Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

Semantic Segmentation Based on Geometric Calibration Using AI and AR in Health Care

G. Ganesan*, **S. Poonkuntran**

School of Computing Science and Engineering, VIT Bhopal University, Bhopal, India
E-mail: ggmukil123@gmail.com

Received: 1 January 2024; **Revised:** 12 March 2024; **Accepted:** 13 March 2024

Abstract: In the medical field, medical imaging is essential for precise diagnosis, treatment planning, and condition monitoring. The goal of this work is to improve the field of healthcare imaging by investigating the complementary applications of augmented reality (AR) and artificial intelligence (AI) in semantic segmentation. More precise identification and delineation of anatomical structures and abnormalities in medical images is made possible by AI-driven semantic segmentation, opening the door to more precise diagnosis and treatment plans. Modern AI algorithms are employed in the suggested methodology to perform semantic segmentation in a variety of medical imaging modalities, including ultrasound, CT, and MRI scans. The detection of anomalies such as tumours, irregularities in organs, and neurological disorders is made easier by these algorithms. The foundation for integrating augmented reality technologies into the healthcare ecosystem is provided by the segmented medical images. AR enhances the interaction and visualization of segmented medical data in the context of healthcare. Real-time augmented reality overlays help surgeons by improving surgical navigation and precision. Additionally, AR apps help with medical education by offering professionals and students alike an immersive learning environment. AR provides interactive visualizations to help patients understand their medical conditions when they are going for therapy and rehabilitation. The difficulties in integrating these technologies in the healthcare industry are discussed in this paper, with a focus on the significance of data privacy, regulatory compliance, and easy integration with current healthcare systems. The successful deployment of AI technologies necessitates interdisciplinary collaboration among AI developers, healthcare professionals, and AR specialists to ensure compliance with ethical, legal, and clinical standards mandated in the healthcare domain. The proposed Segmentation mask is redefined with geometric calibration and used with U-Net for image segmentation. The segmentation is experimented on Cityscapes and PASCAL VOC 2012 datasets. The experimental results show that proposed semantic segmentation based on geometric calibration yields more accuracy than its counterparts.

Keywords: geometric calibration, semantic segmentation, artificial intelligence (AI), augmented reality (AR), healthcare imaging, computer vision

MSC: 62P10, 68U10, 97R40

DOI: https://doi.org/10.37256/cm.5420244218

1. Introduction

The integration of cutting-edge technologies has become a transformative force in the modern healthcare landscape, offering previously unheard-of capabilities in patient care, treatment planning, and diagnosis. The convergence of Augmented Reality (AR) and Artificial Intelligence (AI) in the field of semantic segmentation is one such ground-breaking intersection [1]. This combination has the potential to completely transform medical imaging and give doctors new and effective tools to improve accuracy, productivity, and patient outcomes [2].

The foundation of clinical decision-making is medical imaging, which provides a non-invasive window into the minute details of the human body. But one challenge brought on by the abundance of complex data produced by imaging modalities is the requirement for precise and effective analysis [3, 4]. Here comes semantic segmentation, an advanced AI-powered method that can accurately identify and classify structures in medical images.

AI has proven incredibly adept at semantic segmentation tasks, especially when it comes to sophisticated machine learning algorithms and neural networks [5, 6]. These algorithms open the door to more accurate diagnosis by being able to recognize and separate anatomical structures, anomalies, and diseases on their own in medical images. AI-driven semantic segmentation is revolutionizing medical imaging by being able to identify minute details in ultrasound images and detect subtle nuances in MRI scans. Healthcare imaging takes on a new dimension with the integration of Augmented Reality, as if the revolutionary powers of AI alone weren't enough.

Artificial Intelligence (AI) and Augmented Reality (AR) are increasingly utilized in medical imaging to enhance diagnostic accuracy, improve treatment planning, and facilitate surgical procedures. Here's how AI and AR are currently applied in medical imaging:

• AI-Assisted Diagnosis: AI algorithms are trained on extensive datasets of medical images to detect patterns and anomalies that may elude human observers. For instance, AI systems assist radiologists in identifying abnormalities in X-rays, MRIs, CT scans, and other imaging modalities. These systems contribute to enhancing diagnostic accuracy, reducing interpretation errors, and expediting the diagnostic process.

• Image Segmentation: AI algorithms can segment medical images into different regions or structures, enabling more precise analysis and measurement. This capability is particularly beneficial in treatment planning for conditions such as tumors, where accurate delineation of tumor boundaries is critical for planning radiation therapy or surgical interventions.

• Personalized Treatment Planning: AI analyzes medical imaging data alongside other patient information to customize treatment plans for individual patients. For example, AI algorithms predict tumor response to various treatment modalities based on imaging characteristics, genetic information, and clinical data, enabling oncologists to develop personalized treatment strategies.

• Surgical Navigation: AR technology overlays digital information, like preoperative imaging data or anatomical landmarks, onto a surgeon's field of view during surgery. This allows surgeons to visualize internal structures in real-time and precisely navigate surgical instruments, thereby enhancing surgical accuracy and reducing complications. AR-assisted surgery is particularly valuable for complex procedures such as neurosurgery and orthopedic surgery.

• Medical Education and Training: AR technology offers medical students and healthcare professionals immersive learning experiences by overlaying virtual anatomy models or simulated surgical procedures onto real-world environments. This hands-on approach to medical education enables learners to interact with anatomical structures and practice procedures in a safe and controlled environment.

• Remote Consultation and Telementoring: AR-enabled platforms enable experts to provide remote guidance and mentoring to healthcare providers in real-time. By overlaying virtual annotations or instructions onto live video streams of medical procedures, AR technology facilitates collaboration and knowledge transfer between healthcare professionals, regardless of their physical location.

In summary, AI and AR are revolutionizing medical imaging by improving diagnostic capabilities, treatment outcomes, and medical education and surgical practices. As these technologies advance further, they hold significant promise for continued innovation in healthcare.

Augmented Reality (AR) has found its place in various healthcare realms, such as surgical navigation, medical education, therapy, and rehabilitation, each benefiting uniquely from its integration. Let's delve into specific AR technologies or platforms utilized in these areas and their respective advantages:

1. Surgical Navigation:

- HoloLens by Microsoft: This mixed reality headset overlays holographic images onto the user's real-world view. In surgical navigation, it aids by displaying patient anatomy directly onto the surgeon's field of view during procedures.

- Advantages: AR enhances precision and accuracy in surgical navigation by offering real-time, three-dimensional visualization of patient anatomy. Surgeons can better comprehend complex anatomical structures, leading to more effective surgical planning and reduced risks for patients.

