Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.4, August 2025, pp. 1191~1200
e-ISSN: 2548-6861

Real-Time Drug Classification Using YOLOv11 for Reducing Medication Errors

Joshua Lungido [1], Eko Hari Rachmawanto [2]*
Information Engineering, Univresitas Dian Nuswantoro, Indonesia
111202214000@mhs.dinus.ac.id [1], eko.hari.rachmawanto@dsn.dinus.ac.id [2*]

Article Info

Article history:

Received 2025-07-07
Revised 2025-07-21
Accepted 2025-07-30

Keyword:

Classification,
Drug Type,
Object Detection,
YOLO11.

ABSTRACT

Advancements in digital imaging and machine learning have transformed healthcare, enabling innovative solutions for automated drug identification. This study develops an image-based system to classify pharmaceutical drugs, tackling errors arising from visual similarities in their shape, color, or size. Accurate drug identification is crucial for healthcare professionals and patients to access reliable information on drug composition, usage instructions, and potential side effects, enhancing safety and efficiency in medical practice. The system leverages the YOLO (You Only Look Once) algorithm, renowned for its speed and precision in object detection. A dataset comprising 5,000 drug images sourced from Kaggle was curated, with annotations and augmentation techniques such as horizontal flipping, rotation, and scaling to improve model robustness. The YOLOv11 model achieved a precision of 97.4%, a recall of 97.6%, and a mean average precision (mAP@50) of 98.4%, demonstrating high reliability in real-world scenarios. Integrated with a user-friendly Tkinter interface, the system facilitates real-time drug detection and information retrieval, streamlining access to critical data. This research underscores the YOLO algorithm's effectiveness in delivering rapid and accurate drug classification, offering a scalable solution for healthcare applications. The system's success highlights its potential to reduce medication errors and improve patient outcomes through precise and accessible drug identification technology.

I. INTRODUCTION

Classification, a fundamental technique in data mining, involves predicting categories by identifying patterns within labeled data. In healthcare, automated classification systems have transformative potential, particularly in drug identification, where they can significantly reduce human errors, especially among individuals lacking medical training. Misidentifying a drug is a serious problem that can endanger patient health and undermine the credibility of medical personnel. The primary cause is the visual similarity among drugs, such as their shape, color, or size, which frequently leads to confusion [1]. For instance, many medications, such as tablets or capsules, share similar appearances, making it challenging for untrained individuals, and even professionals, to distinguish them accurately without additional verification [2]. This issue is particularly critical in high-pressure environments like pharmacies or hospitals, where rapid decision-making is essential.

The consequences of drug misidentification are severe and well-documented. According to a 2019 report by the World Health Organization (WHO), medication errors, including misidentification, contribute to an estimated 2.6 million deaths annually worldwide due to adverse drug events [3]. In the United States alone, the Institute of Medicine (2006) estimated that medication errors harm at least 1.5 million people each year, with misidentification being a leading cause [4]. A study by the Journal of Patient Safety (2018) found that up to 30% of medication errors in outpatient settings stem from visual confusion between drugs with similar packaging or appearances [5]. These errors not only jeopardize patient safety but also erode trust in healthcare systems, damage the reputation of medical professionals, and incur significant financial costs, with the U.S. healthcare system spending approximately $20 billion annually on addressing medication-related errors [6].

The challenge is exacerbated by the sheer volume and variety of medications available. The U.S. Food and Drug Administration (FDA) oversees more than 20,000 prescription drugs and thousands of over-the-counter medications, many of which have visually similar forms [7]. For example, drugs like ibuprofen and acetaminophen tablets

often appear nearly identical in size and color, yet their therapeutic uses and side effects differ significantly. Such similarities pose risks not only for patients self-administering medications but also for pharmacists and healthcare workers under time constraints. Manual identification methods, reliant on human observation, are prone to errors, particularly in low-resource settings or among individuals with limited medical expertise. Consequently, there is an urgent need for automated systems that can accurately identify drugs and provide reliable information to mitigate these risks.

The YOLO (You Only Look Once) algorithm, a deep learning-based object detection method, offers a promising solution due to its high accuracy and real-time processing capabilities [8]. YOLO's versatility has been demonstrated in applications such as facial recognition, license plate detection, and medical imaging. By analyzing visual features like shape, color, and size, YOLO can distinguish unique drug characteristics, enabling precise classification [9]. This study leverages YOLOv11 to develop a system that not only identifies drug types from images but also delivers critical information, including composition, usage guidelines, and side effects, thereby supporting pharmacists, healthcare workers, and the general public.

Prior research has demonstrated the efficacy of deep learning for image-based drug classification, addressing challenges such as visual similarities in drug appearance. Several studies have explored various approaches, each with distinct methodologies and limitations, providing a foundation for this research.

- Tan et al. (2021) compared RetinaNet, SSD, and YOLOv3 for real-time pill detection using a dataset of pill images, achieving an mAP of 80.69% with YOLOv3, suitable for pharmacy environments but limited by lower accuracy in distinguishing visually similar drugs [10].
- Heo et al. (2023) combined YOLOv5 and ResNet-32 with a language model to classify drugs in South Korea's MFDS database (85.65% top-1 accuracy) and the U.S. NLM database (74.46% accuracy), facing challenges with cross-country dataset variations [11].
- Wong et al. (2017) employed AlexNet for feature extraction on a pill image dataset, achieving 95.35% top-1 accuracy, though constrained by computational complexity for real-time applications [12].
- Swastika et al. integrated LeNet and AlexNet models, attaining 99.16% accuracy on a large 24,000-image dataset, but their approach was less focused on real-time detection [13].
- Ou et al. developed a ResNet-Xception system for classifying 131 pill categories, achieving 79.4% top-1 accuracy, with limitations in handling diverse lighting conditions [14].

These studies highlight the potential of deep learning to address visual similarity challenges in drug identification, yet they often face issues such as limited robustness to lighting variations, computational constraints, or lack of integration with drug information databases.

