IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

# MEDIINSIGIHT FOR TABLET RECOGNITION WITH INTEGRATED OBJECT DETECTION AND OPTICAL CHARACTER RECOGNITION

**Ms. P. Yasaswani[1], Mr. K. Abhishek[1], Mr. K. Kamal [1], Ms. M. Srilakshmi [1],**
[1]B. Tech Final Year Student Lendi Institute of Engineering and Technology(A),
Jonnada, Vizianagaram, Andhra Pradesh, India

***Abstract:*** This project presents a novel approach to tablet identification, leveraging the YOLO (You Only Look Once) algorithm for half tablet sheets and Optical Character Recognition (OCR) for full tablet sheets. The system dynamically adapts to user-input formats, offering a versatile solution for both half and full tablet sheet scenarios. For half tablet sheets, the YOLO algorithm is employed for object detection, ensuring precise identification and localization of tablet images. On the other hand, full tablet sheets are scanned using OCR to retrieve the tablet names. The integration of these technologies enables accurate and efficient retrieval of tablet-related information, including specifications. The system's adaptability to different input formats enhances accessibility, providing users with a seamless experience in obtaining comprehensive details about tablets. This innovative approach not only streamlines the identification process but also caters to diverse user preferences, contributing to improved usability and effectiveness in accessing tablet information.

**Keywords –** ***Yolo, Optical Character Recoginition, Tablets, Identification.(key words)***

## I. INTRODUCTION

The ubiquity of tablet devices in contemporary society has led to an escalating demand for efficient and user-friendly tools to identify and retrieve specifications for these devices. As tablets come in various shapes and sizes, often presented in half or full tablet sheet formats, the need for a robust computer vision-based solution becomes apparent. In response to this demand, we propose a novel application that harnesses the power of advanced technologies, specifically the YOLO (You Only Look Once) algorithm and Optical Character Recognition (OCR), to seamlessly tackle the challenges associated with tablet identification and specification retrieval. Tablets serve as indispensable tools in various domains, from personal use to professional applications. However, the diversity in tablet dimensions and the sheer volume of available models necessitate automated solutions for accurate identification and efficient access to specifications. Traditional methods often fall short in handling the complexity of tablet sheet images, prompting the integration of state-of-the-art computer vision techniques into our proposed application.The YOLO algorithm, known for its real-time object detection capabilities, offers a robust solution for precisely localizing tablet images within input sheets. This not only enhances the accuracy of tablet identification but also accommodates variations in tablet sizes and orientations. Complementing this, the utilization of OCR facilitates the extraction of textual information, ensuring that crucial specifications are retrieved and presented to users in a comprehensible manner. In this context, our research aims to bridge the gap between the ever-expanding realm of tablets and the need for accessible, automated tools to discern and understand their specifications. By amalgamating YOLO and OCR technologies, our application seeks to provide users with a streamlined and efficient means of tablet identification and specification retrieval, contributing to a more intuitive and user-friendly experience in the digital landscape.

## II. LITERATURE SURVEY

Through an extensive review of relevant literature, this survey aims to provide a thorough understanding of the methodologies, techniques, and recent advancements in the field. By systematically examining existing approaches, the survey intends to shed light on the challenges inherent in tablet sheet identification, critically evaluate the effectiveness of different methods, and pinpoint areas for further improvement. In conducting this literature survey, we endeavor to contribute to the advancement of tablet sheet identification by synthesizing existing knowledge, identifying gaps in the research landscape, and proposing avenues for future investigation. By elucidating the complexities and nuances of tablet sheet identification, we aim to facilitate the development of more efficient, accurate, and reliable solutions. Ultimately, our efforts are geared towards enhancing document management and accessibility in the digital age, thereby enabling individuals and organizations to harness the full potential of tablet sheets for various purposes, including academic research, business analytics, and data management. By doing so, we seek to foster a more efficient, accurate, and reliable framework for working with tablet sheets, thereby catalyzing advancements in various fields and unlocking new possibilities for innovation and productivity.

