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

A Medical Image Assistant Featuring Text-to-Image Retrieval, Deep Learning Annotation, and Description System

Nikhitha R Karanth
Dept. of Computer Science and Engineering
PESITM, Shivamogga

Noor Ayesha Naaz
Dept. of Computer Science and Engineering
PESITM, Shivamogga

Radhika B Kulkarni
Dept. of Computer Science and Engineering
PESITM, Shivamogga

Ramya P
Dept. of Computer Science and Engineering
PESITM, Shivamogga

Shivanand C Maradi
Dept. of Computer Science and Engineering
PESITM, Shivamogga

Abstract- Contemporary medical education has tremendous potential to enhance learning experiences by integrating machine learning technologies with data strategies. This paper presents a new framework designed for educational settings, using machine learning methods to analyze text-based, visual, and language-based medical education in schools to provide detailed information through image generation through medical concepts, observations, and contextual understanding. It enhances the learning environment.

Keywords: Machine Learning, Medical Education, Textual Analysis, Image Generation.

I INTRODUCTION

In today's rapidly evolving educational environment, the integration of technology has become an important factor in enhancing the learning experience. Nowhere is this more evident than in medicine, where the complexity and breadth of issues present significant challenges for teachers and students. With population growth and rapid urbanization, the demand for health professionals increases, emphasizing the growing importance of a strong medical curriculum but traditional methods of medical education is often struggling to match students' diverse needs and learning styles. Moreover, the dynamic nature of medical knowledge requires adaptive and innovative approaches to education. In this context, machine learning (ML) has emerged as a revolutionary force, offering unparalleled opportunities to transform medical education.

This project is the application of ML for the development of a sophisticated teaching tool specifically designed for medical education in schools. The overall goal is to simplify the curriculum to enhance teachers and students, to facilitate a deeper understanding of medical concepts, and to develop competent health professionals.

Additionally, the project seeks to bridge the gap between textual and visual learning styles. Visual aids play an important role in reinforcing conceptual understanding, especially in complex areas such as medicine. By integrating imaging techniques, including YOLO V8M recognition, the system will provide visual images that match text descriptions, enabling the clarity and retention of medical knowledge to the sky.

Furthermore, the project endeavors to bridge the gap between textual and visual learning modalities. Visual aids play a crucial role in reinforcing conceptual understanding, particularly in fields as intricate as medicine. Through the integration of image processing techniques, including YOLO V8M for object detection, the system will provide visual representations corresponding to textual descriptions, thereby enhancing the clarity and retention of medical knowledge.

Specifically, this project represents a concerted effort to use ML technology to improve medical education in schools. Developing a comprehensive and scalable educational tool aims to empower faculty and students to navigate the complexities of medical science with confidence and skill, ultimately training the next generation of health professionals.

II LITERATURE SURVEY

The YOLO (You Look Once) system has been a cornerstone in real-time object recognition systems, and its latest iteration, YOLO V8M, marks a significant leap Paper “A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond ” provides a complete analysis of the evolution of the YOLO framework, where its Innovations and contributions to the current state of affairs are discussed in detail This review is essential to understanding the capabilities of the framework in applications ranging from robotics to video monitoring.

The MedYOLO system represents a breakthrough in medical imaging. It is a 3-D object recognition system that uses the YOLO family of single-view methods, analogous to medical imaging applications. Demonstrating its high performance in medium- and large-scale systems, the effectiveness of the system has been validated on a variety of data2. This highlights the potential of YOLO models in clinical research and treatment planning.

Furthermore, the integration of natural language processing (NLP) into medical applications has opened up new possibilities for enhancing health outcomes. The article “Medical Text and Image Processing: Applications, Issues, and Challenges” explores the critical role of text and image analysis in health care, addressing the challenges and opportunities presented by these technologies.

Cognitive Deep Learning Method for Medical Image Processing, as described in “Cognitive Deep Learning Method for Medical Image Processing”, uses a hybrid model that combines U-Net algorithms with such advanced visualization techniques of human-like cognitive computing systems Medical revolution is being transformed into image analysis by transformation and rationalization.

Collectively, these data highlight rapid advances in object recognition, medical imaging, and NLP. The integration of these technologies is set to transform a wide range of applications, delivering more accurate, efficient, and reliable solutions with the potential to significantly impact patient care and medical research.

III PROPOSED SYSTEM

Our proposed system bridges the gap between textual descriptions and visual information, particularly in the realm of medical images, by harnessing the power of Deep Learning and AI libraries. It is specifically tailored to help school children grasp medical image concepts more easily, fostering a better understanding of the human body and medical science. These young users have two primary ways to engage with our system. First, they can input a word related to an organ, such as "spleen." Second, they can directly upload a picture to the system.