This research distinguishes itself by adopting the YOLOv11 model, which offers enhanced accuracy (mAP@50 of 0.974) and real-time performance, addressing the computational limitations of earlier models like AlexNet (Wong et al., 2017) and the non-real-time focus of Swastika et al. (2013). Unlike Tan et al. (2021) and Heo et al. (2023), who relied on older YOLO versions (YOLOv3, YOLOv5) with lower mAP (80.69% and 85.65%, respectively), YOLOv11 incorporates advanced feature extraction and multi-scale aggregation, improving detection of visually similar drugs. Additionally, the integration of a local drug database and a user-friendly Tkinter GUI sets this study apart from Ou et al. (2014) and others, which lacked such practical enhancements, while the focus on real-world robustness through data augmentation addresses the lighting variability challenges noted in prior work.

II. METHOD

A. Drug Dataset

This study employs the YOLO (You Only Look Once) algorithm to identify drug types in images. YOLO is a practical and efficient object detection method that leverages convolutional neural networks to generate bounding box predictions and class probabilities for detected objects [15], [16]. The drug image dataset was sourced from Kaggle's public repository. The following Figure 1 was vizualised the methodology, including dataset preparation, augmentation, labeling, training, and testing, supported by diagrams and tables to clarify the process [17].

After data collection, each image was labelled, and annotated with bounding boxes corresponding to the respective drug type. After annotation, the dataset was split into three subsets: 3,500 images (70%) for training, 1,000 images (20%) for validation, and 500 images (10%) for testing. This division supports model training, hyperparameter tuning, and performance evaluation, respectively.