2. Medical Education:

- Complete Anatomy by 3D4Medical: This AR platform facilitates detailed exploration of 3D models of the human body. It provides interactive experiences for medical students and professionals to study anatomy.

- Advantages: AR in medical education offers immersive learning experiences, allowing users to visualize and interact with anatomical structures in three dimensions. It fosters better understanding and retention of complex concepts, promotes active learning, and enables simulated surgical procedures without reliance on cadavers.

3. Therapy:

- Rehabilitation Gaming System (RGS) by Jintronix: RGS is an AR platform designed for physical therapy and rehabilitation. It utilizes motion-tracking sensors and AR technology to create interactive exercises for patients recovering from injuries or surgeries.

- Advantages: AR-based therapy platforms provide engaging and personalized rehabilitation exercises, motivating patients to actively participate in their recovery process. Real-time feedback and progress tracking empower therapists to monitor patient performance and adjust treatment plans accordingly, leading to expedited and more effective recovery outcomes.

4. Rehabilitation:

- Virti: Virti is an AR-based training platform simulating real-world scenarios for healthcare professionals, encompassing therapy and rehabilitation scenarios. It offers immersive training experiences using AR technology.

- Advantages: AR in rehabilitation training allows healthcare professionals to practice various procedures and interventions in a secure and controlled environment. It fosters skill development, enhances decision-making abilities, and boosts confidence, ultimately improving patient care and safety.

In conclusion, AR technologies and platforms deliver significant advantages across surgical navigation, medical education, therapy, and rehabilitation. They offer immersive visualization, interactive learning experiences, personalized therapy sessions, and realistic training simulations, thereby enhancing patient outcomes, healthcare professional training, and the efficiency of healthcare delivery. As AR continues to advance, its applications in healthcare are poised to expand, revolutionizing medical procedures, educational methodologies, and therapeutic practices.

Clinicians can experience an immersive and interactive medical data overlay with augmented reality (AR) [7]. The combination of AI and AR has the potential to completely change how medical professionals use medical images, allowing for improved decision-making and real-time insights.

This paper explores the various uses of AI and AR-powered semantic segmentation in healthcare. We investigate its possible effects on disease diagnosis, planning of treatments, surgical guidance, medical education, patient education, and rehabilitation. We want to highlight the revolutionary potential these technologies hold for raising the bar for healthcare delivery as we work through the difficulties and complexities of integrating them in a healthcare setting [8].

Here are some key challenges associated with integrating AI and AR in healthcare:

1. Privacy and Security:

- AI: Utilizing vast datasets containing sensitive patient information raises concerns about data privacy and security. Compliance with regulations like HIPAA is essential.

- AR: AR devices capturing and displaying real-time patient data necessitates robust safeguards against unauthorized access and cyber threats to uphold patient confidentiality.

2. Data Interoperability:

- AI: Healthcare systems often use varied formats for storing patient data, hindering seamless integration with AI algorithms.

- AR: AR applications must integrate with Electronic Health Records (EHRs) and other healthcare IT systems, requiring compatibility and smooth data exchange for effective implementation.

3. Regulatory Compliance:

- AI: Evolving regulatory frameworks for AI in healthcare necessitate compliance with existing regulations and adaptation to new guidelines for widespread adoption.

- AR: Regulatory challenges arise regarding AR usage in medical procedures and training, highlighting the need for clear guidelines and standards to ensure ethical practices.

4. Standardization:

- AI: The absence of standardized protocols for AI model development and validation impedes interoperability and collaborative efforts.

- AR: Standardizing AR technologies and interfaces is crucial to promote compatibility across various devices and applications.

5. Ethical Concerns:

- AI: Addressing ethical issues such as bias in AI algorithms, transparency, and accountability is vital to ensure fairness and unbiased healthcare outcomes.

- AR: Ethical considerations regarding patient consent, proper use, and potential psychological impacts arise with the use of AR in sensitive medical procedures.

6. Integration with Existing Workflows:

- AI: Seamless integration of AI tools into clinical workflows without disruption is challenging, as physicians may resist adoption if it disrupts established practices.

- AR: Implementing AR in healthcare requires adjustments to established procedures to ensure that AR applications enhance rather than impede clinical workflows.

7. Limited Clinical Validation:

- AI: Robust clinical validation of AI algorithms is necessary to gain trust among healthcare professionals and foster widespread acceptance.

- AR: Thorough validation of AR applications in various medical scenarios is crucial to ensure safety and effectiveness.

8. Cost and Resource Constraints:

- AI: Developing and implementing AI solutions can be resource-intensive, posing challenges for smaller healthcare facilities in terms of cost and expertise.

- AR: Acquiring and maintaining AR devices and infrastructure can be costly, requiring considerations for accessibility across different healthcare settings.

9. User Training and Acceptance:

- AI: Healthcare professionals may require training to effectively utilize AI tools, and overcoming resistance to adoption is essential.

- AR: Adequate training for healthcare professionals on AR device usage is necessary to ensure user acceptance and address potential resistance.

10. Technical Limitations:

- AI: Current AI models may struggle with complex medical scenarios or rare conditions, necessitating ongoing research for algorithm improvement.

- AR: Technical challenges such as device limitations and real-time tracking accuracy require continuous refinement to optimize performance in diverse medical settings.

Addressing these challenges requires collaboration among healthcare professionals, technology developers, regulators, and other stakeholders. Overcoming these hurdles will pave the way for the responsible and effective integration of AI and AR technologies in healthcare.

To foster interdisciplinary collaboration among AI developers, healthcare professionals, and AR specialists, a strategic approach and effective communication channels are essential. Here's how this collaboration can be achieved:

• Establish Common Goals and Objectives: Defining shared goals and objectives that align with the needs and challenges of healthcare creates a common purpose, motivating collaboration among different disciplines.

• Create Cross-Disciplinary Teams: Form interdisciplinary teams consisting of AI developers, healthcare professionals (such as physicians, nurses, and researchers), and AR specialists. Each team member brings unique expertise and perspectives, fostering innovation and problem-solving.

• Promote Open Communication: Encourage open communication and collaboration channels among team members through regular meetings, brainstorming sessions, and collaborative tools. This facilitates idea exchange and problem-solving.

• Provide Interdisciplinary Training: Offer training programs or workshops to allow team members to better understand each other's disciplines. This promotes mutual respect and appreciation for the challenges and opportunities faced by each discipline.

• Facilitate Knowledge Sharing: Create platforms or forums where AI developers, healthcare professionals, and AR specialists can share knowledge, best practices, and insights from their respective fields. This encourages cross-pollination of ideas and fosters a culture of learning.

