INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 4, April 2026 | ISSN: 2456-4184 | IJNRD.ORG

# *Integrated brain tumor segmentation via yolov11 with an interactive patient supportive chatbot*

**Ms.J.Lethisia Nithiya,M.E,**
**Assistant professor,**
**Computer Science and Engineering,**
**Bharath Institute of Science and technology (BIST),**
**173, Agaram road, Selaiyur, Tambaram,**
**Chennai-600073,Tamil Nadu,India.**

**Vimalkannan B,**
**Computer Science and Engineering,**
**Bharath Institute of Science and technology (BIST),**
**173, Agaram road, Selaiyur, Tambaram,**
**Chennai-600073,Tamil Nadu, India.**

**Vigneshwaran J,**
**Computer Science and Engineering,**
**Bharath Institute of Science and technology (BIST),**
**173, Agaram road, Selaiyur, Tambaram,**
**Chennai-600073,Tamil Nadu, India.**

**L Vijayvarman,**
**Computer Science and Engineering,**
**Bharath Institute of Science and technology (BIST),**
**173, Agaram road, Selaiyur, Tambaram,**
**Chennai-600073,Tamil Nadu, India.**

**D Vijayaraghavan,**
**Computer Science and Engineering,**
**Bharath Institute of Science and technology (BIST),**
**173, Agaram road, Selaiyur, Tambaram,**
**Chennai-600073,Tamil Nadu, India.**

***Abstract:*** **In Brain tumor treatment process, segmentation plays a major vital role in treating tumor and observing it through MRI scan images of brain. This research exhibits the automated method of segmentation with deep learning approaches. This research contains preprocessing and segmentation strategies that enhance the boundaries of tumor. It deals with the challenges of tumor size variability and irregular shape which requires a enhancement in accuracy and efficiency. Essential idea of research is to study yolov8 model's inefficiency in brain tumor segmentation which is claimed as prominent architecture in existing system. Additionally in this research, we implement a supportive chatbot feature for brain tumor patients to educate and guide them which is never done for brain tumor patients in existing researches. The chatbot system convey the patients about their treatment processes and symptoms in case of unable to contact doctor and doubtfulness in their mind, so they can know about their conditions and can make decision to visit hospital immediately. This improves patient trustability on treatment processes. This research utilizes yolov11 architecture which contains several**

**segmentation techniques to enhance accuracy, efficiency and provide detailed insights on tumor location and position.**

***Keywords: Brain tumor segmentation, deep learning, MRI, yolov11, Chatbot, Automated.***

## I. INTRODUCTION

Basically tumors are one of the most vulnerable diseases in the world, especially brain tumors are threatening to life. Brain tumor treatment processes were vital in saving patients from the threat, tumor segmentation is important for planning the treatment. In the field of neurology, automated frameworks is very required to handle the tasks where accurate detection and segmentation is essential for treatment planning. Existing tools use a outdated technique yolov8 which does not contain accuracy enhancing frameworks, this research uses yolov11 for improving efficiency.

Apart from existing simple yolov11 architecture, we additionally implemented a supportive chatbot system to the research which no other researches upfronted to implement it. MRI (magnetic resonance imaging) is most common and efficient imaging method to segment the tumors, so we too used mri images.

The research setup is to train the yolov11 model with mri images and then to provide segmentation results. They also determine the type of tumors based on the location (meningioma, glioma, pituitary). The research integrated with supportive chatbot framework implemented by llama and groq api with principles of natural language processing. Patient who diagnosed with brain tumor lacks information on their treatment procedure and staying in collapsed or doubtful mindset, our research discovered and cared about that fact, we overcome with an automated chatbot which conveys them the information and knowledge of their symptoms and treatment procedure. That kind of information lacking often leads to anxiety development, stress and fear that would affect their mentality, hope and decision making abilities which affect patient trust in healthcare. This research Ultimately, improve patient outcomes by enabling them to make informed decisions about their care and seek medical attention when necessary.