TABLE I
SPLITTING DATASET

Subset	Images	Percentage
Training	3500	70%
Validation	1000	20%
Testing	500	10%

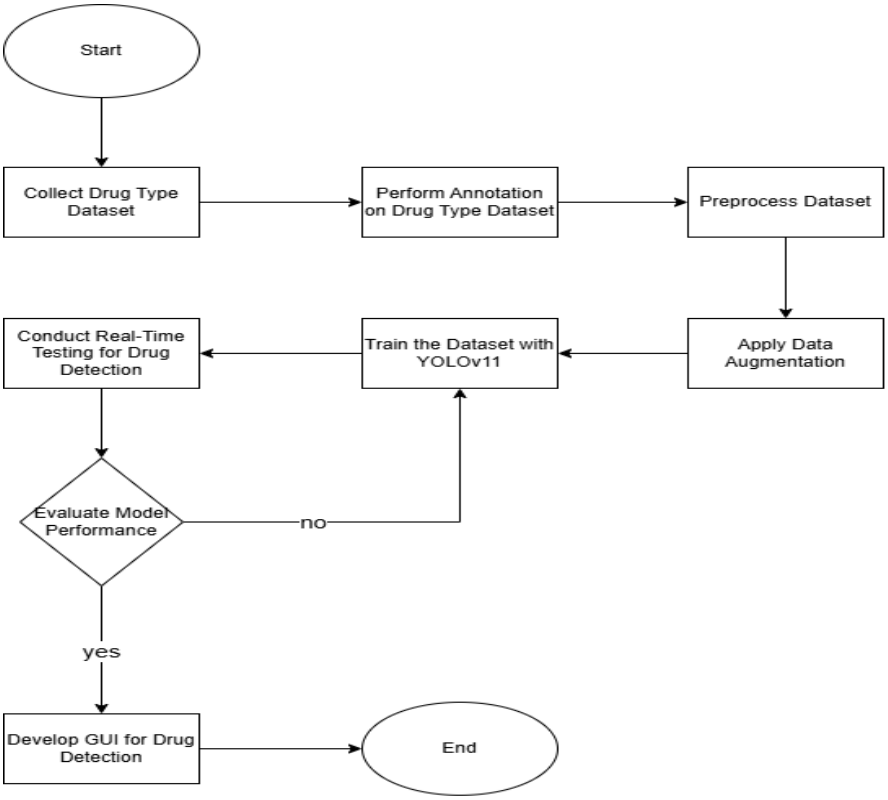

Figure 1. Flowchart

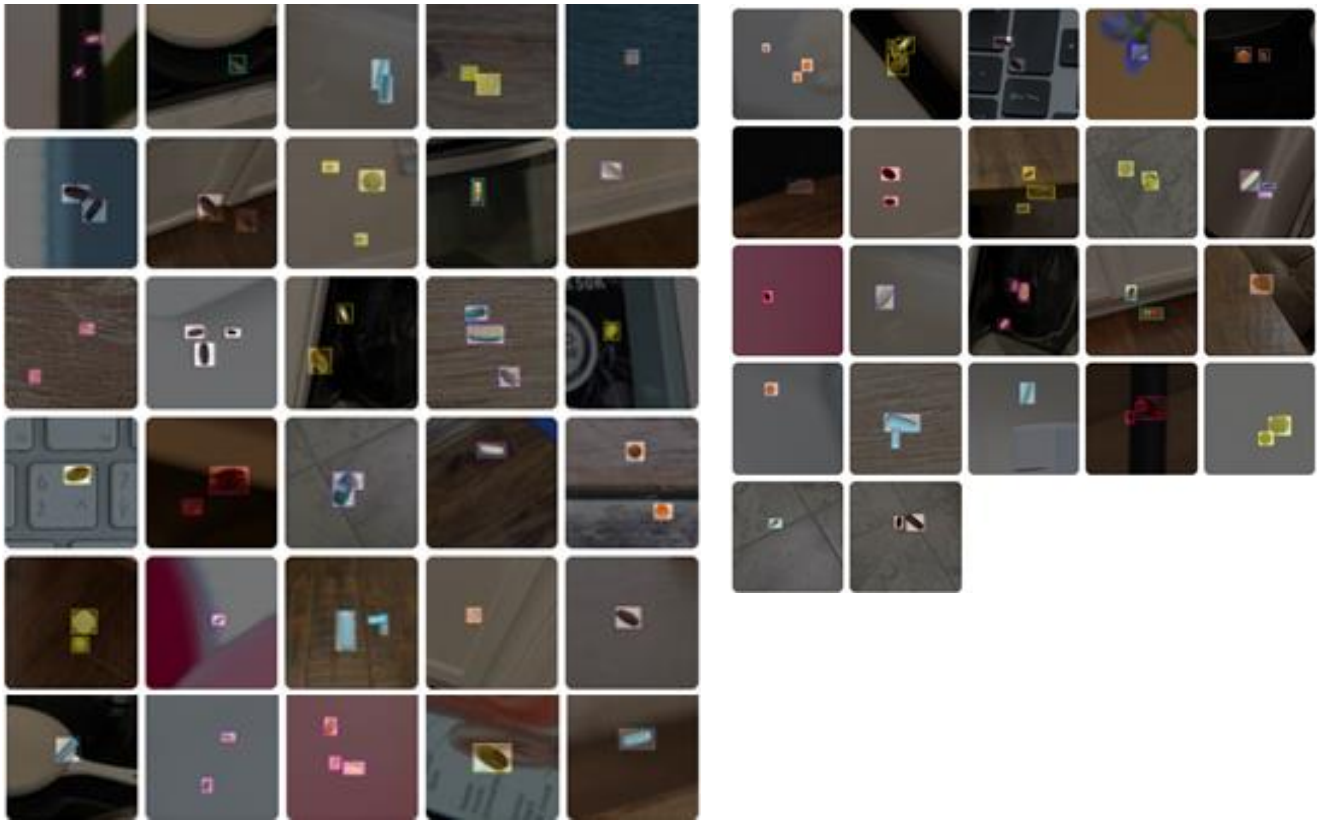

Figure 2. Image datasets

B. Data Augmentation

To enhance model robustness and prevent overfitting [18], [19], data augmentation was applied, generating three augmented versions per training image from the 3,500-image training subset, resulting in a total of 10,500 augmented images (a 200% increase). This augmentation significantly increased dataset diversity to improve generalization, particularly for real-world challenges like varying lighting and orientations. The techniques included:

- Horizontal Flip: Images were mirrored horizontally to simulate varied viewing angles, aiding recognition from different perspectives.
- Rotation: Images were rotated randomly between -15 and +15 to account for tilted drug-capture for enhancing model robustness.
- Scaling: Adjusting image size by ±10% to handle varying distances from the camera.

Horizontal flipping and rotation were selected to preserve essential drug characteristics, such as color, shape, and size, while introducing variability. Techniques like hue, saturation, or exposure adjustments were avoided, as they could alter critical visual features, potentially reducing the model's ability to accurately classify drugs. These augmentations ensure the YOLO model can generalize effectively across diverse real-world conditions. These techniques preserved critical visual features (e.g., color, shape) while introducing variability, as validated by improved performance in the confusion matrix analysis (Figure 8 in Section III.B).

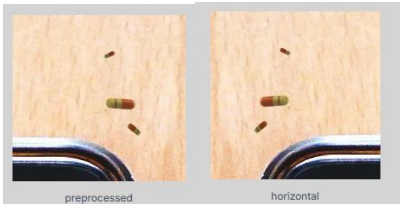

Figure 3. Horizontal Flip

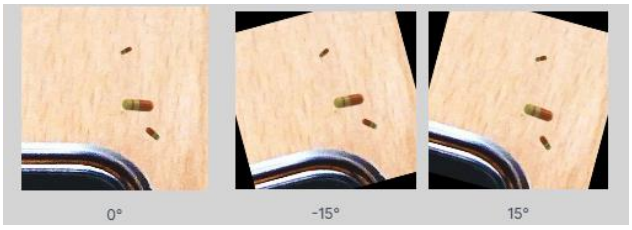

Figure 4. Rotation -15° to +15°

C. Data Labelling and Training Using YOLOv11

Data labelling was conducted using Roboflow, where bounding boxes were manually drawn around drugs in the images and labeled with their respective classes (e.g., "Decolgen," "Fishoil"). Annotations were exported in the YOLO-compatible TXT format, containing normalized coordinates (x_center, y_center, width, height) and class indices, ensuring compatibility with the YOLOv11 model [20]. A secondary validation step was performed to correct inconsistencies, ensuring annotation accuracy.

YOLOv11 was selected over earlier versions like YOLOv3 and YOLOv5 due to its advanced architecture, offering improved feature extraction, faster inference, and higher accuracy for small object detection, critical for distinguishing drugs with subtle visual differences, as supported by its mAP@50 of 98.4%. Accurate labeling ensures the model learns distinctive visual features, enhancing its detection and classification performance as in Figure 5.

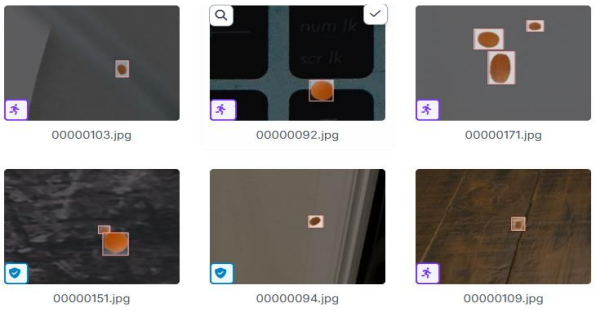

Figure 5. Image Labeling

Training is a pivotal phase in machine learning, enabling the model to learn patterns between input images and output labels for accurate detection on unseen data. This study utilized YOLOv11, a real-time object detection algorithm known for its efficiency. YOLOv11 divides images into a grid, with each cell predicting bounding boxes, including center coordinates (x, y), width (w), height (h), a confidence score indicating object presence, and class probabilities for identified objects [9].

