images using mod-YOLO

Figure 9 depicts the result of applying Mod-YOLOv11 on processing images of medicines. The images are classified into two categories: real and fake, from the learned features and the predicted bounding boxes. Every picture is labeled with four separate coordinates of the rectangular areas of interest and the respective probabilities of the picture being fake or not.

Conclusion

The proposed system for counterfeit medicine detection using modified YOLOv11 represents a significant advancement in combating the global issue of counterfeit pharmaceuticals. By leveraging a modified YOLOv11 architecture, which incorporates an efficient backbone network and an attention mechanism, the system enhances feature extraction and classification accuracy. This approach offers a reliable and real-time solution for identifying counterfeit products, even in adverse environments. The use of advanced techniques like adaptive spatial partitioning and efficient feature pyramid networks allows the system to handle high-definition images of medicines with minimal delay, ensuring that detection occurs swiftly and accurately. The preprocessing steps further improve model performance by increasing precision, recall, and F1-scores. By integrating lightweight architectures, the system minimizes computational complexity, making it feasible for real-time applications. Experimental results show an impressive accuracy of 83.23% and a quick processing time of 234ms, even with 500 epochs of training. This system provides a robust, practical solution to the pressing issue of counterfeit medicines, offering an effective method to safeguard public health while reducing the risks associated with fake drugs. The proposed model is not only innovative but also scalable, making it a promising tool for widespread deployment in healthcare settings.

REFERENCE

1. An-Bing, Z., Hui-Hua, Y., Xi-Peng, P., Li-Hui, Y., & Yan-Chun, F. (2020). On-site identification of counterfeit drugs based on near-infrared spectroscopy Siamese-network modeling. *IEEE Access*, *9*, 3195-3206.
2. Alsallal, M., Sharif, M. S., Al-Ghzawi, B., & al Mutoki, S. M. M. (2018, August). A machine learning technique to detect counterfeit medicine based on X-Ray fluorescence analyser. In *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)* (pp. 118-122). IEEE.
3. Ting, H. W., Chung, S. L., Chen, C. F., Chiu, H. Y., & Hsieh, Y. W. (2020). A drug identification model developed using deep learning technologies: experience of a medical center in Taiwan. *BMC health services research*, *20*, 1-9.
4. Carou-Senra, P., Ong, J. J., Castro, B. M., Seoane-Viano, I., Rodríguez-Pombo, L., Cabalar, P., ... & Goyanes, A. (2023). Predicting pharmaceutical inkjet printing outcomes using machine learning. *International Journal of Pharmaceutics: X*, *5*, 100181.
5. Ibrahim, A. M., Hendawy, H. A., Hassan, W. S., Shalaby, A., & ElMasry, M. S. (2020). Determination of terazosin in the presence of prazosin: Different state-of-the-art machine learning algorithms with UV spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *236*, 118349.
6. Banerjee, S., Sweet, J., Sweet, C., & Lieberman, M. (2016, March). Visual recognition of paper analytical device images for detection of falsified pharmaceuticals. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1-9). IEEE.
7. Ramos, R. R. T., Samonte, K. R. B., & Manlises, C. O. (2024, March). Medicine Authentication Based on Image Processing Using Convolutional Neural Networks. In *2024 16th International Conference on Computer and Automation Engineering (ICCAE)* (pp. 278-282). IEEE.
8. Islam, I., & Islam, M. N. (2022). Digital intervention to reduce counterfeit and falsified medicines: A systematic review and future research agenda. *Journal of King Saud University-Computer and Information Sciences*, *34*(9), 6699-6718.
9. Santos, M., Kahmann, A., Caffarate, L. M., Ucha, L. R., Limberger, R. P., & Ortiz, R. S. (2020). Counterfeit medicines: a pilot study for chemical profiling employing a different proposal of an usual technique. *Drug Analytical Research*, *4*(2), 19-23.
10. Rasheed, H., Höllein, L., & Holzgrabe, U. (2018). Future information technology tools for fighting substandard and falsified medicines in low-and middle-income countries. *Frontiers in pharmacology*, *9*, 995.
11. An-Bing, Z., Hui-Hua, Y., Xi-Peng, P., Li-Hui, Y., & Yan-Chun, F. (2020). On-site identification of counterfeit drugs based on near-infrared spectroscopy Siamese-network modeling. *IEEE Access*, *9*, 3195-3206.

12. do Prado Puglia, F., Anzanello, M. J., Scharcanski, J., de Abreu Fontes, J., de Brito, J. B. G., Ortiz, R. S., & Mariotti, K. (2021). Identifying the most relevant tablet regions in the image detection of counterfeit medicines. *Journal of Pharmaceutical and Biomedical Analysis*, *205*, 114336.
13. Sansone, A., Cuzin, B., & Jannini, E. A. (2021). Facing counterfeit medications in sexual medicine. A systematic scoping review on social strategies and technological solutions. *Sexual Medicine*, *9*(6), 100437-100437.
14. Ozawa, S., Evans, D. R., Bessias, S., Haynie, D. G., Yemeke, T. T., Laing, S. K., & Herrington, J. E. (2018). Prevalence and estimated economic burden of substandard and falsified medicines in low- and middle-income countries: a systematic review and meta-analysis. *JAMA network open*, *1*(4), e181662-e181662.
15. Thomson, B. S., & Varuna, W. R. (2023, October). Analyzing the Counterfeit Medicines Based on Classification Using Machine Learning Techniques. In *International Conference on Computer & Communication Technologies* (pp. 355-361). Singapore: Springer Nature Singapore.
16. https://go.drugbank.com/data_packages
17. https://www.drugs.com/image/famotidine-images.html
18. https://www.kaggle.com/datasets/vencerlanz09/pharmaceutical-drugs-and-vitamins-dataset-v2