*A. Project objective*

1. To develop an automated, real time, and highly accurate brain tumor segment classification system using the advanced YOLOv11 deep learning architecture.
2. To educate patients by explaining their symptoms and tumor types in an accessible manner to enhance patient engagement.
3. To enhance accuracy and reliability in identifying and delineating brain tumor regions from MRI images and provide a chatbot service to patients to support their mental health , helps them to be aware of their own health.
4. To overcome challenges of tumor size variability and complexity in imaging.
5. To provide precise tumor boundaries, size, and location for improved diagnosis and treatment planning.
6. To support clinicians with a fast, auto, and efficient decision-support tool for better segmentation and patient care.

*B. Project motivation*

1. To Enhancing patient engagement by Empowering patients to take education in their care by providing them with proper information and support.
2. To Reduce anxiety and stress of patients by providing them with trustable and accessible safe information about their symptoms and treatment processes.
3. To ultimately improve patient outcomes by enabling them to make informed decisions about their health and get medical attention when necessary.
4. Manual tumor segmentation from MRI scans is time-consuming, labor-intensive, and prone to human error, this research helps in eliminating these.
5. Advanced model YOLOv11 enables real time, automated, and highly accurate tumor segmentation.
6. Providing doctors with a fast and reliable decision-support system can improve treatment planning and ultimately save lives.
7. The challenge is to design a model that can not only detect the presence of a tumor but also segment it precisely, providing clear and accurate tumor boundaries.
8. The proposed yolov11 model seeks to address these challenges by improving segmentation accuracy and providing detailed insights into tumor location.
9. The task is to develop a deep learning based system that can accurately segment brain tumors in MRI and integrate an chatbot feature with natural language processing properties.

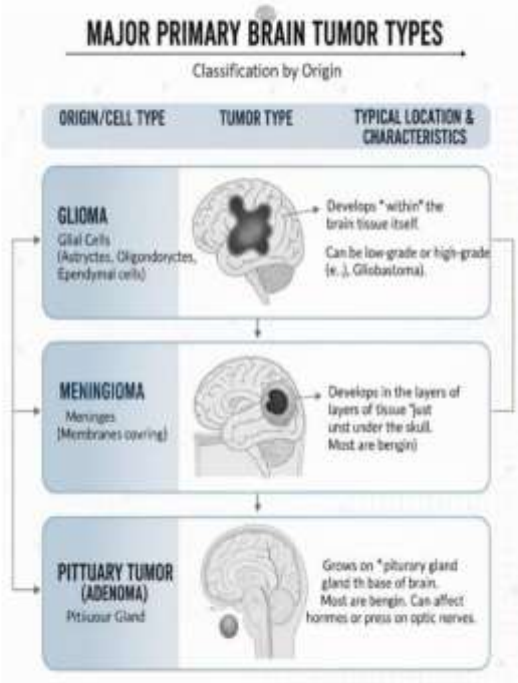


*Fig 1.1 Major brain tumor types*

## II. RELATED RESEARCH

Essential idea of research is to study yolov8 model's inefficiency in brain tumor segmentation which is claimed as prominent architecture in existing system. Additionally in this research, we implement a supportive chatbot feature for brain tumor patients to educate and guide them which is never done for brain tumor patients in existing researches. In the field of neurology, automated frameworks is very required to handle the tasks where accurate detection and segmentation is essential for treatment planning. Existing tools use a outdated technique yolov8 which does not contain accuracy enhancing frameworks, this research uses yolov11 for improving efficiency. Apart from existing simple yolov11 architecture, we additionally implemented a supportive chatbot system to the research which no other researches upfronted to implement it. Patient who diagnosed with brain tumor lacks information on their treatment procedure and staying in collapsed or doubtful mindset, our research discovered and cared about that fact, we overcome with an automated chatbot which conveys them the information and knowledge of their symptoms and treatment procedure. That kind of information lacking often leads to anxiety development, stress and fear that would affect their mentality, hope and decision making abilities which affect patient trust in healthcare.

YOLOv5 is based on PyTorch framework native, which employed CSPDarknet53 as a backbone and a PANet neck to combine features of various scales. It did not present a brand new detection theory, but its feature of Auto Anchor was crucial; it trained ideal anchor box sizes on a given dataset (e.g. MRIs) automatically. In the case of medical imaging, Yolov5 led in popularity due to its easy deployment and flexibility, but could not handle pixel-level detail covering the smallest of tumors. YOLOv7 emphasized a lot on trainable bag-of-freebies and architecture efficiency. The greatest innovation it made was its E-ELAN (Extended Efficient Layer Aggregation Network) that enabled the model to be learnt better through regulating the shortest and longest gradient. This assisted in converting more networks with no information lost. Regarding your research, the YOLOv7 is identified to possess a bigger ratio of Accuracy to Parameter with the fewer computational resources. Nevertheless, it is anchored-based in nature, which may be a weakness with more irregular, amoeba-like imaging of some brain tumors.

