Augmented Reality-Assisted Pediatric Surgery: A Machine Learning-Based Approach for Enhancing Accuracy

Sumedha Bhardwaj
Department of School of Computing
DIT University,
Dehradun, Uttarakhand, India.
sumedhabhardwaj719@gmail.com

Swapna Narla
Tek Yantra Inc,
California, USA.
swapnanarla@ieee.org

Md Nurul Absur
Computer Science, CUNY Graduate Center,
New York, New York, USA.
mabsur@gradcenter.cuny.edu

Shuvo Dutta
Department of Physics.
Western Michigan University,
USA.
shuvo.dutta@wmich.edu

Hyzil J B
Department of Zoology, Govt College for Women,
Thiruvananthapuram.
Affiliated to University of Kerala.
hyzilsyam@gmail.com

Krishna Ganesh S M
Department of Artificial Intelligence and Data Science
New Prince Shri Bhavani College of Engineering and Technology
Chennai, India.
krishnaganeshsm@gmail.com

Abstract: **Numerous advancements have been identified as potentially disruptive innovations in the surgical workplace as surgical workflow continues to improve in the digital age. The application of virtual reality (VR) and augmented reality (AR) within health care to enhance the use of data for medical reasons is unavoidable given their rapid expansion in availability, affordability, and accessibility. Using augmented reality may help paediatric orthopaedic and trauma surgeons overcome the difficult problem of comprehending the spatial linkages among the implants, anatomy, and their instruments. As needs be, the current examination proposes another AR of the surgical scene that deftly incorporates the various information sources given by a mobile C-arm and a Kinect RGB-depth sensor. This paper present a learning-based worldview to perceive the significant items or life structures in both Kinect and data of X-ray, and to produce an object unambiguous pixel-wise alpha guide that permits importance based combination of the X-ray pictures and the video inside a solitary view. We demonstrate very encouraging results in 10 simulated procedures with the goal of giving physicians a better grasp of the surgical environment and an enhanced perception.**

Keywords: Augmented reality, Accuracy, Pediatric surgeries, Kinect .

I. Introduction

In order to enhance their working circumstances, surgeons are always looking for novel methods. They frequently take the lead in implementing new technology that improves the surgery and patient experience in their profession. Examples of these developments included fiber-optics, which enabled limited access surgery, and robotic surgery, which paved the way for the creation of devices such as the da Vinci robot over a decade ago [1]. These tools used to be very expensive, but in the past ten years, they have grown more affordable and accessible. The quick development and shrinking of real-time visualisation platforms is giving rise to new computational paradigms. With micro processing capabilities comparable to desktop computers, smartphones are increasingly widely used. The nearly universal usage of smartphones by doctors is contributing to an increase in the application of electronic devices for medical purposes [2]. Clinical practice is now guided by immediate access to web-based information and applications for medicine [3]. Future apps to improve the surgical experience should take into account the recently developed role of VR and AR in healthcare, which is ready to be translated into this data-rich industry.

The move from open operations to minimally invasive laparoscopic surgeries, which only give the surgeon visual information through the laparoscope camera and do not permit direct organ touch, is one movement in surgery.

Minimally invasive surgery, or MIS, has become more popular and significant since these methods allow surgeons to reach the patient's interior anatomy in less intrusive ways and allow for the completion of whole operations with the least amount of stress to the patient.

Driven by the advantages that MIS may provide patients, numerous research teams are currently concentrating on creating systems that will help surgeons perform surgeries and complete their work more quickly and safely. Other examination bunches have made answers for help preoperative surgical preparation as well as the intraoperative careful activity. Despite the fact that deciphering computed tomography (CT) or magnetic resonance imaging (MRI) is as yet troublesome, the most recent improvements in clinical imaging handling permit the reproduction of 3D portrayals of the organs, giving physiological information that is not really perceivable by CT and X-ray cuts or ultrasound filters, as well as an exhaustive comprehension of the patient's pathologies and life structures.

A good use of such simulators could enhance patient care by directing the devices across the body without the doctor's direct sight. These models can also be used as the basis for developing lifelike virtual worlds for VR and AR technologies.