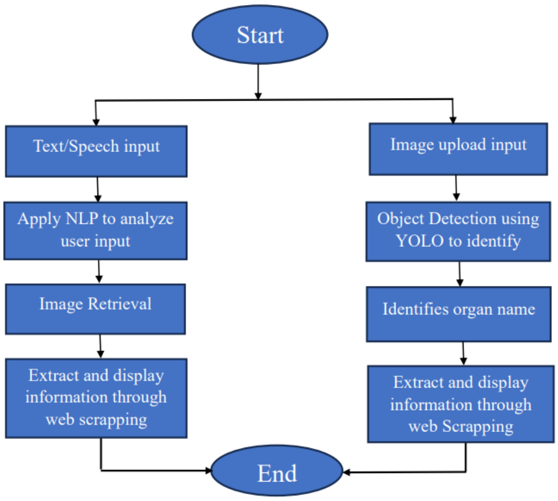

Fig 1. Architecture of proposed system

When a user inputs a word, our system utilizes Natural Language Processing (NLP) to grasp their intended meaning. This process is akin to teaching the computer to comprehend human language. Once the user's intent is understood, we employ an AI Library called Simple Image Download to retrieve relevant pictures matching their input. Subsequently, system scour the internet that is web scrapping to gather information about the organ depicted in the picture, including its function within the body and its location.

In the case of a user uploading a picture, Deep Learning algorithm, YOLOv8, analyzes the image to identify its contents. For instance, if a school child uploads a picture of a spleen, YOLOv8 will recognize it as such. We then proceed to web scrapping to provide the user with more detailed information about the spleen, enriching their understanding.

Through the combined use of NLP, YOLOv8, and internet searching, our system simplifies the process of understanding medical images for users, whether they are school children typing in words or uploading pictures

IV METHODOLOGY

YOLOv8, a powerful Deep Learning algorithm, excels in medical image detection due to its speed and accuracy. Trained on vast datasets of labeled medical images, YOLOv8 can swiftly identify and potentially localize objects within new images. To train YOLOv8, several organ images are collected and annotated using a tool called Labelling. This involves drawing bounding boxes around each organ to facilitate the model's learning process. The annotated images are then divided into training and validation sets, and the YOLO configuration was adjusted to accommodate the various organ classes within the dataset.

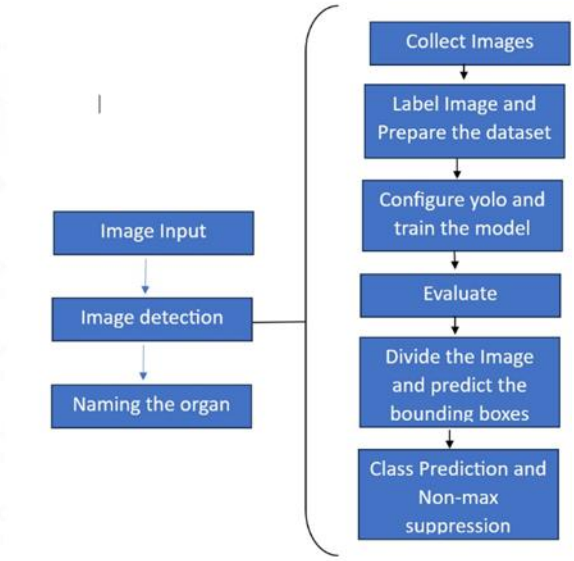

Fig 2. Methodology

Subsequently, the model is trained using the training dataset, allowing it to learn the distinctive features of each organ and how to accurately delineate bounding boxes around them. Testing is conducted using the validation set to measure the model's performance. During inference, the trained YOLOv8 model is presented with a new image, which it divides into a grid. Each grid cell predicts bounding boxes and assigns a confidence score to these predictions. For each bounding box, the model predicts class probabilities, indicating the likelihood of the detected object belonging to a particular class. This information is then utilized to display the name of the detected object, facilitating the interpretation of medical images with enhanced accuracy and efficiency. By leveraging technology to bridge the gap between textual descriptions and visual information, system empower school children to explore the fascinating world of medical imaging and anatomy, ultimately promoting lifelong learning and curiosity in the field of healthcare.

V EXPECTED RESULT

Our system empowers school children to delve into the fascinating world of human anatomy through a user-friendly system. Users can interact with our system through two primary modes: inputting text for search and image retrieval or directly uploading images.