• Encourage Interdisciplinary Research: Support interdisciplinary research projects that combine AI, healthcare, and AR expertise to address complex healthcare challenges. Funding opportunities and research grants can incentivize collaboration and innovation.

• Provide Resources and Infrastructure: Ensure interdisciplinary teams have access to necessary resources, infrastructure, and technology to support their collaborative efforts. This includes access to datasets, computing resources, AR devices, and healthcare facilities for testing and validation.

• Promote Leadership and Collaboration Skills: Develop leadership and collaboration skills among team members to effectively manage interdisciplinary projects. Training in project management, conflict resolution, and team dynamics enhances collaboration effectiveness.

• Foster a Culture of Innovation: Cultivate a culture that values innovation, creativity, and experimentation. Encourage risk-taking and exploration of novel approaches to healthcare challenges, fostering an environment where interdisciplinary collaboration thrives.

• Celebrate Success and Recognize Contributions: Acknowledge and celebrate the achievements of interdisciplinary teams. Recognizing individual contributions and team successes fosters a sense of accomplishment and motivation for future collaboration.

By implementing these strategies, interdisciplinary collaboration among AI developers, healthcare professionals, and AR specialists can be effectively nurtured, leading to innovative solutions that address complex healthcare challenges. Certainly, ensuring data privacy and regulatory compliance is paramount when integrating AR technologies into healthcare. Here are potential solutions and strategies for addressing challenges while upholding these principles:

1. Addressing Technical Challenges:

- Interoperability: Encourage collaboration between AR developers and healthcare system providers to establish interoperability standards, ensuring seamless integration with existing IT infrastructure.

- Hardware Limitations: Forge partnerships with hardware manufacturers to create affordable AR solutions tailored for healthcare, leveraging ubiquitous devices like smartphones and tablets to mitigate costs.

- Training and Adoption: Develop comprehensive training programs for healthcare professionals to proficiently utilize AR technologies, providing ongoing support to facilitate adoption and overcome potential hurdles.

2. Ensuring Regulatory Compliance and Data Privacy:

- HIPAA Compliance: Collaborate with legal experts to embed HIPAA-compliant security features into AR applications, including data encryption, access controls, and regular audits to maintain compliance.

- GDPR Compliance: Adhere to GDPR regulations by implementing transparent data processing practices, obtaining explicit consent for data collection, and providing mechanisms for data subject access and erasure.

- Ethical Data Use: Establish clear policies for ethical data usage in AR applications, prioritizing patient consent, anonymization when feasible, and bias mitigation to ensure fairness.

3. Managing Cost and ROI:

- Cost-Benefit Analysis: Conduct thorough cost-benefit assessments to showcase the potential ROI of AR implementation in healthcare, emphasizing benefits such as improved patient outcomes and enhanced efficiency in education and therapy.

- Reimbursement Policies: Advocate for reimbursement policies that incentivize healthcare providers to adopt AR technologies, highlighting potential cost savings and enhanced patient care to justify investment.

In conclusion, a collaborative effort involving stakeholders from technology development, healthcare, regulation, and advocacy is essential for successfully integrating AR into healthcare while upholding data privacy and regulatory compliance. Prioritizing these principles will foster trust among stakeholders and ensure the responsible and ethical use of AR in healthcare environments.

The approaches for AI-based semantic segmentation, integrating augmented reality into healthcare workflows, and interdisciplinary cooperation necessary for successful implementation will all be covered in the following sections of this paper [9]. We hope that this investigation will shed light on how healthcare imaging is developing and the significant implications that these technologies will have for patient care in the future.

2. Semantic segmentation using AI and AR

The application of augmented reality (AR) and artificial intelligence (AI) to semantic segmentation in healthcare has the potential to significantly improve medical imaging, diagnosis, and treatment planning. Semantic segmentation driven by artificial intelligence (AI) can precisely recognize and delineate various anatomical structures and abnormalities in medical images. When analyzing ultrasound images, segmentation can help with the identification and measurement of organs, tumors, and other abnormalities.

Radiologists can receive assistance in early detection and accurate diagnosis by using AI algorithms that can segment and highlight tumors in medical images. Brain structure segmentation can aid in the diagnosis of diseases such as multiple sclerosis and Alzheimer's. Planning for radiation therapy becomes more precise and focused thanks to semantic segmentation, which aids in defining the target and surrounding tissues [10]. Segmented images help surgeons plan and visualize procedures, increasing accuracy and reducing harm to healthy tissues.

During a procedure, augmented reality superimposes segmented medical images onto the surgeon's view, improving accuracy and offering real-time guidance. By visualizing important structures and pathways, surgeons can perform minimally invasive surgeries with greater precision. By superimposing segmented 3D models over real anatomy, augmented reality (AR) can be utilized in medical education to provide a more engaging and dynamic learning environment. In a safe virtual setting, trainees can understand anatomical variations and practice procedures.

Patients can better understand their conditions and treatment plans by using augmented reality (AR) applications to visualize segmented medical data. This facilitates shared decision-making between patients and healthcare professionals and raises patient engagement [11]. By superimposing visual cues onto the actual environment, AR can assist patients with rehabilitation exercises and guarantee correct form and adherence to prescribed exercises. Additionally, it can monitor development over time, encouraging patients and giving medical professionals useful information.

Strong safeguards must be in place to protect patient information because healthcare data is sensitive. It is imperative to comply with regulatory frameworks like HIPAA in order to guarantee the lawful and moral application of AI and AR technologies in the healthcare industry [12]. Widespread adoption requires seamless integration with medical imaging platforms, electronic health records (EHRs), and current healthcare systems. In order to develop solutions that effectively address clinical needs, cooperation between AI developers, healthcare professionals, and AR specialists is necessary for successful implementation [13].

3. Literature survey analysis

Semantic segmentation using AI and AR in healthcare is a rapidly developing field with significant advancements and a wide range of applications, according to a literature survey analysis. Numerous facets, such as algorithm development, integration strategies, and clinical implementations, have been investigated by researchers and practitioners.

Creating strong AI algorithms for semantic segmentation in medical imaging has been the subject of much research. Across various imaging modalities, Convolutional Neural Networks (CNNs), U-Net architectures, and attention mechanisms have proven to perform better at accurately segmenting organs, tumors, and anomalies. Research acknowledges the significance of amalgamating data from various imaging modalities to augment the precision of segmentation [14]. Improved accuracy in identifying intricate anatomical structures and abnormalities is demonstrated by fusion approaches that incorporate data from ultrasound, CT, and MRI scans.