YOLOv8 is the latest popular architecture whose most significant advancement was the transition to an anchor-free detector. Unlike the attempt to apply a tumor to a box that is already shaped, YOLOv8 predicts the center of the object directly. It has also presented the C2f module (Cross-stage Partial bottleneck with two convolutions), which enhances the circulation of spatial information. Although it became much more efficient at the detectability of small objects compared to its predecessors, your work rightly refers to its inefficiency in the medical segmentation, as even faster, it may sometimes give noisy masks, or otherwise have trouble with the accuracy of the border delineation that is required to make the difference between tumor tissue and healthy edema in an MRI. To solve those problems, we suggest using YOLOv11 that incorporates special segmentation strategies to improve the boundaries tracing, space precision and computing efficiency.
Moreover, it is a patient-centered feature introduced in this research: an inbuilt AI chatbot to educate and support patients in their treatment process.

*A. Drawbacks of existing system*

Computer-Aided Diagnosis (CAD) solutions and segmentation techniques require large amounts of computing resources (memories and processors) to run in real-time, which is essential in providing clinical diagnoses in a timely manner. The complexity of most CNN models also implies that they require high computing (memory and processing power) to execute in real time. Segmentation models (such as U-Net, or Mask R-CNN) are able to locate the affected area correctly with the help of a mask, but are more costly to compute. On the other hand, less complex CNN models are usually inherently restricted, with respect to the accuracy with which the exact location of the tumor can be identified. Previous CNN-based approaches and variants of R-CNN are usually characterized by a poor segmentation performance. Deep CNNs can be complicated in nature, thus slow inference time. This slow processing speed is a major practical disadvantage to clinical uses where a high turn-around time of processing image data is essential to surgical planning and timely diagnosis. This drawback is the major reason why real-time models such as the YOLO family were developed. Low accuracies in the localization of tumor regions - It is hard to exactly segment tumors, particularly in cases where they are to HIV and when tumors are to be homogeneous in size and shape. Sensitivity to artifact and noise - The systems can be sensitive to noise and image artifact and can decrease the quality of segmentation.

YOLOv7 architecture and multifaceted training can introduce additional complexity to the model to comprehend, reduce and optimize to close-specific training requirements, unlike the simpler architecture of the YOLOv8. YOLOv7 is mainly a detector of objects. It frequently needs individual, community-oriented implementations to extend it to other tasks such as instance segmentation or pose estimation, without the integrated and multi-task architecture of more recent models such as YOLOv8 and YOLOv11YOLOv8 still needs significant computational resources (powerful GPUs) to train itself, which may be limited in researchers or organizations with limited hardware. Good quality brain tumor samples, especially those that are subdivided in pixel level (ground truth masks), are usually small, patchy and can be varied with scanning protocol or error by human annotation. Constrained access to standardized and large-scale medical data can cause overfitting and bad generalization to the data that has not been seen. Emphasis Detection, Lack Segmentation Quality: These older versions are normally based on detection and rarely do they deeply evaluate segmentation and provide a direct comparison among the advanced versions. It was mentioned that one of the studies employing YOLOv5 failed to execute the segmentation process that deters precise identification of tumor boundaries. High False Positives of Glioma, Glioma Model, models that are used to segment the glioma tumor have high false positives.