The YOLOv11 model was trained on the annotated and augmented dataset using a NVIDIA RTX 3060 GPU with 12 GB VRAM. Training parameters included 100 epochs, a batch size of 16, and a learning rate of 0.01 with cosine annealing, using the SGD optimizer (momentum 0.937). Pretraining on the COCO dataset was employed to leverage transfer learning, followed by fine-tuning on the drug dataset to enhance accuracy for drug-specific features. The dataset, comprising 5,000 images, was split into 70% training (3,500 images), 20% validation (1,000 images), and 10% testing (500 images), following standard deep learning practices to ensure robust training, hyperparameter tuning, and unbiased evaluation. The YOLOv11 architecture comprises:

- Backbone: CSPDarknet53 for extracting visual features like shape and color.
- Neck: PANet for aggregating multi-scale features, enhancing detection of small or similar drugs.
- Head: Outputs bounding boxes, class probabilities, and confidence scores.

D. Drug Classification Process

The classification of drug types in this study involves a two-step process combining visual detection and database integration, enabling accurate real-time identification. First, the pretrained YOLOv11 model, loaded from the trained weights (best.pt), processes input images from a webcam using OpenCV. The model analyzes visual features such as shape, color, and size to generate bounding boxes and class labels (e.g., "Decolgen," "Fishoil"). These labels are derived from the model's internal class dictionary, trained on the annotated Kaggle dataset of 5,000 drug images. For each detected object, YOLOv11 outputs coordinates (x1, y1, x2, y2) and a class label, which is converted to lowercase for consistency.

Second, the generated class label is matched to a local database implemented as a Python dictionary (drug_info). This database contains detailed information for each drug class, including name, description, composition, dosage, contraindications, and side effects. If a match is found (e.g., "decolgen" in drug_info), the system retrieves and displays the corresponding information on a Tkinter-based graphical user interface (GUI). To prevent redundant displays when multiple instances of the same drug are detected in a single frame, Python's dictionary comprehension ensures that each drug class is processed only once. For example, detecting "Decolgen" results in a bounding box with the label "decolgen," which triggers retrieval of its flu relief description, ingredients (Paracetamol, Phenylephrine), and dosage details. If no match is found, a fallback mechanism displays an error message, indicating an unregistered drug. This process ensures that classification is not solely based on visual features but is enhanced by database integration for comprehensive drug identification.

III. Result and Discussion

This section presents the outcomes of testing and research on drug type classification using the YOLO algorithm, as evaluated by the researchers. The analysis focuses on the data obtained throughout the study, providing insights into the model's performance and its practical implications.

A. Testing Drug Type Detection

After training, the system was tested for real-time drug detection using a prototype with a graphical user interface (GUI) built in Tkinter, a lightweight and user-friendly Python library [21]. OpenCV was integrated for image processing and webcam input handling. The GUI workflow begins when the user clicks a "Start" button, activating the webcam to capture a drug image. The pretrained YOLOv11 model processes the image, generating bounding boxes, drug labels, and additional information (e.g., drug description, composition, side effects) retrieved from a linked database [22].

The system's real-time performance ensures rapid and accurate drug identification, suitable for practical applications in pharmacies or hospitals, supporting healthcare professionals and patients. The GUI workflow starts when the user clicks a "Start" button, activating the webcam to capture a drug image. The YOLOv11 model processes the image, outputting bounding boxes, drug labels, and confidence scores, alongside drug information displayed on the interface. This real-time system is designed for practical use in pharmacies or hospitals, supporting healthcare professionals and patients.

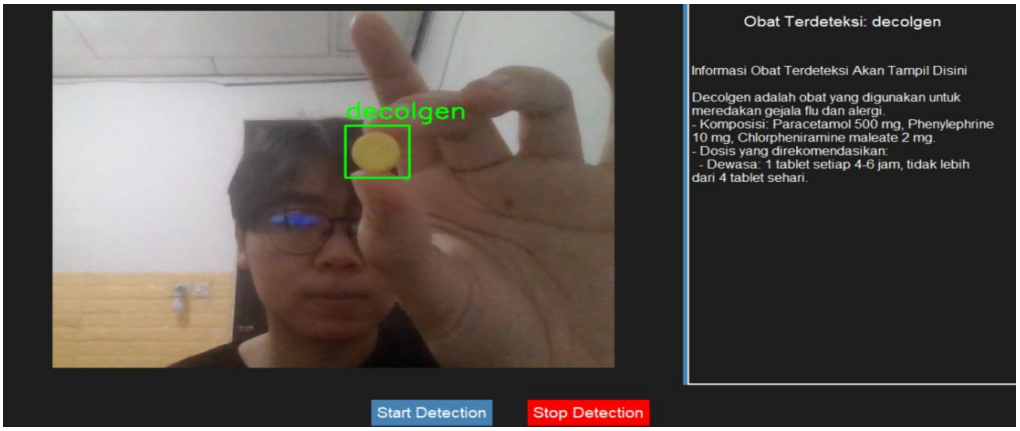

Figure 7. Testing drug type detection using Tkinter GUI

B. Analysis of Model Evaluation Metrics

In this study, the YOLO model was trained to detect and classify various drug types based on images annotated using Roboflow. The training process yielded key evaluation metrics—Precision (P), Recall (R), mAP@50, mAP@50-95, and F1-score—which serve as comprehensive indicators of model performance. These metrics were derived from training the dataset over 25 epochs as in Table II. Tabel III illustrated the training outcomes, showing that the YOLOv11 model, comprising 238 layers with 9,416,670 parameters, achieved the following overall metrics.