The potential for augmented reality to provide real-time guidance during surgical procedures has drawn attention to its integration into surgical workflows. AR overlays that are produced by preoperative semantic segmentation provide surgeons with better visualization of important structures, which helps them, operate more precisely during minimally invasive procedures [15]. The importance of AI-driven semantic segmentation in medical education is highlighted in the literature. Through the use of segmented 3D models in AR-based simulations, learners can engage with and comprehend intricate anatomical structures, improving their abilities and spatial awareness [16].

Studies examine augmented reality apps intended for patient involvement and education. Patients become more knowledgeable and empowered when they are able to understand their conditions, treatment options, and possible results thanks to the segmented medical data that is displayed through augmented reality interfaces [17]. Notwithstanding encouraging developments, research notes issues with data security, privacy, and the moral application of AI and AR in healthcare. Regulatory compliance becomes apparent as a crucial factor that needs careful thought, particularly when it comes to patient data.

In March 2023, Microsoft released a software to automate clinical documentation using AI named as Dragon Ambient eXperience Express (DAX). It's an automated clinical documentation tool seamlessly integrated into existing workflows, marking the first instance where established conversational and ambient AI technologies are combined with the advanced reasoning and natural language capabilities of OpenAI's GPT-4. Building upon the success of Nuance's Dragon Medical solutions and leveraging the groundwork laid by the DAX ambient solution introduced in 2020, DAX Express represents the next step forward in Nuance's ongoing mission to alleviate administrative burdens and enable clinicians to devote more of their time to patient care rather than paperwork [18].

AR glasses attracted many surgeons around the globe. This glasses will be more powerful and useful when its paired with AI [19].

For an example of healthcare services, Thi Thi Zin et al. proposed elderly people care monitoring system using image processing [20] and sensors [21].

Collaboration between computer scientists, healthcare professionals, and AR specialists is necessary for the successful implementation of semantic segmentation using AI and AR. In order to ensure practical utility, the literature emphasizes the significance of seamless integration with current healthcare systems, such as Electronic Health Records (EHRs). Future research directions are covered in a number of papers. These include developing standardized protocols for AR integration, improving algorithms for particular medical domains, and investigating the use of AI and AR in telemedicine and personalized medicine applications.

4. Existing approches

FCNs allow end-to-end pixel-wise predictions, extending the concept of convolutional neural networks for semantic segmentation tasks. In order to improve the focus on pertinent regions during segmentation, attention mechanisms have been implemented. When it comes to capturing spatial relationships, spatial attention is helpful, but channel attention concentrates on key channels.

Increased segmentation accuracy has been demonstrated by combining data from several imaging modalities, such as CT and MRI scans. Fusion approaches take advantage of the complimentary qualities of various modalities. Semantic segmentation and other image-to-image translation tasks have been performed with GANs. High-resolution segmented images have been successfully produced by Pix2Pix GANs in particular. More accurate segmentation has been achieved by developing approaches that take into account the volumetric nature of the data, particularly in applications like medical imaging where the data is intrinsically three-dimensional. Markerless tracking is used by AR systems for surgical navigation to overlay segmented medical data in real-time onto a surgeon's view. This improves accuracy when performing surgeries. Interactive 3D models are made for medical training simulations using AI-driven segmentation. Through augmented reality interfaces, trainees can interact with these models and enhance their comprehension of intricate anatomical structures. In order to customize segmentation models to specific patient data, AI algorithms are currently being developed. This tailored strategy improves segmentation accuracy for particular patient attributes.

Efficient integration of AI and AR applications with current EHR systems has been attempted in an attempt to make the technology more useful for medical practitioners. Federated learning techniques have been investigated to address privacy concerns. With the help of these techniques, AI models can be trained on several dispersed devices without requiring the exchange of raw patient data.

The importance of healthcare decisions has led to an increasing focus on creating AI models that can explain predictions and improve communication and trust between medical professionals. Improvements in semantic segmentation in the healthcare industry are facilitated by collaborative efforts, which are frequently in the form of open-source projects that allow the sharing of algorithms, datasets, and tools. Semantic segmentation AR applications have been investigated for remote consultations, where medical practitioners can visualize and discuss medical images with patients in real-time.

5. Proposed method

Semantic segmentation plays a critical role in computer vision, aiming to categorize each pixel within an image into specific classes or categories. This task varies across different imaging modalities, including optical images, medical images like MRI or CT scans, and satellite images, each necessitating distinct approaches. Here's an overview of several prevalent AI algorithms used for semantic segmentation across various imaging modalities:

• Convolutional Neural Networks (CNNs):

CNNs serve as the foundation for many cutting-edge semantic segmentation algorithms due to their capacity to automatically extract hierarchical features from data. Architectures such as U-Net, Fully Convolutional Networks (FCNs), and DeepLab are frequently chosen for semantic segmentation tasks. CNNs can be customized for different modalities by adjusting input data preprocessing, network architecture, and loss functions.

• U-Net:

U-Net stands out as a widely adopted CNN architecture tailored for biomedical image segmentation tasks. Comprising a contracting path for contextual capture and a symmetric expanding path for precise localization, U-Net has demonstrated success across various medical imaging modalities like MRI, CT scans, and histopathology images.

• DeepLab:

DeepLab represents another prominent CNN architecture recognized for its robust performance in semantic segmentation tasks. It utilizes atrous convolution (also known as dilated convolution) to effectively capture multi-scale contextual information. DeepLab includes variants like DeepLabv3 and DeepLabv3+, which integrate enhancements for improved performance and efficiency.

• Graph-based Methods:

Graph-based methods frame semantic segmentation as a graph partitioning problem, with nodes representing image elements (e.g., pixels) and edges denoting pairwise affinities between elements. Techniques like graph cuts, random walks, and conditional random fields (CRFs) are commonly employed in this approach. While adaptable to various modalities, graph-based methods excel in scenarios with structured or interconnected elements, such as medical images.

• Region-based Methods:

Region-based methods segment images into regions sharing similar properties using approaches like region growing, watershed segmentation, or mean-shift clustering. These methods often rely on handcrafted features or clustering algorithms to group pixels into cohesive segments. Although not as prevalent in modern deep learning-based approaches, region-based methods remain effective, particularly in scenarios with limited training data or computational resources.

● Transfer Learning and Pretrained Models:

Transfer learning entails leveraging knowledge from pre-trained models on extensive datasets (e.g., ImageNet) and fine-tuning them for specific segmentation tasks. Pretrained CNN architectures like VGG, ResNet, and MobileNet serve as feature extractors in semantic segmentation pipelines. Transfer learning proves advantageous when confronted with limited annotated data, facilitating better generalization from large datasets to specific tasks or modalities.