The performance of YOLOv5, in comparison, included higher segment validation loss than YOLOv7 meaning that segmentation was poor. Although the accuracy is quite high, the inference time of YOLOv9 is 23.5ms and thus not suitable in diagnosing tumors in emergency situations in real time because of the high rate of false positive detection than the latest version. Moreover, YOLOv10 has a less better detection accuracy on the whole. The original YOLOv8 algorithm cannot detect small objects. The minimal pixel sizes of tumors do tend to lose important characteristics in the down sampling processes of the model, thus causing either no detections or a dramatic effect on accuracy. This is one of the problems with single stage detectors. Brain tumors frequently possess complicated, indistinct edges (particularly diffuse tumors such as gliomas) or can appear to be a part of the adjacent normal tissue or edema. The architecture of YOLOv8, including a segmentation head, can have a hard time defining these unclear boundaries. YOLOv8 scans the image on a global basis, with an ability to still lack sufficient context understanding between complex areas.

| SYSTE M | PRECISIO N(%) | RECALL (%) | mAP@0.5 (%) |
|---|---|---|---|
| YOLO v11 | 97.6 | 97.2 | 97.4 |
| YOLO v10 | 95.7 | 94.3 | 95.5 |
| YOLO v9 | 94.4 | 92.1 | 94.2 |
| YOLO v8 | 93.7 | 90.6 | 93.5 |
| YOLO v7 | 92.5 | 90.3 | 92.1 |
| YOLO v5 | 91.6 | 89.6 | 91.4 |

*Table 2.1 System performance comparison*

## III. PROPOSED METHODOLOGY

The application of yolov11 architecture with deep learning approaches for segmentation system and applying llama as large language model and groq API as language processing unit with natural language processing principles for chatbot feature.

The main element of the diagnosis uses a YOLOv11 architecture, the latest development in the line of You Only Look Once. As compared to its predecessors, YOLOv11 uses C3k2 block and C2PSA (Cross Stage Partial with Spatial Attention) modules. They contain these architectural improvements to enhance segmentation accuracy and precision in pointing tumor boundaries. It is made to train preprocessed MRI images in which noise reduction and contrast enhancement have been done to give the spatial attention mechanisms a chance to differentiate between the necrotic cores, edema, and healthy brain tissue. In the case of the supportive chatbot, the methodology is based on Llama (Large Language Model Meta AI) as the engine. This model is tailored to the medical field based on the principles of Natural Language Processing (NLP) applied to it.
It works based on the principle of translating complicated medical words into the human-friendly language, which is understanding and empathetic. It relies on Retrieval-Augmented Generation (RAG) pipeline, during which the Llama model will be able to read the exact outputs of the underlying segmentation system (e.g. a tumor location or size change) and describe them to the patient in real-time, giving them a step-by-step rundown about their treatment process. The Groq LPU (Language Processing Unit) API is incorporated into the methodology to render the chatbot responsive and live. Conventional GPUs usually cannot keep up with the sequential nature of language and so result in slow AI dialogues. The Groq special hardware architecture is specialized to perform LLM inference and enables Llama model to run tokens at over 300 tokens/sec. This is the language processing unit that takes the heavy, computation-intensive task of ensuring continuity of a flowing conversations.

### *A. Benefits of proposed system*

The biggest benefit of working with YOLOv11 in comparison to the previous ones (such as YOLOv8) is that it delivers the capability of spatial attention. Abnormal tumors in the brain are known to have non geometric shapes. The C2PSA block of YOLOv11 enables the model to pay close attention to minor textural differences in scans of MRI and greatly lowers the number of false positives, in addition to the margins of the tumor being scanned being detected with some accuracy. This accuracy is an imperative to neurosurgeons that use precise limits when planning the surgery or when aiming for radiation beams. Through using the Groq API, Patients with high anxiety wishes to have quick responses in their queries regarding their symptoms or treatment schedules. Groq has deterministic execution, and this feature makes the chatbot react to queries almost immediately resembling a real-human interaction. This speed does not only seem like a luxury, but it will save frustration on the part of the patients, and vital suggestions like visiting the hospital so urgently can be conveyed without any postponements occurring. Llama has made a cold, technical diagnostic tool warmly different and a supportive companion.

This enhances patient trustability, when a patient knows the reason as to why particular treatment is taking place, or whether a scan result signifies something, his or her adherence to the treatment procedures increases, resulting in the overall improved health.