Surgeons are constantly searching for new technology to increase productivity. The use of VR and AR in medicine, particularly in surgery, has grown [4–6]. Users can fully immerse themselves in a simulated world thanks to VR's creation of an artificial three-dimensional simulated environment, which is one of the two types of reality [7, 8].VR cannot be employed in actual procedures since it lacks reality. Immersion in the scenario is challenging for clinicians [9, 10]. AR and VR are not the same thing. AR is commonly defined as the immediate interaction of extra information or visual components with the user's surroundings [11]. Instead of operating in a completely artificial setting, AR carries out the task's interaction focus in the real world [12].However, the majority of AR solutions rely on heavy devices and intricate external navigation systems, which restricts their application in standard surgical procedures [13].

In order to provide a better augmented reality representation, the current study recommends utilising more depth information: It (I) characterizes a learning-based procedure based with respect to variety and profundity data to recognize objects of interest in Kinect information; (ii) distinguishes foreground objects in X-beam pictures utilizing a cutting edge random forest; and (iii) utilizes an article explicit stirring look-into table to make a pixel-wise alpha guide.

II. Literature Review

Generally speaking, pre-operative planning, training, and advanced visualisation throughout the actual treatment can all benefit from the use of AR in MIS.

A number of image-guided surgery technologies have been created, and multiple groups of researchers are investigating the application of AR in surgery.

For robotically assisted minimally invasive heart surgery, studies suggest using an endoscopic AR system [14].

Others offer customised solutions for a variety of therapeutic applications by utilising techniques based on cutting-edge ideas in robotics, haptics, and visualisation [15].

Technologies like VR and AR have drawn a lot of attention from a variety of sectors, including entertainment, industry, education, and medicine. An examination of earlier research indicates that the field of study on the applications, challenges, and potential advancements of VR and AR in surgical training and medical education is growing quickly. An in-depth analysis of 3D printing and anatomical engineering for surgery and medical equipment is provided by [16]. The study emphasises how 3D printing technology may be used to create anatomical models tailored to each patient for preoperative planning and surgical training.

A setting "where real world and virtual world objects are presented together within a single display, that is, anywhere between the extrema of the virtuality continuum" is what [17] characterised as mixed reality.

The virtuality continuum, which extends from the fully real to the entirely virtual world, is where AR and enhanced virtuality are situated.

With real-time interaction, virtual reality creates an immersive, entirely synthetic computer-simulated visual and environment. For over ten years, virtual reality has been utilised for endoscopic instruction and evaluation [18]. One of the earliest endoscopic education platforms was Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR). A recent meta-analysis of randomised controlled studies shown that using VR training for novice trainees or in conjunction with traditional laparoscopic training can reduce operational time, mistake rate, and accuracy [19].

To properly overlay synthetic data over real-world images, researchers consider both the visual complexity of the computer-generated augmentations applied to the view and the information that would be hidden by the augmentations [20].

Studies provide an system of AR that can guide the physician throughout the surgical phase to prevent unintentionally disturbing specific organs throughout surgery [21].

[22] Introduce a novel method of medical in-situ visualisation that improves their capacity to comprehend 3D data of medical imaging and move instruments of surgery in relation to the patient's anatomy, thereby addressing the problem of inaccurate perceptions of depth and spatial organisation in medical AR. A head-mounted video see-through display can be used to alter the transparency of video images, as they explain.

III. Methodology

A. System configuration: mobile C-arm enhanced by Kinect

In this study, it suggest adding a sensor of Kinect which contains of a depth sensor connected to a camera video —to a standard intraoperative mobile C-arm. This RGB-D sensor's video camera optical centre is positioned to align with the projection centre of X-ray. The depth sensor works by projecting infrared light patterns into the image, a technique known as "structured light." The depth is deduced from the distortions of those patterns caused by the scene's three-dimensional structure using an infrared camera. In order to be processed the images on depth into the video camera coordinates, the sensor contains an integrated calibration. This system is made up of a mirror system, a Kinect sensor, and an aluminium frame that realistically resembles a C-arm. Here, we use a single mirror to mimic the requirement that the X-ray source and the optical video centre of the C arm camera-augmented mobile system essentially coincide.

B. Suggested method

First, let's look at the colour I and depth D images that the Kinect sensor provided, both of which were registered using the built-in calibration. Finding pertinent items in the surgical scene—that is, items that belong to the front and certain classes of interest, like the surgeon's hands or surgical instruments—is the aim of this work. In this context, it suggests dividing the work of recognising relevant things into two smaller tasks: (1) using the depth image's content to detect candidate foreground objects, and (2) using the colour image's content to identify relevant objects.