| Authors | Title of the Project | Methodology | Inference | Limitations |
|---|---|---|---|---|
| Hsien-Wei Ting et al. | Drug Identification model developed using deep learning | You Only look Once (YOLO) | The project demonstrates a convolution-based deep learning network's ability to address 'look-alike' errors in drug identification, achieving better accuracy. It offers practical applications for pharmacists to prevent medication errors with similar-looking blister packages. The findings can also serve as core software for robotic prescription filling, enhancing efficiency and safety. | The model's performance could be sensitive to variations in image quality and packaging, impacting its generalization. Additionally, the deployment of the deep learning network in dynamic, real-time pharmacy environments may face challenges such as unexpected variations in drug presentations |
| Chengji Liu et al. | Object Detection Based on YOLO Network | You Only look Once (YOLO) | the model was trained using degraded images which resulted in improved average precision. It was proved that the average precision for degraded images was better in general degenerative model compared to the standard mode. | The model may lack generalization to standard high-quality images, potentially limiting its applicability in broader contexts. The reliance on a specific degraded image dataset may hinder performance in diverse real-world conditions. |
| Hyuk-Ju Kwon et al. | Pill Detection Model for Medicine Inspection Based on Deep Learning | Mask R-CNN is an extended model of Faster R- CNN, which is an instance segmentation algorithm that simultaneously predicts the bounding box that informs the existing object location and the mask of the object area. | The algorithm addresses challenges posed by the exponential increase in pill combinations when dealing with multiple pill types in an image. Despite constraints in imaging and dataset size, the algorithm demonstrates superior performance compared to existing methods. | the algorithm's scalability to handle an exponentially increasing number of pill combinations might be a concern, especially if confronted with novel or uncommon combinations not well-represented in the training data. |
| Isha Patel et al | Optical Character Recognition using Deep learning | Convolutional Neural Networks (CNNs) process input data through convolutional layers, where filters | It specifically addressing problems like line segments, half characters, same characters, different thicknesses, and discontinuation or distortion of characters. It highlights | the need for extensive, representative datasets for effective deep learning, potential challenges in adapting to diverse |

| | | | | |
|---|---|---|---|---|
| | | extract features. Pooling layers reduce dimensionality, and fully connected layers classify extracted features, enabling CNNs' efficacy in image recognition and pattern detection | the necessity of large datasets for applying deep learning techniques to character recognition | regional language characteristics, and uncertainties in real-world applications due to varying script complexities. |

## III. DESIGN AND IMPLEMENTATION

The implementation process for tablet image identification and specification retrieval is a comprehensive endeavor that involves various stages, each crucial for the successful operation of the system. The process begins with user interaction, where users provide tablet images for analysis. These images serve as the input data for the system, initiating the data processing pipeline. Upon receiving the image, the system undergoes several preprocessing steps to ensure the data's quality and consistency. Initially, the system standardizes image resolutions, formats, and lighting conditions to ensure uniformity across the dataset. This standardization is essential for accurate analysis and comparison of tablet images. Additionally, the system employs image segmentation techniques to isolate tablet images within the input, effectively separating them from the background or surrounding elements. This step prepares the data for further analysis by focusing specifically on the tablets within the image.

Following image preprocessing, the system proceeds to annotate the extracted tablet images with bounding boxes, indicating their precise locations within the input sheet. This annotation process is crucial for supervised learning, providing the algorithm with labeled data to learn from. By annotating the images, the system enables the algorithm to accurately detect and localize tablets within various contexts, laying the foundation for subsequent training steps.

With the dataset prepared and annotated, the system trains the YOLO (You Only Look Once) algorithm for tablet detection. During training, the model learns to recognize and localize tablets within images, adapting to factors such as different tablet dimensions, orientations, and backgrounds. Fine-tuning techniques are employed to optimize the algorithm's performance specifically for tablet detection, ensuring high accuracy and robustness.

Validation checks are conducted regularly during training using a portion of the dataset set aside for validation purposes. These checks assess the model's accuracy and generalization on unseen data, identifying and addressing any overfitting or underfitting issues that may arise. Once the model is validated, the system integrates Optical Character Recognition (OCR) with the trained YOLO model. OCR is utilized to extract textual information, such as serial numbers or product codes, from the identified tablets, enhancing the depth of information retrieval.

After extracting textual information, the system processes the text to identify relevant keywords or specifications using Natural Language Processing (NLP) techniques. These techniques analyze the text and extract structured data, such as tablet brand, model, and specifications. The extracted specifications are then used to query a database containing information about various tablet models, retrieving relevant details such as technical specifications, usage instructions, and compatibility details.

Finally, the system presents the retrieved tablet specifications to the user in a user-friendly format, which may include displaying the specifications on a graphical user interface or providing them in a structured document format. Throughout the entire process, ethical considerations are paramount. The system ensures the ethical use of user-provided data, addresses potential biases and privacy concerns, and adheres to ethical guidelines and regulations governing data handling and AI usage.