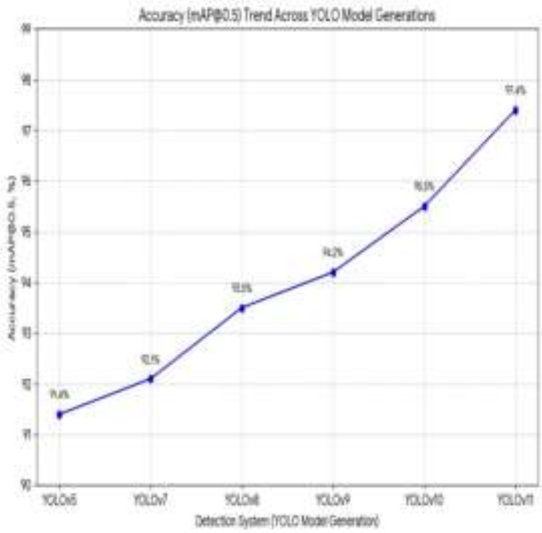


*Fig 3.1 Yolo performance graph*

*B. Implementation stages*

**1. Data collection**

The mri images of brain with and without tumor are collected from an open-source website for training the model.

**2. Data preprocessing**

The raw MRI images are preprocessed to improve the quality of images and minimize noise. This step is critical to enhance the level of segmentation and to guarantee that the model works properly and the data is annotated.

**3. Model Training**

The yolov11 model is trained with preprocessed and annotated mri images to provide segmentation results with the help of its convolutional. layers, spatial pyramid pooling, and a head of segmentation.

**4. LLM Integration with LPU**

For implementing chatbot API, the large language model (Llama) application which will act as a database for the chatbot framework is integrated with language processing unit (GUAQ cloud) which will allow developers to act and utilize LPU

**5. Prompt Engineering**

Prompt Engineering is implemented to fix response limitations (hallucinations), you must use a System Prompt that enforces a clinical persona and structured reasoning.

**6. Flask API Integration with frontend UI** A Flask web interface that handles backend logic, interacts with the database integrated with frontend UI that interacts with the Flask API to display data and handle user input as a single interface that enables chatbot feature.

**7. Prediction System**

After training, the YOLOv11 model is deployed in a clinical environment where it can be used to segment tumors in real-time MRI scans. The system generates binary masks that delineate tumor regions, assisting clinicians in diagnosis and treatment planning.

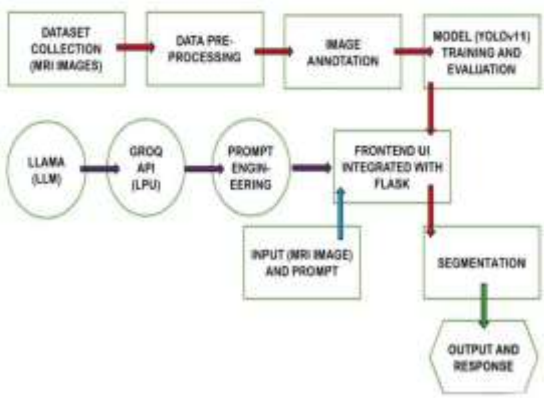


*Fig 3.2 System architecture*

## IV. EXPERIMENTAL RESULTS

The supportive chatbot feature integration brought positive qualitative feedback about the use of the chatbot by the patients. The chatbot was effective in alleviating uncertainty in anxiety related to uncertain symptoms because it acted as a translation between medical information and patient knowledge.

Patients could be provided with real-time information on their condition, and it directly affected the way they made the decisions to visit the hospital. This two-facet approach to bridging the gap between a high-quality computer vision and a patient-central communication tool has created a new standard of certainty, known as trustability within the automated medical treatment systems.

As the experimental outcome has proven, the change made between YOLOv8 and YOLOv11 is far more effective in improving the accuracy of brain tumor segmentation. The YOLOv11 model was able to overcome the properties of irregular shapes of tumors and the variability of sizes that were unmanageable by simple use of the previous architecture by addressing the underlying inefficiencies of the former architecture. The pipeline of preprocessing was effective in enhancing tumor boundaries of the MRI Scan, which the deep learning model was able to realize higher Dice coefficients and intersection-over-union (IoU) scores than the existing baseline methods. In addition, the automated segmentation strategy was implemented and it yielded more information on the localization and spatial positioning of tumors. This granular data is essential in clinical decision making because it is possible to have a more precise examination of how the tumor behaves with the surrounding brain tissue. These high-accuracy results were also only made possible by the efficiency of YOLOv11, which controlled the computing latency to a minimum, thus it becomes a viable candidate to serve as real-time clinical diagnostic support.