C. Using depth photos for background modelling

To achieve foundation deduction in variety photographs for following applications, foundation demonstrating has been broadly explored. The main concept with a fixed-camera system is to use a collection of backdrop photos to learn a colour distribution for every pixel. Several techniques for adaptive real-time subtraction background have been proposed in the last decade, according to [23]. These techniques include kernel density estimation, mixture simulations, Gaussian average, and so-called foundation of Eigen. In this work, we suggest learning an established framework for the depth background by using a set of frames at the surgery starting.

D. Identifying objects with colour pictures

Lesion identification, segmentation, and anatomy localisation are only a few of the many uses for random forests in medical image analysis, according to [24]. As an assortment of decision trees, they give piecewise portrayals of any dissemination in high-layered spaces.

E. Performance Measure

The correctness of the model influences both its training performance and its real-world behaviour. It will not, however, specify how it would be applied to the problem. Accuracy merely tells us how well the trained model works.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (1)$$

When assessing if there are more false positives than real positives, precision is essential..

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

Recall is useful when the number of false negatives is higher. The efficacy of the approach we use decreases when false negatives happen more frequently than.

$$Recall = \frac{TP}{TP + FN} \qquad (3)$$

The weighted mean of recall and precision is known as the F1-score.

$$F1 - score = \frac{2. precision. recall}{precision + recall} \qquad (4)$$

IV. Empirical Results

In this study, we use our proof-of-concept system to show the potential of the proposed approach. Employing phantom surgical and actual images of X-ray taken from various orthopaedic procedures, it simulates ten distinct paediatric orthopaedic surgeries. Please be aware that the images of X-ray are carefully orientated within the view of the proposed surgical scene before we start the acquisitions. Different tasks utilising various surgical instruments, such as a scalpel, drill, and hammer, are carried out in each sequence. It obtain two sequences with varying light circumstances, while the simulations in ten sequences are conducted in comparable light settings.

It annotated each of the ten sequences, four video and depth frames, and twenty X-ray images for this experiment. It used two-fold cross-validation to assess our methods in both data of X-ray and RGB-D. Please take note that the two sequences with varying light conditions were not used to train our RGB-D identification classifier. Each CIElab channel retrieves fifty context characteristics that show the context of visual of every pixel in the image of colour. To tackle the object recognition task, it utilise 25 trees of depth 20 to train a random for-est classifier.

Before the surgery starts, the first 30 frames of each sequence are used to create the depth backdrop model. The classifier consists of a random forest of 25 trees with depth 20 and extracts 50 context characteristics to characterise pixel context for item identification in X-ray images.

The overall identification-fusion speed for our proof-of-concept device is 2.7 frames per second. Figs. 1, 2, and 3 present a quantitative evaluation of our item recognition algorithm in both Kinect and X-ray data. The job of detecting products of interest towards the background was evaluated in terms of recall, F-score and precision.

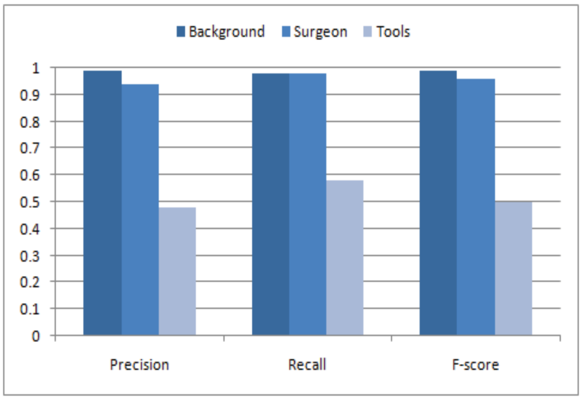

Fig 1. Using a broad tool, the objects' quantitative outcomes for standard sequences