In summary, the implementation process for tablet image identification and specification retrieval involves a series of interconnected steps, ranging from data preprocessing and algorithm training to information extraction and presentation. Each step plays a crucial role in ensuring the accuracy, reliability, and ethical use of the system, ultimately providing users with valuable insights into tablet specifications in a user-friendly manner..

**Table 1.** Training and testing rules of the deep learning network

| | |
|---|---|
| Size of input image<br>224 × 224 pixels | Adjusted to |
| Network built-in data augmentation function | disabled |
| Pre-trained model | no |
| Batch size | 16 |
| Highest no of training Epochs (168,800 iterations) | 80 Epochs |
| Training weight file storage timing | 1 Epoch (1688 iterations) |

## IV. RESULTS

The study involves training the input images using the YOLO (You Only Look Once) algorithm, a state-of-the-art object detection algorithm, to accurately identify tablet images within half or full tablet sheet formats. Through a process of training, where the algorithm learns from annotated images, it becomes proficient in recognizing tablets of varying dimensions. During training, the YOLO algorithm is exposed to a large dataset of annotated tablet images, where each tablet is labeled with its corresponding class (tablet name) and bounding box coordinates. Through iterative optimization processes, the algorithm learns to detect tablets within images and predict their respective classes with high accuracy.

Once trained, the system is capable of efficiently identifying tablets in new, unseen images. This involves passing the test images through the trained YOLO model, which analyzes the images and predicts the tablet's class (name) along with its location within the image. Additionally, the system utilizes Optical Character Recognition (OCR) to extract textual information from the identified tablets. This OCR component further enhances the system's capabilities by extracting crucial specifications from the tablet images, such as brand, model, dosage, and other relevant details.By integrating YOLO for precise tablet detection and OCR for textual information extraction, the system provides a comprehensive solution for automating tablet identification and retrieving detailed specifications. This approach enhances user experiences by offering quick and accurate access to pertinent tablet details in our digital and fast-paced world.


fig1. Ofloxacin Tablets IP Labelling


fig2.Telmikind Tablets Labelling


fig3.Pentab Tablets Labelling

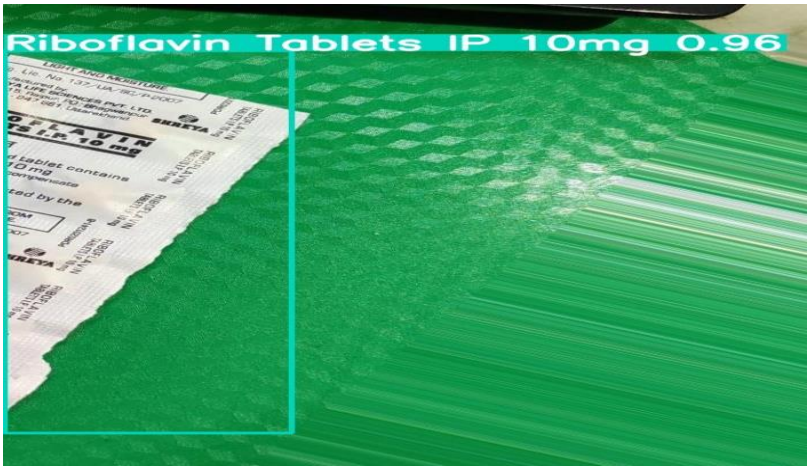


fig4.Riboflavin Tablets Labelling



## V. RESULT ANALYSIS

The integration of advanced technologies such as the YOLO (You Only Look Once) model and Optical Character Recognition (OCR) has revolutionized the process of tablet identification and specification retrieval. Once trained, the system harnesses the power of the YOLO model to efficiently detect tablets in new, unseen images. YOLO operates by analyzing the entirety of an image in a single pass, making it highly efficient for real-time applications. Through this process, the model not only identifies the presence of tablets but also predicts their class or name along with their precise location within the image.

However, the identification of tablets alone may not suffice for comprehensive analysis or user utility. Therefore, the system employs OCR technology to extract textual information from the identified tablets. OCR enables the extraction

of crucial specifications such as the brand, model, dosage, and other relevant details from the images. This extraction process significantly enhances the system's capabilities, providing users with detailed and pertinent information about the tablets they are interested in.By integrating YOLO for precise tablet detection and OCR for textual information extraction, the system offers a holistic solution for automating tablet identification and specification retrieval. This approach not only streamlines the process but also ensures accuracy and efficiency, thereby enhancing user experiences. In today's digital and fast-paced world, where quick access to information is paramount, such a system proves invaluable.Users benefit from the system's ability to quickly and accurately identify tablets and provide comprehensive details without the need for manual input or extensive research. Whether it's for medical professionals seeking to verify medications or individuals looking to understand the specifics of their prescriptions, the system offers a reliable and user-friendly solution.