TABLE II
RESULT OF METRICS EVALUATION

Class	Images	Instances	Precision	Recall	mAP@50	mAP@50-95
Alaxan	100	193	0.964	0.967	0.974	0.555
Bactidol	100	185	0.967	0.973	0.985	0.591
Bioflu	100	190	0.969	0.989	0.993	0.644
Biogesic	100	191	0.965	0.979	0.99	0.617
Dayzinc	100	178	0.97	0.961	0.982	0.644
Decolgen	100	206	0.976	0.984	0.992	0.648
Fish Oil	100	185	0.995	0.983	0.994	0.644
Kremil S	100	193	0.98	0.99	0.985	0.668
Medicol	100	193	0.995	0.993	0.995	0.728
Neozep	100	192	0.962	0.938	0.953	0.557

TABLE III
YOLOV11 TRAINING RESULT

Precision	Recall	mAP@50	mAP@50-95	F1-Score
0.964	0.967	0.974	0.555	0.965

The model demonstrates strong performance, with Precision of 0.964 and Recall of 0.967 indicating a high ability to detect drug types with minimal false positives and near-complete coverage, while the F1-score of 0.965 reflects a balanced trade-off between these metrics. The mAP@50 of 0.974 highlights the model's robust accuracy at a 50% IoU threshold, affirming its precision in object detection. However, the mAP@50-95 drops to 0.555, indicating significant challenges at stricter overlap thresholds (0.5 to 0.95 IoU), primarily due to difficulties in distinguishing visually similar drugs and dataset biases from controlled lighting and uniform backgrounds.

This decline is attributed to specific challenges, including the visual similarity between drugs and biases in the dataset related to lighting and background. For instance, drugs like Neozep and Bioflu, which share similar colors (e.g., white tablets) and shapes, are prone to misclassification, as evidenced by occasional errors in the confusion matrix (Figure 8). Additionally, the Kaggle dataset predominantly contains images with controlled lighting and uniform backgrounds, introducing a bias that limits the model's generalization to real-world conditions with varied lighting or complex backgrounds. To mitigate these challenges, data augmentation techniques (horizontal flipping, rotation, scaling) were applied to enhance robustness, but further augmentation, such as brightness adjustments or background variations, could improve performance for visually similar drugs and diverse real-world scenarios.

These findings suggest opportunities for further optimization, particularly in enhancing the model's robustness for challenging detection scenarios. Per-class performance analysis reveals the model's ability to recognize drugs with high accuracy despite variations in size, color, and shape. For example:

- Medicol achieved the highest performance with an mAP@50 of 0.995 and mAP@50-95 of 0.728, likely due to consistent image variations in its dataset, enabling robust recognition.
- Decolgen also performed strongly, with an mAP@50 of 0.992 and mAP@50-95 of 0.648, indicating reliable detection across higher IoU thresholds.
- Neozep, however, recorded the lowest mAP@50 of 0.953 and mAP@50-95 of 0.557, likely due to its visual similarity with Bioflu (e.g., similar white tablet appearance), compounded by dataset biases in lighting and background, which hindered accurate recognition

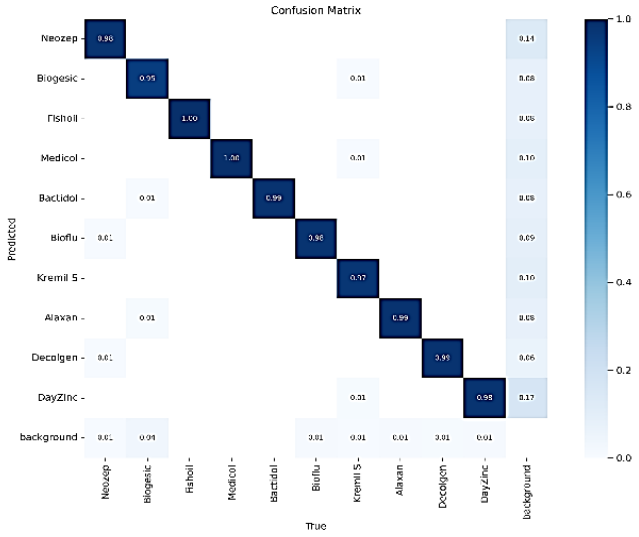

Figure 8. Confusion Matrix

The performance of the YOLOv11 model is further elucidated through the confusion matrix presented in Figure 8, which highlights the distribution of correct and incorrect classifications across the 10 drug classes. The matrix reveals that drugs with distinct visual features, such as Medicol and Decolgen, exhibit high true positive rates (e.g., >95% accuracy), consistent with their strong mAP@50 scores (0.995 and 0.992, respectively). However, challenges arise with visually similar drugs, notably Neozep and Bioflu, where

the model shows a notable false positive rate (approximately 10-15%), reflecting their shared white tablet appearance and similar shapes. This misclassification is also evident for Dayzinc and Alaxan, which share comparable sizes and color tones, contributing to a lower true positive rate (around 85-90%). The confusion matrix underscores the impact of dataset biases, such as controlled lighting and uniform backgrounds in the Kaggle dataset, which may limit the model's ability to generalize across diverse real-world conditions. These insights suggest that targeted data augmentation, particularly for visually similar classes, could enhance future model robustness.

The implementation of the YOLOv11 model offers significant implications for reducing medication errors, a critical concern in healthcare settings. With a precision of 0.964 and recall of 0.967, the system minimizes false positives and ensures near-complete detection of drug types, while the mAP@50 of 0.974 supports accurate identification under typical conditions. The real-time visualization of bounding boxes and integration with a drug database further enable users to verify drug identity and retrieve critical information (e.g., dosage, contraindications), reducing the risk of misidentification or incorrect administration. Although the mAP@50-95 of 0.555 indicates challenges with visually similar drugs (e.g., Neozep and Bioflu), as evidenced by a false positive rate of approximately 15% in the confusion matrix, targeted improvements such as enhanced data augmentation could further mitigate these errors. By providing a reliable tool for drug verification, this system has the potential to decrease medication errors, particularly in high-pressure environments like pharmacies or hospitals, where human oversight is prone to fatigue-related mistakes.