● Attention Mechanisms:

Attention mechanisms empower models to concentrate on relevant image regions while suppressing irrelevant information, thus enhancing segmentation accuracy and efficiency. Techniques like self-attention mechanisms, spatial attention, and channel-wise attention have been seamlessly integrated into CNN architectures for semantic segmentation. Attention mechanisms prove particularly beneficial in scenarios with intricate backgrounds or where accurate segmentation of specific regions of interest is paramount.

In conclusion, semantic segmentation across diverse imaging modalities relies on a range of AI algorithms and techniques, encompassing CNNs, graph-based methods, region-based methods, transfer learning, and attention mechanisms. The selection of an appropriate algorithm hinges on factors such as data characteristics, availability of labeled training data.

To take advantage of learned features, applied transfer learning to pretrained models on sizable datasets (like Cityscapes and PASCAL VOC). Used fusion techniques to combine information from various imaging modalities when working with multi-modal data to improve segmentation accuracy. Given the nature of semantic segmentation tasks, define a suitable loss function (e.g., dice loss, cross-entropy loss). For instance, the U-Net architecture is depicted in figure 1. The blue boxes correspond to feature maps blocks with their denoted shapes. The white boxes correspond to the copied and cropped feature maps.

To train the model, used optimization techniques such as Adam or SGD. For stable convergence, used learning rate scheduling. To keep an eye on the model's performance during training, set aside a validation dataset. To increase segmentation accuracy, tweak the model and modify hyperparameters in light of validation outcomes. Selected an augmented reality framework, like ARKit for iOS or ARCore for Android, that is appropriate for healthcare apps.

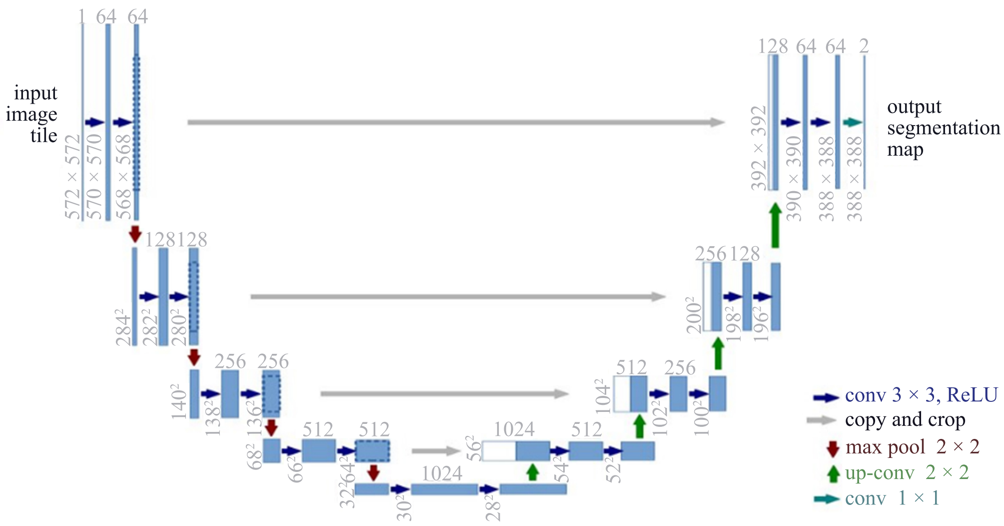

Figure 1. U-Net Architecture. Source: https://arxiv.org/pdf/1505.04597.pdf

Align segmented results with real-world coordinates during AR visualization by implementing coordinate transformation algorithms. In order to enable the overlay of segmented data onto the surgeon's view, develop markerless tracking algorithms for real-time augmented reality surgical navigation. Combine AR segmentation with current surgical navigation systems to improve accuracy.

On a different test dataset, assess the segmentation model using quantitative metrics (such as the IoU and dice coefficient). Examine the segmentation results' accuracy visually in relation to the ground truth to conduct qualitative evaluations. Secure patient data by implementing encryption techniques, especially when handling sensitive medical data.

Assure adherence to healthcare laws and guidelines (such as HIPAA) pertaining to the privacy of patient data. Create user interfaces that make it simple for patients and healthcare professionals to interact with augmented reality visualizations. Integrate informative elements into augmented reality interfaces to offer understanding of the divided medical data.

Work together with medical experts to confirm the suggested method's efficacy in actual clinical settings. Optimize the approach by making modifications based on input from medical professionals to improve clinical usability. Given the suggested method comprehensive documentation that covers all of the steps, algorithms, and technologies used. Think about sharing the code with the larger research community by making it open-source.

Definition 5.1 Let X be the input image and Y be the corresponding ground truth segmentation mask. The goal is to learn a mapping $F : X \to Y$using a neural network.

$$Y_{predicted} = F(X; \theta) \tag{1}$$

Y is the predicted segmentation mask.

F represents the neural network (e.g., a convolutional neural network).

X is the input medical image.

θ denotes the parameters of the neural network that are learned during training.

The training of the model involves minimizing a loss function that quantifies the difference between the predicted segmentation mask and the ground truth mask.

$$(L(\theta) = Loss\left(Y_{predicted}, Y_{groundtruth}\right) \tag{2}$$

$L(\theta)$ is the loss function.

$Y_{groundtruth}$ is the ground truth segmentation mask.

θ represents the parameters of the neural network.

Commonly used loss functions for semantic segmentation include cross-entropy loss, pixel-wise cross-entropy loss, or more advanced options like the dice loss or focal loss.

The model is trained by updating the parameters (θ) using an optimization algorithm, typically stochastic gradient descent (SGD) or one of its variants, to minimize the loss function over the training dataset.

$$(\theta \leftarrow \theta - y\nabla_{\theta} L(\theta) \tag{3}$$

y is the learning rate.

$\nabla_{\theta} L(\theta)$ is the gradient of the loss with respect to the parameters.

The integration of semantic segmentation with augmented reality involves overlaying the predicted segmentation mask onto the real-world view. The transformation from image coordinates to real-world coordinates is often performed using geometric calibration and registration techniques.

$$(AR_{output} = AR_{input} + Transformation\left(Y_{predicted}\right) \tag{4}$$

AR_{output} is the augmented reality output.

AR_{input} is the real-world view captured by the AR device.

$Transformation\left(Y_{predicted}\right)$ involves transforming the predicted segmentation mask into the appropriate real-world coordinates.

The integration of augmented reality with semantic segmentation for medical data visualization. It outlines the process of learning a mapping using a neural network to predict segmentation masks from medical images. The training of the model involves minimizing a loss function using optimization algorithms like stochastic gradient descent. The integration of semantic segmentation with augmented reality involves overlaying the predicted segmentation mask onto the real-world view using geometric calibration and registration techniques. The above comparison Tables 1-3 presents the accuracy (mIOU) along with frame rate results for the proposed method with the methods from the literature tested on the images of Cityscapes and PASCAL VOC 2012 database.