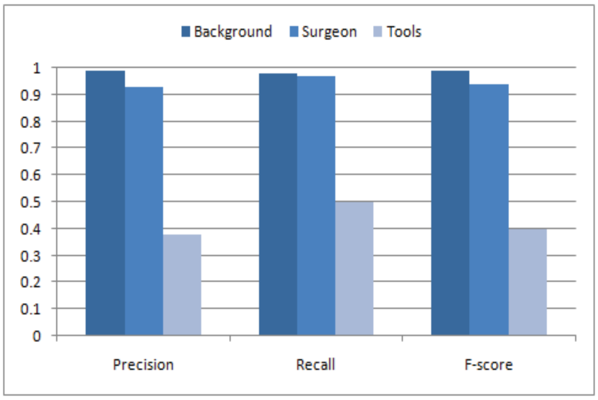

Fig 2. Categorisation outcomes using thin metallic tools on typical sequences

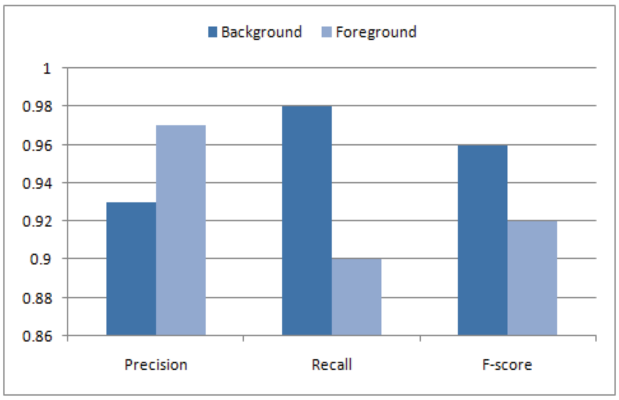

Fig 3. Quantitative findings from the identification of objects in X-ray data

Tables I and II show the all included matrix of confusion for the various objects of interest compared to the backdrop. The annotated frames for all 50 sequences, 42 photos of 10 standard sequences with comparable lighting conditions, and the final 4 sequences with perturbations, such as dynamic changes in lighting conditions, are used to calculate these metrics. The surgeon's segmentation works well, but the tools' segmentation is more difficult, according to the results. Indeed, a variety of instruments with varying shapes, sizes, colours, and material properties—such as drills, clamps, or scalpels—were employed in our studies.

TABLE I. CONFUSION MATRIX FOR IDENTIFYING OBJECTS IN ALL SEQUENCES OF KINECT RGB-D DATA

	Predicted background	Predicted Surgeon	Predicted tool
Actual background	99.72	.52	.85
Actual Surgeon	6.76	96.15	.28
Actual tool	19.04	.29	85.26

TABLE II. CONFUSION MATRIX FOR IDENTIFYING OBJECTS IN X-RAY DATA

	Predicted background	Predicted foreground
Actual background	93.28	6.74
Actual foreground	3.76	97.48

Ten different sequences were acquired from five clinical participants, including few last-year medical students and three experienced surgeons. A 5-point Likert scale, is used to evaluate their questionnaire responses. The majority of participants (4.7 ± 0.5) agreed that our approach resolves the depth ordering. The response is neutral (3.3±1.4) and somewhat positive (3.6±1.1) with respect to the visibility of the instrument tip or implant in X-rays. The overall impression of the visualisation has changed, which makes it more appropriate for potential inclusion into surgical workflow, according to their unanimous agreement (4.2 ± 1.4). In the end, every participant strongly agreed (4.6 ± 0.9) that they would rather use our innovative visualisation than traditional alpha blending.

VI. CONCLUSION

In this study, we proposed new learning methodologies and strategies for AR visualisation to enhance depth perception and comprehension of paediatric surgical scenes. The primary findings included the identification of learning-based techniques for identifying objects of interest in Kinect and data of X-ray, the suggestion of a mixed Kinect and C-arm sensor for obtaining depth and colour details, and the creation of an object-specific pixel-wise alpha map for enhanced image fusion. It provides encouraging results for enhanced depth perception and surgical scene comprehension in 10 simulated surgeries. Additionally, our innovative Kinect-enhanced C-arm system paves the way for a number of fascinating future projects, including workflow analysis, tool tracking, position estimation, and navigation.

REFERENCES

[1] N. E. Samalavicius et al., "Robotic surgery using Senhance® robotic platform: single center experience with first 100 cases," *J. Robotic Surgery*, vol. 14, pp. 371-376, 2020.

[2] N. V. Shah, R. Gold, Q. A. Dar, B. G. Diebo, C. B. Paulino, and Q. Naziri, "Smart technology and orthopaedic surgery: current concepts regarding the impact of smartphones and wearable technology on our patients and practice," *Curr. Rev. Musculoskeletal Med.*, pp. 1-14, 2021.

[3] J. N. Grow, J. D. Vargo, N. Nazir, and R. Korentager, "Smartphone applications in plastic surgery: a cross-sectional survey of 577 plastic

surgeons, fellows, residents, and medical students," *Aesthetic Surgery Journal*, vol. 39, no. 12, pp. NP530-NP537, 2019.