Moreover, the integration of these technologies opens up possibilities for further enhancements and applications. For instance, the system could be expanded to include additional features such as medication reminders, dosage tracking, or drug interaction warnings, thereby promoting better health outcomes and medication management. Overall, the combination of YOLO and OCR technologies represents a significant advancement in the field of automated tablet identification and specification retrieval, catering to the needs of users in today's digital age.

## VI. CONCLUSION

In conclusion, this project represents a significant advancement in computer vision applications, particularly in the realm of tablet identification and specification retrieval. By harnessing the power of the YOLO algorithm for precise detection of tablet images and integrating Optical Character Recognition (OCR) for extracting textual information, the system offers an innovative solution that streamlines the process of identifying and accessing tablet details.The system's ability to accommodate both half and full tablet sheet formats, along with its robust performance in handling varying tablet dimensions, underscores its versatility and practicality in real-world scenarios. Moreover, the seamless integration of YOLO and OCR technologies ensures accurate and efficient processing, enabling users to obtain crucial tablet specifications swiftly and with ease. This research not only contributes to the advancement of computer vision applications but also significantly enhances user experiences by providing a user-friendly and automated solution for tablet identification. In our digital and fast-paced world, where access to detailed information is paramount, the system's capabilities hold immense value in facilitating informed decision-making and improving productivity.

Moving forward, further research and development in this area could explore additional enhancements to the system, such as incorporating advanced natural language processing techniques for more comprehensive text analysis or integrating additional data sources to enrich the information retrieved. Additionally, efforts to address ethical considerations, such as ensuring user privacy and data security, should remain a priority in the continued development and deployment of such technologies.Overall, this project lays a solid foundation for future innovation in computer vision applications and underscores the potential of technology to enhance user experiences and streamline processes in various domains.

## VII. REFERENCES

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, “You Only Look Once: Unified, Real-Time Object Detection”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788

[2] Ren, S., He, K., Girshick, R., and Sun, J. Faster r-CNN: Towards real time object detection with region proposal networks. In
Advances in neural information processing systems (2015), pp. 91–99

[3] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, “Selective search for object recognition,” International Journal of Computer Vision (IJCV), 2013.

[4] Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E (2018) Deep learning for computer vision: a brief review. Comput
Intell Neurosci 2018:1–13

[5] Yuming He, Research on Text Detection and Recognition Based on OCR Recognition Technology,2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), 27-29 September 2020

[6] Rey J (2017) Object detection with deep learning: the definitive guide

[7] Liu L, Ouyang W, Wang X, Fieguth P, Chen J, Liu X, Pietikäinen (2020) M. Deep learning for generic object detection: a survey. Int J Comput Vis 28(2):261–318

[8] Jiao L, Zhang F, Liu F, Yang S, Li L, Feng Z, Qu R (2019) A survey of deep learning-based object detection. IEEE Access
7:128837–128868.

[9] Ridvy Avyodri, Samuel Lukas, Hendra Tjahyadi, Optical Character Recognition (OCR) for Text Recognition and its Post-Processing Method, 2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA)

[10] Ch. Rupa, Gautam Srivastava, Bharath Ganji, Sai Praveen Tatiparthi, Karthik Maddala, Srinivas Koppu and Jerry Chun-Wei
Lin , Medicine Drug Name Detection Based Object Recognition Using Augmented Reality

[11]Healthline Medication Safety: Pill Identification, Storage, More.2021. Available online:
https://www.healthline.com/health/pill-identification (accessed on 31 July 2022).

[12] Z. Yaniv et al., "The national library of medicine pill image recognition challenge: An initial report", Proc. IEEE Appl. Imag. Pattern Recognit. Workshop (AIPR), pp. 1-9, Oct. 2016.

[13] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, Published in Advances in Neural Information Processing Systems 28 (NIPS 2015).

[14] Information Processing Systems 28 (NIPS 2015). 8. Joseph Redmon, Ali Farhadi, “YOLO9000: Better, Faster, Stronger”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263-7271.

[15] YOLO Juan Du1,” Understanding of Object Detection Based on CNN Family”, New Research, and Development Centre of Hisense, Qingdao 266071, China.