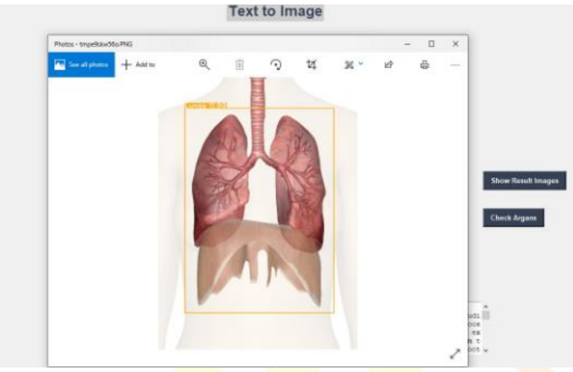

Fig 3. Organ Detection

The core technology behind this exploration is Deep Learning. YOLOv8 acts like a highly trained expert, swiftly and accurately identifying organs within medical images. This approach may enhance the accuracy and reliability. This involves drawing bounding boxes around each image with annotations.

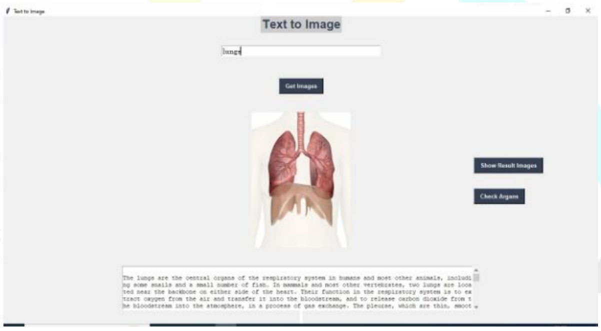

Fig 4. Retrieval of Image and Description based on the user input

After detection of the image, web scraping is used to gather detailed descriptions, providing school children with a comprehensive understanding of the identified organ's function and location within the body. This combination of NLP, YOLOv8's object detection and Web scraping for descriptions empowers young learners to explore medical images in an interactive and informative way.

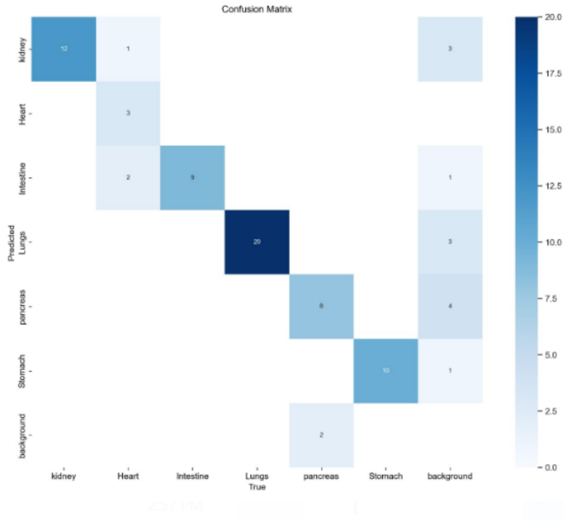

Fig 5. Confusion Matrix

In conclusion, our project endeavors to revolutionize medical education in schools through the integration of machine learning technologies, including natural language processing and image recognition. By developing a versatile system capable of analyzing medical-related text, images, and speech inputs, we aim to address the challenges faced by educators and students in comprehending complex medical concepts. Through intuitive interfaces and accessibility features such as speech recognition, our system seeks to empower both educators and students, fostering inclusive and personalized learning experiences. With a focus on simplicity, effectiveness, and inclusivity, we envision our project as a catalyst for democratizing access to medical education and nurturing the next generation of healthcare professionals, while also highlighting the potential for further advancements in machine learning to reshape the landscape of medical education globally.

REFERENCES

[1] Viola P and Jones M, "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, 2001, pp. 511-518.

[2] Jawaid Nasreen., Warsi Arif. and Asad Ali Shaikh, Yahya Muhammad (2019), "Object Detection and Narrator for Visually Impaired People", IEEE International Conference on Engineering Technologies and Applied Science.

[3] Ortiz J. A. R., Beltran B. A. G. Dan Lopez L. G., 2015, Clinical Decision Support Systems: A survey of NLP-based Approaches from Unstructured Data, 26th International Workshop on Databased ang Expert Systems Applications, pp. 163-167.

[4] W. Lan, J. Dang, Y. Wang, and S. Wang, "Pedestrian detection based on yolo network model," in 2018 IEEE international conference on mechatronics and automation (ICMA), pp. 1547–1551, IEEE, 2018.

[5] W.-Y. Hsu and W.-Y. Lin, "Adaptive fusion of multi-scale yolo for pedestrian detection," IEEE Access, vol. 9, pp. 110063–110073, 2021.

[6] A. Benjumea, I. Teeti, F. Cuzzolin, and A. Bradley, "Yolo-z: Improving small object detection in yolov5 for autonomous vehicles," arXiv preprint arXiv:2112.11798, 2021.