Table 1. Accuracy (mIOU) for the proposed method with the methods from the literature tested on the images of Cityscapes database

Semantic Segmentation Methods	mIOU
PSPNet	81.2
ResNet-38	80.6
DeepLabv3	81.3
DeepLabv3+	82.1
Mapillary	82.0
Proposed Method	83.7

Cityscapes is a large scale database widely used for image segmentation. It consists of 5,000 fine annotated images and 20,000 coarse annotated ones. It provides annotations for 30 classes grouped into 8 categories.

PACAL VOC(Visual Object Classes) is a well-known dataset which is commonly used for object detection, semantic segmentation and classification. It is XML file created for each image. It contains three subsets of datasets such as 1,464 training images, 1,449 validation images and a private test set. The proposed method was tested on both Cityscapes and PACALVOC dataset.

Table 2. Frame rate result for the proposed method with the Cityscapes database images

Database	Images	Time(s)	Frame Rate (fps)
Cityscapes	1525	25	61

Table 3. Accuracy (mIOU) for the proposed method with the methods from the literature tested on the images of PASCAL VOC 2012 database

Method	mIOU
Large_Kernel_Matters	83.6
Multipath-RefineNet	84.2
Deep Layer Casecade (LC)	82.7
TuSimple	83.1
ResNet-38_MS_COCO	84.9
PSPNet	85.4
IDW_CNN	86.3
DeepLabv3-JET	86.9
DeepLabv3+(Xception)	87.8
DeepLabv3+(Xception-JET)	89.0
DIS	86.8
$CASIA_IVA_SDA$	86.6
Proposed Method	89.5

Table 4. Confirmed and Testing values for different parameters and their difference

	Dice	Loss	Mean_iou	Sensitivity	Precision	Specify	Recall
Training	0.8769	-0.8769	0.7850	0.8716	0.8856	0.9453	0.8716
Testing	0.9207	-0.9207	0.8570	0.8900	0.9583	0.9537	0.8900
Difference	0.0435	-0.0435	0.070	0.0145	0.0678	0.0334	0.0145

6. Experimental results

Figures 2-4 shows the comparison of training Vs. validation in terms of loss, accuracy, mean_iou, specificity, recall and precision respectively.

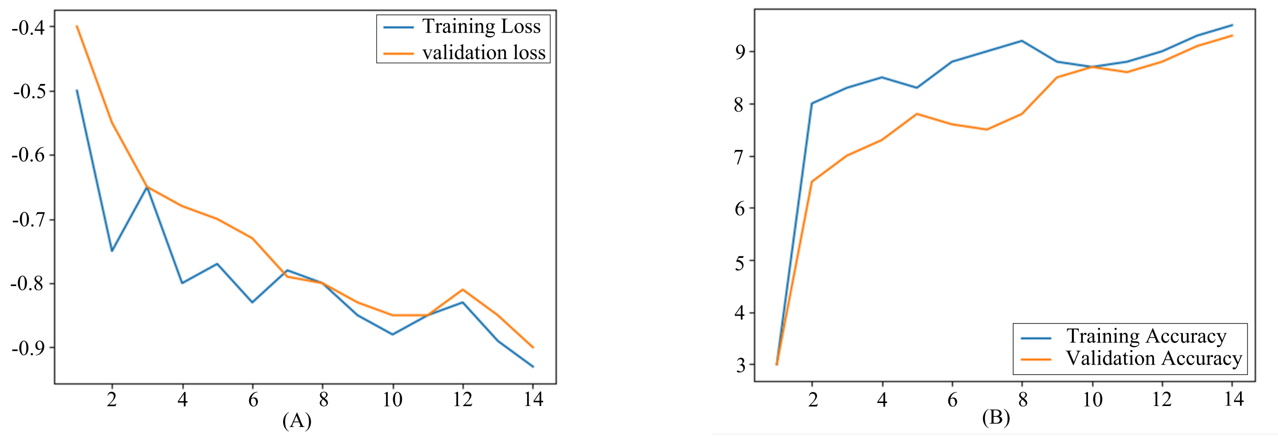

Figure 2. Training Vs. Validation loss (A), Training Vs. Validation accuracy (B)

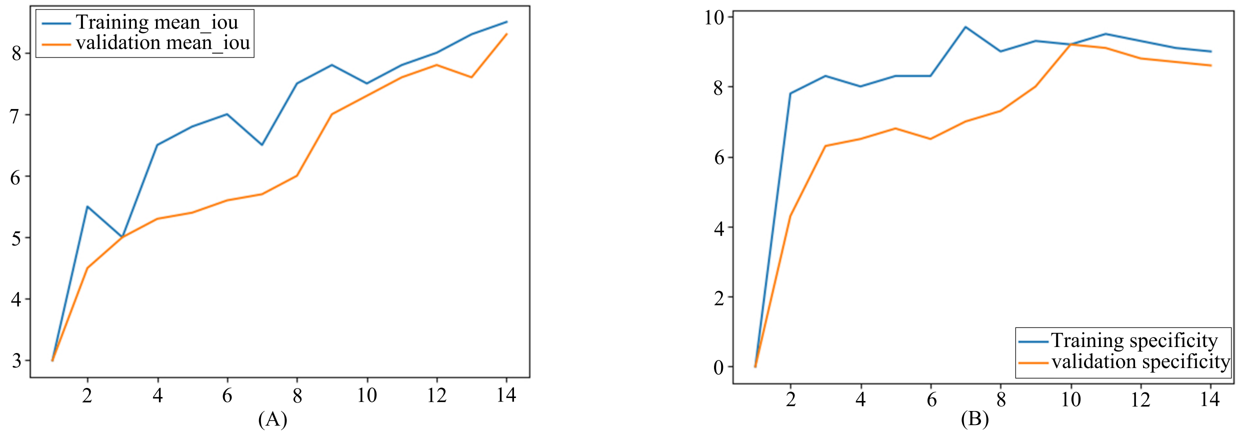

Figure 3. Training Vs. Validation mean_iou (A), Training Vs. Validation specificity (B)