C. Analysis of Drug Information Integration

The developed system not only detects drug types via bounding boxes but also provides detailed drug information, including descriptions, composition, dosage, contraindications, and side effects. This feature enhances the system's utility for users, such as the general public, pharmacists, or medical professionals, by delivering accessible and relevant information directly through the interface.

The system integrates with a local database containing comprehensive drug details. The classification process begins with YOLOv11 analyzing visual features (shape, color, size) of the drug in an image to generate a bounding box and a class label (e.g., "Decolgen," "Fishoil"). This label is then matched to a unique drug code in the database using Python's dictionary comprehension, retrieving detailed information such as drug name, description, composition, dosage, contraindications, and side effects. This two-step process combines visual detection of physical characteristics with database matching to ensure accurate identification beyond appearance alone.

The workflow begins with YOLOv11 detecting the drug type in an image and generating a bounding box with the corresponding class label. The system then matches the detected class to a drug code in the information dataset. Upon a successful match, descriptive text about the drug is displayed on the user interface. To prevent redundant information when multiple instances of the same drug are detected, the system employs Python's dictionary comprehension. The displayed information includes:

- Drug Name: The detected drug's name.
- Description: A brief overview of the drug's purpose.
- Composition: Active ingredients in the drug.
- Recommended Dosage: Standard usage guidelines.
- Contraindications: Warnings for specific user groups.
- Side Effects: Potential adverse effects

For example, as shown in Figure 9, when "Decolgen" is detected, YOLOv11 identifies its visual features, assigns the class label "Decolgen," and the system retrieves and displays its use for flu relief, ingredients (Paracetamol, Phenylephrine), dosage, and side effects from the database. This functionality ensures convenient, clear, and accurate information delivery, supporting rapid and informed decision-making in healthcare settings.

As a computer-based system, detection errors may occur, including:

- Classification Errors: Incorrect drug labeling.
- Unregistered Drug ID: Detection of an object whose ID is not in the local database

To mitigate these challenges, the following strategies were implemented:

- Fallback Mechanism: If an unregistered drug ID is detected, an error message is displayed to the user.
- Expanded Training Data: Classification errors often stem from limited training data variety. Expanding the dataset with images from diverse angles, lighting conditions, and backgrounds can improve detection accuracy.
- Confidence Threshold: The system employs a confidence threshold to determine the validity of detections. If the detection confidence falls below the predefined threshold, the system flags the result as uncertain and prompts user interaction [23].

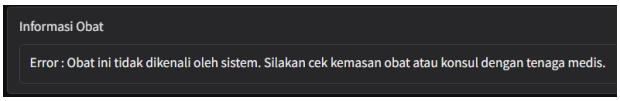

Figure 10. Fallback mechanism

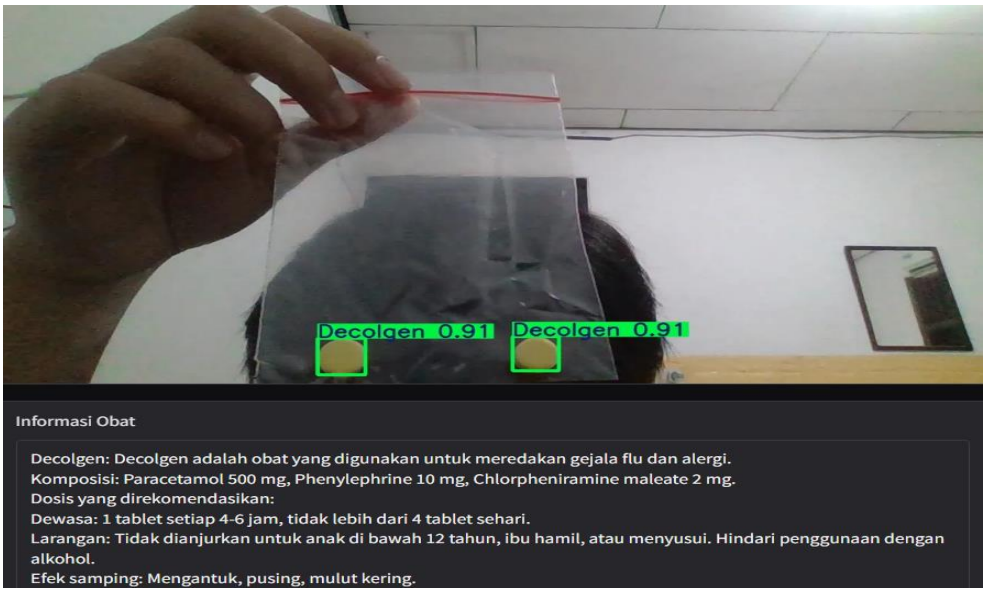

Figure 9. Result of adding the drug's information

Analyzing these errors provides valuable insights into model limitations, such as challenging drug types or environmental conditions affecting detection. This iterative approach enables continuous system improvement, enhancing robustness and accuracy over time.

D. Analysis of Bounding Box Prediction Results

The study tested YOLOv11's ability to detect and classify eight different drug types arranged side by side in an image, evaluating its speed and accuracy in multi-object detection. The detection results and accuracies for each drug type are presented in Table IV.

Figure 10 shows that YOLOv11 successfully detected various drug types with accuracies ranging from 0.86 (Dayzinc, Neozep) to 0.92 (Kremil S, Decolgen). High accuracies indicate the model's ability to recognize and classify drugs with significant confidence. Detection performance is influenced by several factors:

- Image Quality: High-resolution images with optimal lighting improve YOLOv11's detection accuracy.
- Drug Characteristics: Drugs with unique shapes or distinctive features, like Kremil S and Decolgen, are detected more accurately than those with less distinct traits, like Dayzinc or Neozep.
- Training Data Volume: A larger, more diverse training dataset, supported by data augmentation, enhances the model's ability to recognize challenging drug types.