*A. Conclusion*

This study concludes that YOLOv11 is a better architecture to segment the brain tumors using because it is able to surpass some limitations that YOLOv8 has when dealing with complex MRI morphologies. The study, through inclusion of an innovative patient-education chatbot, does not just cover the technical accuracy, but it also aims at covering the psychological and logistical requirements of the patient. Finally, having advanced deep learning disease segmentation tools as well as having an interactive patient support system will only make the process of tumor treatment more precise clinically, as well as increase the overall accuracy of the healthcare experience.

## REFERENCES

(1) Key Statistics for Brain and Spinal Cord Tumors, May 2023, [online]Available: https://www.cancer.org/cancer/types/brain-spinal- cord-tumors-adults/about/key-statistics.html.

(2) S. Lapointe, A. Perry and N. A. Butowski, "Primary brain tumours in adults", Lancet, vol. 392, no. 10145,pp. 432-446, Aug. 2018.

(3) E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, et al., "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme", Magn. Reson. Med., vol. 62, no. 6, pp. 1609-1618, Dec. 2009.

(4) J. Liu, M. Li, J. Wang, F. Wu, T. Liu and Y. Pan, "A survey of MRI-based brain tumor segmentation methods", Tsinghua Sci. Technol., vol. 19, no. 6, pp. 578-595, Dec. 2014.

(5) N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani and T. R. Mengko, "Brain tumor classification using convolutional neural network",pp. 183-189, 2019.

(6) J. Amin, M. Sharif, A. Haldorai, M. Yasmin and R. S. Nayak, "Brain tumor detection and classification using machine learning: A comprehensive survey", Aug. 2022.

Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, et al., "Brain tumor classification for MR images using transfer learning and fine-tuning", Computerized Med. Imag. Graph., vol. 75, pp. 34-46, Jul. 2019.

(7) J. Watts, G. Box, A. Galvin, P. Brotchie, N. Trost and T. Sutherland, "Magnetic resonance imaging of meningiomas: A pictorial review", Insights Imag., vol. 5, no. 1, pp. 113-122, Feb. 2014.

(8) G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, et al., "A survey on deep learning in medical image analysis", Med. Image Anal., vol. 42, pp. 60-88, Dec. 2017.

(9) A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI", Zeitschrift Für Medizinische Physik, vol. 29, no. 2, pp. 102-127, May 2019.

(10) G. Alwakid, W. Gouda, M. Humayun and N. Z. Jhanjhi, "Diagnosing melanomas in dermoscopy images using deep learning", Diagnostics, vol. 13, no. 10, pp. 1815, May 2023.

(11) M. F. Almufareh, S. Tehsin, M. Humayun and S. Kausar, "A transfer learning approach for clinical detection support of monkeypox skin lesions", Diagnostics, vol. 13, no. 8, pp. 1503, Apr. 2023.

(12) M. Humayun, M. I. Khalil, S. N. Almuayqil and N. Z. Jhanjhi, "Framework for detecting breast cancer risk presence using deep learning", Electronics, vol. 12, no. 2, pp. 403, Jan. 2023.

(13) A. Khan, J. Bukhari, J. I. Bangash, A. Khan, M. Imran, M. Asim, et al., "Optimizing connection weights of functional link neural network using APSO algorithm for medical data classification", J. King Saud Univ.- Comput. Inf. Sci., vol. 34, no. 6, pp. 2551-2561, Jun. 2022.

(14) Z.-Q. Zhao, P. Zheng, S.-T. Xu and X. Wu, "Object detection with deep learning: A review", IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 11, pp. 3212- 3232, Nov. 2019.

(15) M. Imran, H. Anwar, M. Tufail, A. Khan, M. Khan and D. A. Ramli, "Image-based automatic energy meter reading using deep learning", Comput. Mater. Continua, vol. 74, no. 1, pp. 203-216, 2023.

(16) G. Alwakid, W. Gouda and M. Humayun, "Deep learning-based prediction of diabetic retinopathy using CLAHE and ESRGAN for enhancement", Proc. MDPI, pp. 863, 2023.

(17) W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, Y. Fu, et al., "SSD: Single shot MultiBox detector" in Computer Vision—ECCV, Amsterdam, The Netherlands:Springer, pp. 21-37, 2016.

(18) R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 580-587, Jun. 2014.

(19) S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, Jun. 2017.