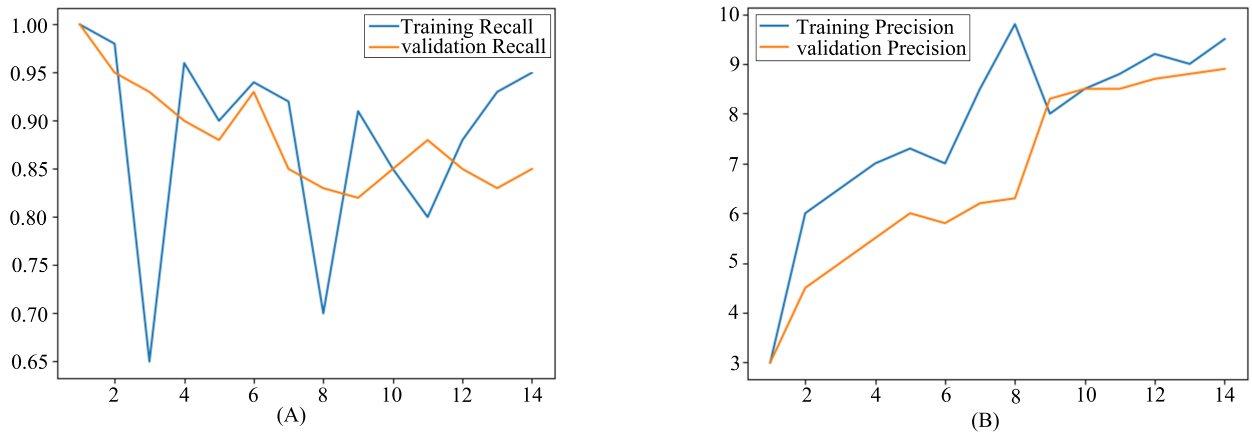

Figure 4. Training Vs. Validation Recall (A), Training Vs. Validation precision (B)

The Figure 5 proves that the proposed semantic segmentation method had a higher accuracy than its counterparts which was tested on Cityscapes database.

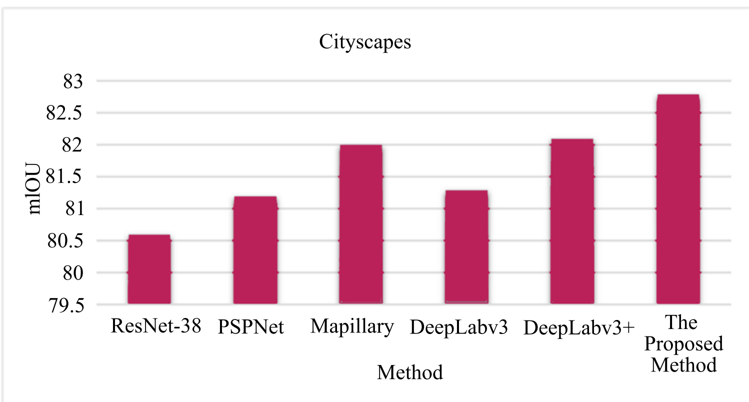

Figure 5. Accuracy (mIOU) of the proposed method Vs. its counterparts on images in the Cityscapes database

The Figure 6 proves that the proposed semantic segmentation method had a higher accuracy than its counterparts which was tested on PASCAL VOC 2012 database.

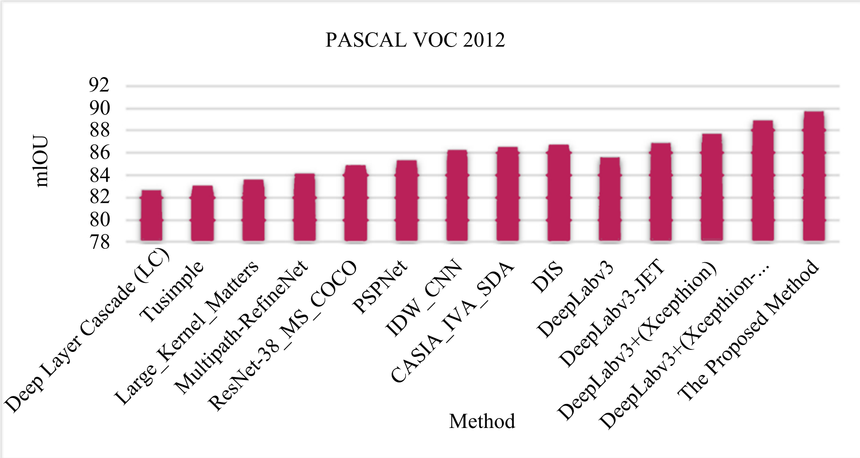

Figure 6. Accuracy (mIOU) of the proposed method Vs. its counterparts on images in the the PASCAL VOC 2012 database

7. Conclusion

Let's sum up by saying that the application of artificial intelligence (AI) and augmented reality (AR) to semantic segmentation in healthcare is a revolutionary step that could change treatment planning, diagnosis, and imaging. Along with improving the precision of the delineation of anatomical structures, this convergence of state-of-the-art technologies also presents new methods for training, visualization, and patient involvement. Through accurate and automated identification of anatomical structures and abnormalities across multiple imaging modalities, AI-driven semantic segmentation enables medical professionals to provide care. Early detection and individualized treatment plans may be possible as a result of this substantial improvement in diagnostic accuracy and efficiency.

During surgical procedures, real-time visualization of segmented medical data is made possible by the integration of AR into workflows in the healthcare industry. Augmented overlays have the potential to improve patient outcomes by helping surgeons become more precise and spatially aware during complex surgeries. A dynamic platform for medical education is offered by the combination of AR technologies and semantic segmentation. In order to comprehend complex anatomy and diseases better, trainees can interact with interactive 3D models. The training process could be speed up and healthcare professionals' skills could be enhanced by this immersive learning opportunity.

Patients can gain a better understanding of their conditions through AR interfaces that present segmented medical data in an understandable manner. This promotes better communication between patients and healthcare professionals, facilitating well-informed choices and raising patient satisfaction levels all around. Strong security and privacy measures are necessary due to the sensitive nature of medical data. Respecting legal frameworks like HIPAA is essential to protecting patient data.

Interdisciplinary collaboration among AI developers, healthcare professionals, and AR specialists is essential for addressing complex healthcare challenges. To achieve this, the document outlines several key strategies, including establishing common goals and objectives, creating cross-disciplinary teams, promoting open communication, providing interdisciplinary training, facilitating knowledge sharing, encouraging interdisciplinary research, providing resources and infrastructure, promoting leadership and collaboration skills, fostering a culture of innovation, and celebrating success and recognizing contributions. These strategies aim to foster a collaborative environment where unique expertise and perspectives can come together to drive innovation and problem-solving in healthcare.

Healthcare professionals, AR specialists, and AI experts must work together seamlessly for the implementation to be successful. It should take a team effort to develop and implement these technologies in order to successfully meet clinical needs. The current healthcare systems and electronic health records (EHRs) need to be seamlessly integrated with AI and AR solutions. Widespread adoption depends on compatibility and interoperability.