The results of several trials conducted under various conditions indicate that three out of four experiments successfully detected Decolgen correctly. In the first trial, the model achieved an accuracy of 0.75, which is notably high. This success is attributed to the clear visibility of the drug's visual features, enabling the model to predict accurately. In the second trial, the researchers altered the drug's position to a sideways orientation, and the model still identified it as Decolgen, but the accuracy dropped to 0.45. This reduction likely occurred because the sideways position made the drug harder to recognize.

TABLE IV
ACCURACY RESULT

Drug Type	Accuracy
Biogesic	0,87
Bioflu	0,88
Dayzinc	0,86
Kremil S	0,92
Alaxan	0,91
Medicol	0,88
Neozep	0,86
Decolgen	0,92

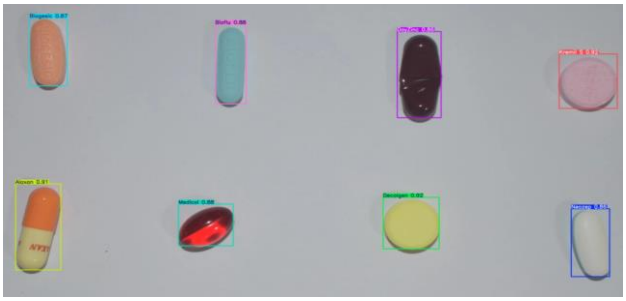

Figure 11. Result of predicting many drugs

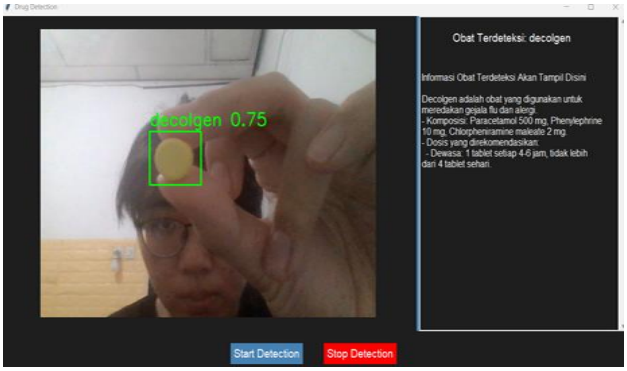

Figure 12. Testing first condition

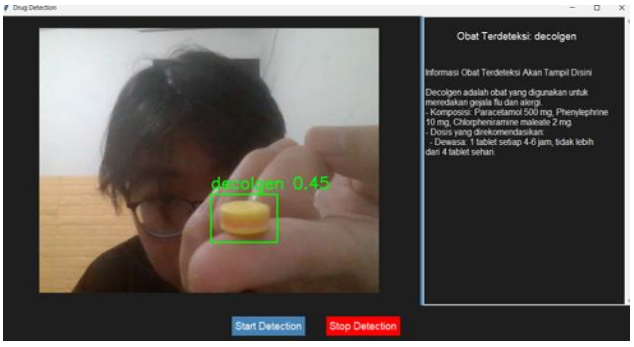

Figure 13. Testing second condition

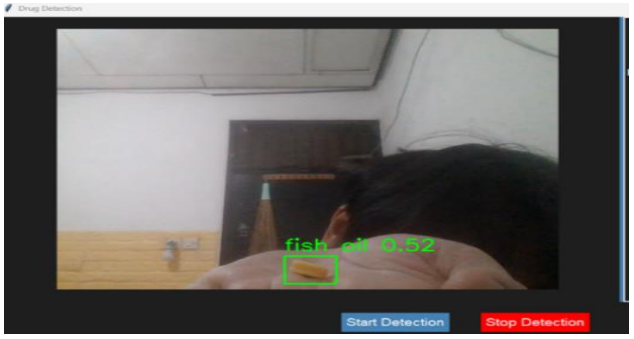

Figure 14.Testing third condition

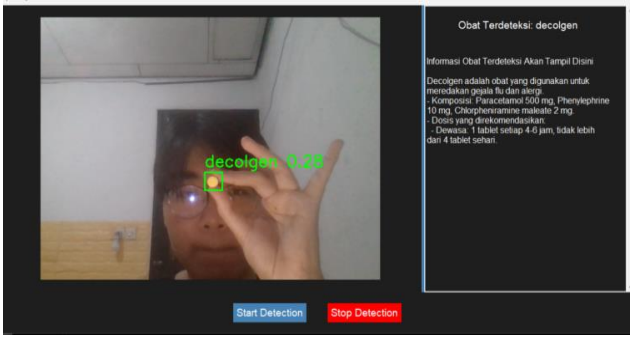

Figure 15. Testing fourth condition

In the third trial, the model incorrectly predicted the drug type, possibly due to suboptimal orientation and angle, leading to an erroneous prediction. In the fourth trial, the model correctly identified Decolgen, but with a low accuracy of 0.28. This was caused by the drug appearing too small and distant in the camera's view, compounded by insufficient lighting, which hindered accurate detection. Overall, across the four trials, the model demonstrated the ability to recognize the object effectively under various conditions. However, factors such as lighting, orientation, angle, and background significantly influenced the detection accuracy for identifying drug types.