[4] Y. Pulijala, M. Ma, M. Pears, D. Peebles, and A. Ayoub, "An innovative virtual reality training tool for orthognathic surgery," *Int. J. Oral Maxillofac. Surg.*, vol. 47, no. 9, pp. 1199-1205, 2018.

[5] M. Zahiri, C. A. Nelson, D. Oleynikov, and K. C. Siu, "Evaluation of augmented reality feedback in surgical training environment," *Surgical Innovation*, vol. 25, no. 1, pp. 81-87, 2018.

[6] S. Sheik-Ali, H. Edgcombe, and C. Paton, "Next-generation virtual and augmented reality in surgical education: a narrative review," *Surgical Technol. Int.*, vol. 33, 2019.

[7] N. Ahmadpour et al., "Virtual Reality interventions for acute and chronic pain management," *Int. J. Biochem. Cell Biol.*, vol. 114, p. 105568, 2019.

[8] R. Eijlers et al., "Systematic review and meta-analysis of virtual reality in pediatrics: effects on pain and anxiety," *Anesthesia Analgesia*, vol. 129, no. 5, pp. 1344-1353, 2019.

[9] M. R. Paro, D. S. Hersh, and K. R. Bulsara, "History of virtual reality and augmented reality in neurosurgical training," *World Neurosurg.*, vol. 167, pp. 37-43, 2022.

[10] C. Fertleman et al., "A discussion of virtual reality as a new tool for training healthcare professionals," *Front. Public Health*, vol. 6, p. 44, 2018.

[11] B. Z. Hameed et al., "Application of virtual reality, augmented reality, and mixed reality in endourology and urolithiasis: an update by YAU Endourology and Urolithiasis Working Group," *Front. Surg.*, vol. 9, p. 866946, 2022.

[12] J. G. Kovoor, A. K. Gupta, and M. A. Gladman, "Validity and effectiveness of augmented reality in surgical education: a systematic review," *Surgery*, vol. 170, no. 1, pp. 88-98, 2021.

[13] S. C. Lee et al., "Multi-modal imaging, model-based tracking, and mixed reality visualisation for orthopaedic surgery," *Healthcare Technol. Lett.*, vol. 4, no. 5, pp. 168-173, 2017.

[14] S. Eom, S. Kim, S. Rahimpour, and M. Gorlatova, "AR-assisted surgical guidance system for ventriculostomy," in *2022 IEEE Conf. Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2022, pp. 402-405.

[15] P. E. Edwards, M. Chand, M. Birlo, and D. Stoyanov, "The challenge of augmented reality in surgery," in *Digital Surgery*, pp. 121-135, 2021.

[16] J. Cornejo et al., "Anatomical Engineering and 3D printing for surgery and medical devices: International review and future exponential innovations," *BioMed Res. Int.*, vol. 2022, p. 6797745, 2022.

[17] J. S. Roo and M. Hachet, "One reality: Augmenting how the physical world is experienced by combining multiple mixed reality modalities," in *Proc. 30th Annual ACM Symp. User Interface Softw. Technol.*, 2017, pp. 787-795.

[18] O. H. Hasan et al., "The need for simulation in surgical education in developing countries. The wind of change. Review article," *J. Pak. Med. Assoc.*, vol. 69, no. Supl. 1, p. S62, 2019.

[19] C. Jin, L. Dai, and T. Wang, "The application of virtual reality in the training of laparoscopic surgery: A systematic review and meta-analysis," *Int. J. Surg.*, vol. 87, p. 105859, 2021.

[20] D. Ren, T. Goldschwendt, Y. Chang, and T. Höllerer, "Evaluating wide-field-of-view augmented reality with mixed reality simulation," in *2016 IEEE Virtual Reality (VR)*, 2016, pp. 93-102.

[21] T. Huber et al., "Highly immersive virtual reality laparoscopy simulation: development and future aspects," *Int. J. Comput. Assisted Radiol. Surg.*, vol. 13, pp. 281-290, 2018.

[22] C. Gsaxner et al., "Inside-out instrument tracking for surgical navigation in augmented reality," in *Proc. 27th ACM Symp. Virtual Reality Softw. Technol.*, 2021, pp. 1-11.

[23] Goyal, K., & Singhai, J. (2018). Review of background subtraction methods using Gaussian mixture model for video surveillance systems. Artificial Intelligence Review, 50, 241-259.

[24] Zhao, X., Wu, Y., Lee, D. L., & Cui, W. (2018). iforest: Interpreting random forests via visual analytics. IEEE transactions on visualization and computer graphics, 25(1), 407-416.