This research experiments on only two datasets (PASCAL VOC and Cityscapes) and considered only two AR frameworks for App development(ARKit and ARCore). It can be extended for experimenting in other datasets also (COCO and Open Images). Other App development (Kudan, Vuforia and Unreal Engine) support also can be extended. In order to provide more precise and patient-specific diagnosis and treatment planning, future research can investigate the applications of personalized medicine by customizing semantic segmentation models to specific patient characteristics. Continuous research and development are necessary to enhance segmentation accuracy, optimize algorithms, and investigate novel applications in the healthcare field as technology advances. Remote diagnostics and consultations can be made possible by the addition of AI-driven semantic segmentation and augmented reality to telemedicine platforms.

Conflict of interest

There is no conflict of interest in this study.

References

[1] Amara K. Augmented reality for medical practice: A comparative study of deep learning models for Ct-scan segmentation. *In 2022 International Conference ofAdvanced Technology in Electronic and Electrical Engineering (ICATEEE)*, Msila, Algeria; 2022. Available from: http://dx.doi.org/10.1109/ICATEEE57445.2022.10093691.

[2] Gite S, Mishra A, Kotecha K. Enhanced lung image segmentation using deep learning. *Neural Computing for IOT Based Intelligent Healthcare Systems*. 2023; 35: 22839-22853. Available from: http://dx.doi.org/10.1007/s00521-021-06719-8.

[3] Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, et al. Artificial intelligence for intraoperative guidance: Using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy: Using semantic segmentation to identify surgical anatomy during laparoscopic cholecystectomy. *Annals of Surgery*. 2022; 276(2): 363-369. Available from: http://dx.doi.org/10.1097/SLA.0000000000004594.

[4] Anitha T, Aanjankumar S, Poonkuntran S, Nayyar A. A novel methodology for malicious traffic detection in smart devices using BI-LSTM-CNN-dependent deep learning methodology. *Neural Computing and Applications*. 2023; 35(27): 20319-20338. Available from: http://dx.doi.org/10.1007/s00521-023-08818-0.

[5] Meghavi R, Bhushan M. Advancements in healthcare services using deep learning techniques. *In 2022 International Mobile and Embedded Technology Conference (MECON)*, Noida, India; 2022. Available from: http://dx.doi.org/10.1109/MECON53876.2022.9752020.

[6] Sathiyamoorthi V, Ilavarasi AK, Murugeswari K, Syed Thouheed Ahmed B, Devi M. A deep convolutional neural network based computer aided diagnosis system for the prediction of Alzheimer's disease in MRI images. *Measurement*. 2021; 171: 108838. Available from: http://dx/doi.org/10.1016/j.measurement.2020.108838.

[7] Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011-2022). *Computer Methods and Programs in Biomedicine*. 2022; 226: 107161. Available from: http://dx.doi.org/10.1016/j.cmpb.2022.107161.

[8] Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*. 2022; 15(1): 210. Available from: http://dx.doi.org/10.1186/s13104-022-06096-y.

[9] Arabahmadi M, Farahbakhsh R, Rezazadeh J. Deep learning for smart healthcare-A survey on brain tumor detection from medical imaging. *Sensors*. 2022; 22(5): 960. Available from: http://dx.doi.org/10.3390/s22051960.

[10] Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y, et al. MONAI: An open-source framework for deep learning in healthcare. *arXiv:2211.02701*. 2022. Available from: http://arxiv.org/abs/2211.02701.

[11] Samee NA, Ahmad T, Mahmoud NF, Atteia G, Abdallah HA, Rizwan A. Clinical decision support framework for segmentation and classification of brain tumor MRIs using a U-Net and DCNN cascaded learning algorithm. *Healthcare*. 2022; 10(12): 2340. Available from: http://dx.doi.org/10.3390/healthcare10122340.

[12] Tedeschini B. Decentralized federated learning for healthcare networks: A case study on tumor segmentation. *IEEE Access*. 2022; 10: 8693-708. Available from: http://dx.doi.org/10.1109/ACCESS.2022.3141913.

[13] Kaviani S, Han KJ, Sohn I. Adversarial attacks and defenses on AI in medical imaging informatics: A survey. *Expert Systems with Applications*. 2022; 198(116815): 116815. Available from: http://dx.doi.org/10.1016/j.eswa.2022.116815.

[14] Neelakandan S, Beulah JR, Prathiba L, Murthy GLN, Irudaya Raj EF, Arulkumar N. Blockchain with deep learning-enabled secure healthcare data transmission and diagnostic model. *International Journal of Modeling, Simulation, and Scientific Computing*. 2022; 13(04): 2241006. Available from: http://dx.doi.org/10.1142/s1793962322410069.

[15] Abdou MA. Literature review: Efficient deep neural networks techniques for medical image analysis. *Neural Computing and Applications*. 2022; 34(8): 5791-812. Available from: http://dx.doi.org/10.1007/s00521-022-06960-9.

[16] Cui Z, Fang Y, Mei L, Zhang B, Yu B, Liu J, et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nature Communications*. 2022; 13(1): 2096. Available from: http://dx.doi.org/10.1038/s41467-022-29637-2.

[17] Amara K. Augmented reality for medical practice: A comparative study of deep learning models for Ct-scan segmentation. *In 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE)*, Msila, Algeria; 2022. Available from: http://dx.doi.org/ 10.1109/ICATEEE57445.2022.10093691.

[18] Microsoft Corporate Blogs. *Breaking new ground in healthcare with the next evolution of AI*. Redmond , WA, USA: The Official Microsoft Blog; 2024. Available from: https://blogs.microsoft.com/blog/2023/03/20/breaking-new-ground-in-healthcare-with-the-next-evolution-of-ai/ [Accessed 6th March 2024].

[19] MedTech Pulse. *The newest surgical assistants in operating rooms? AR glasses*. Paris, French: MedTech Pulse; 2024. Available from: https://www.medtechpulse.com/article/insight/the-newest-surgical-assistants-in-operating-rooms [Accessed 6th March 2024].

[20] Htun SNN, Zin TT, Tin P. Image processing technique and hidden Markov model for an elderly care monitoring system. *Journal of Imaging*. 2020; 6(6): 49. Available from: https://www.mdpi.com/2313-433X/6/6/49.

[21] Thi Zin T, Htet Y, Akagi Y, Tamura H, Kondo K, Araki S, et al. Real-time action recognition system for elderly people using stereo depth camera. *Sensors*. 2021; 21(17): 5895. Available from: https://www.mdpi.com/1424-8220/21/17/5895.