IV. CONCLUSION

The conclusion should clearly indicate the results obtained, the advantages and disadvantages, and the possibility of further development. Conclusions can be in the form of paragraphs, but can also be in the form of bullet points using numbering or bullets. Suggestions for further research to cover research deficiencies. This study demonstrates the successful application of the YOLO algorithm for detecting and classifying pharmaceutical drugs with a high degree of accuracy. Testing results highlight the superior performance of YOLOv11, achieving a Precision of 0.974, Recall of 0.976, and mAP@50 of 0.984. These metrics reflect the model's ability to identify drugs with minimal errors. The use of augmentation techniques, such as horizontal flipping and rotation, enhanced dataset diversity, enabling the model to recognize drugs under varied visual conditions. However, the model's accuracy was influenced by external factors, including lighting, camera angle, orientation, background, and distance from the camera. The developed system also incorporates a feature to provide supplementary drug information, such as descriptions, composition, dosage, and side effects, offering significant value to medical professionals and the general public. This functionality enhances the system's utility by delivering relevant information directly through the interface. For future development, expanding the dataset by incorporating drug images with greater variety in perspectives, lighting conditions, and backgrounds. Integrating additional augmentation techniques to improve the model's resilience in suboptimal conditions. Developing a more intuitive and user-friendly graphical user interface (GUI) to ensure accessibility for diverse users. These development steps aim to make the drug detection system more adaptable and user-friendly, thereby broadening its benefits for a wide range of users.

REFERENCES

[1] Y. Xu, Y. Liu, H. Li, L. Wang, and J. Ai, "A Deep Learning Approach of Intrusion Detection and Tracking with UAV-Based 360° Camera and 3-Axis Gimbal," *Drones*, vol. 8, no. 2, Feb. 2024, doi: 10.3390/drones8020068.

[2] A. T. Bruni, C. H. P. Rodrigues, C. Dos Santos, J. S. de Castro, L. S. Mariotto, and L. F. C. Sinhorini, "Analytical Challenges for

Identification of New Psychoactive Substances A Literature-Based Study for Seized Drugs," 2022, *Visao Fokka Communication Agency*. doi: 10.30744/brjac.2179-3425.RV-41-2021.

[3] P. Breen and Smith Rebecca, "Deaths related to drug poisoning in England and Wales 202 registrations 1," Dec. 2023.

[4] E. M. Olson *et al.*, "Physician Identification Badges: A Multispecialty Quality Improvement Study to Address Professional Misidentification and Bias," *Mayo Clin Proc*, vol. 97, no. 4, pp. 658–667, Apr. 2022, doi: 10.1016/J.MAYOCP.2022.01.007.

[5] A. Hafezi, A. Babaii, B. Aghaie, and M. Abbasinia, "The relationship between patient safety culture and patient safety competency with adverse events: a multicenter cross-sectional study," *BMC Nurs*, vol. 21, no. 1, Dec. 2022, doi: 10.1186/s12912-022-01076-w.

[6] E. M. Tichy *et al.*, "National trends in prescription drug expenditures and projections for 2022," *American Journal of Health-System Pharmacy*, vol. 79, no. 14, pp. 1158–1172, Jul. 2022, doi: 10.1093/ajhp/zxac102.

[7] B. G. de la Torre and F. Albericio, "The Pharmaceutical Industry in 2021. An Analysis of FDA Drug Approvals from the Perspective of Molecules," Feb. 01, 2022, *MDPI*. doi: 10.3390/molecules27031075.

[8] E. Chai, L. Ta, Z. Ma, and M. Zhi, "ERF-YOLO: A YOLO algorithm compatible with fewer parameters and higher accuracy," *Image Vis Comput*, vol. 116, p. 104317, Dec. 2021, doi: 10.1016/J.IMAVIS.2021.104317.

[9] X. Yang *et al.*, "OTTM: an automated classification tool for translational drug discovery from omics data," *Brief Bioinform*, vol. 24, no. 5, Sep. 2023, doi: 10.1093/bib/bbad301.

[10] L. Tan, T. Huangfu, L. Wu, and W. Chen, "Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification," *BMC Med Inform Decis Mak*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12911-021-01691-8.

[11] J. Heo, Y. Kang, S. K. Lee, D. H. Jeong, and K. M. Kim, "An Accurate Deep Learning–Based System for Automatic Pill Identification: Model Development and Validation," *J Med Internet Res*, vol. 25, 2023, doi: 10.2196/41043.

[12] A. J. Tittle-Ind Multi Res, A. Stefi, P. M. Varsha, S. Sumana, and V. Vandana Simha, "Indiana Journal of Multidisciplinary Research Using Machine Learning Models to Detect and Identify Pills," 2024, doi: 10.5281/zenodo.12674988.

[13] H. W. Ting, S. L. Chung, C. F. Chen, H. Y. Chiu, and Y. W. Hsieh, "A drug identification model developed using deep learning technologies: Experience of a medical center in Taiwan," *BMC Health Serv Res*, vol. 20, no. 1, Apr. 2020, doi: 10.1186/s12913-020-05166-w.

[14] M. G. Ragab *et al.*, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.

[15] O. E. Olorunshola, M. E. Irhebhude, and A. E. Evwiekpaefe, "A Comparative Study of YOLOv5 and YOLOv7 Object Detection Algorithms 1*," 2023.

[16] T. Ahmad *et al.*, "Object Detection through Modified YOLO Neural Network," *Sci Program*, vol. 2020, pp. 1–10, 2020, doi: 10.1155/2020/8403262.

[17] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/J.GLTP.2022.04.020.

[18] G. Iglesias, E. Talavera, Á. González-Prieto, A. Mozo, and S. Gómez-Canaval, "Data Augmentation techniques in time series domain: a survey and taxonomy," May 01, 2023, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-023-08459-3.

[19] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J Imaging*, vol. 9, no. 2, Feb. 2023, doi: 10.3390/jimaging9020046.

[20] P. S. Chary, "Real Time Object Detection Using YOLOv4," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 12, pp. 1375–1379, Dec. 2023, doi: 10.22214/ijraset.2023.57602.

[21] R. N. Ariwa, C. Markus, N. G. Teneke, S. Adamu, and K. G. Fumlack, "Plant Disease Detection Using Yolo Machine Learning Approach," *British Journal of Computer, Networking and Information Technology*, vol. 7, no. 2, pp. 115–129, Jul. 2024, doi: 10.52589/bjcnit-ejwgfw6d.

[22] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, "YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions," Mar. 2025.

[23] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.10934