



The Application of the Preoperative Image-Guided 3D Visualization Supported by Machine Learning to the Prediction of Organs Reconstruction During Pancreaticoduodenectomy via a Head-Mounted Displays

Klaudia Proniewska^{1,2}, Radek Kolecki^{3,4}, Anna Grochowska^{4,5}, Tadeusz Popiela^{4,5}, Tomasz Rogula⁴, Krzysztof Malinowski^{1,2}, Damian Dołęga-Dołęgowski⁶, Jakub Kenig^{4,5,6}, Piotr Richter^{4,5,7,8}, Julianna Dąbrowa⁴, MHD Jafar Mortada^{1,9}, Peter van Dam^{1,10}, and Agnieszka Pregowska¹¹✉

- ¹ Center for Digital Medicine and Robotics, Jagiellonian University Medical College, 7E Street, 31-034 Krakow, Poland
- ² Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, Medyczna 7 Street, 30-688 Krakow, Poland
- ³ Department of Orthopedic Surgery, University Hospital in Krakow, Jakubowskiego 2 Street, 30-688 Krakow, Poland
- ⁴ Jagiellonian University Medical College, Anny 12 Street, 31-008 Krakow, Poland
- ⁵ University Hospital in Krakow, Jakubowskiego 2 Street, 30-688 Krakow, Poland
- ⁶ Project LIDER XI NCBiR-HoloMed, Jagiellonian University Medical College, Anny 12 Street, 31-008 Krakow, Poland
- ⁷ Department of General, Oncologic and Geriatric Surgery, Jagiellonian University Medical College, Anny 12 Street, 31-008 Krakow, Poland
- ⁸ Chair and First Department of General Surgery, Jagiellonian University Medical College, Anny 12 Street, 31-008 Krakow, Poland
- ⁹ Department of Information Engineering, Università Politecnica delle Marche, 60121 Ancona, Italy
- ¹⁰ Department of Cardiology, University Medical Center Utrecht, Lundlaan 4, 3584 EA, Utrecht, The Netherlands
- ¹¹ Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawinskiego 5B, 02-106, Warsaw, Poland
apregow@ippt.pan.pl

Abstract. Early pancreatic cancer diagnosis and therapy drastically increase the chances of survival. Tumor visualization using CT scan images is an important part of these processes. In this paper, we apply Mixed Reality (MR) and Artificial Intelligence, in particular, Machine Learning (ML) to prepare image-guided 3D models of pancreatic cancer in a population of oncology patients. Object detection was based on the convolution neural network, i.e. the You Only Look Once (YOLO) version 7 algorithm, while the semantic segmentation has been done with the 3D-UNET algorithm. Next, the 3D holographic visualization of this model as an

interactive, MR object was performed using the Microsoft HoloLens2. The results indicated that the proposed MR and ML-based approach can precisely segment the pancreas along with suspected lesions, thus providing a reliable tool for diagnostics and surgical planning, especially when considering organ reconstruction during pancreaticoduodenectomy.

Keywords: Extended Reality · Mixed Reality · Augmented Reality · Head-Mounted Displays · Artificial Intelligence · Image-guided surgery

1 Introduction

Pancreatic cancer is most often diagnosed in a locally advanced stage, where local vessel involvement is more probable (Kenner et al. 2021), (Rawla et al. 2019). Since the early symptoms of the disease are mostly non-specific, an efficient diagnosis in the initial stages of the disorder is crucial in improving survival rates (Gheorghe et al. 2020). Experienced oncological surgical centers are capable of performing radical local resection with vessel reconstruction; however, this is often an intraoperative decision. Advances in 3D visualization of CT imagery provide a crucial preoperative decision-making tool to identify possible regions of vascular involvement, the feasibility of resection, and the extent of potential vessel reconstruction. This is particularly useful given that vascular invasion necessitating reconstruction is itself a negative prognosticator for long-term outcomes. Preoperative planning that shows the extent of invasion may avoid an operation, instead opting for alternatives such as neoadjuvant chemotherapy. Moreover, image-guided navigation fits in with the assumptions of an alternative to traditional solutions, minimally invasive surgeries (MIS) (Kochanski et al. 2019). Thus, early pancreatic diagnosis is of high importance.

Recently the development of head-mounted displays (HMDs) has opened many possibilities for the use of immersive technologies like Augmented Reality (AR), and Mixed Reality (MR) in various fields, particularly medicine with a special emphasis on the medical education sector (Klinker et al. 2020), (Garlinska et al. 2023), (Pregowska et al. 2022) and surgery (Wang et al. 2023). Since HMDs are the portable alternative to the standard computer monitors, they have the potential to revolutionize surgery by enabling augmenting the operating field through digital data visualization (Quero et al. 2019), (Acidi et al. 2023), (Brockmeyer et al. 2023). Overlaying an image, such as radiography (X-ray), magnetic resonance imaging (MRI), or computed tomography (CT) on anatomical landmarks allows the operator to visualize organs, vessels, and irregularities such as tumors (England et al. 2023) compared the perception of images by radiologists using HMDs or traditional monitor, finding that HMDs provided a higher quality and thus more actionable image. The application of HMDs in surgery may help to improve efficiency by providing the doctor with a more accurate preoperative image. (Patric & Baowei 2023) described the interface of the AR-based system for visualization of prostate biopsy with the color and transparency of each element of the 3D model being adjustable in real-time. (Pan et al. 2020) described AR-based navigation in deep anterior lamellar keratoplasty, finding that AR enables accurate detection and tracking of the corneal contour complex in real-time. Another interesting application of HMDs

is a medical ultrasound-guided 3D visualization (Ruger et al. 2020) (Nguyen et al. 2022). (von Haxthausen et al. 2023) presented an MR-based real-time volumetric in situ ultrasound for guiding vascular punctures. The holographic remote mode was used to send the rendered images to the HMDs. One of the main issues in the wide application of AR/MR in surgery is providing a sense of depth perception for the operator. The main limitation of the proposed approach is the difficulty in distinguishing relevant structures from noise and occlusion.

Machine Learning (ML) and particularly deep neural networks (DNNs) have revolutionized image segmentation (Chen et al. 2018), (Polyviou & Zamani 2023), (Hasan et al. 2023), (Chen et al 2023). Precise segmentation of lesions may contribute to efficient diagnostics and more effective, targeted therapy. For example, (Hosny et al. 2019) presented an algorithm for the segmentation of pigmented skin lesions, which enables diagnoses at an earlier stage without invasive medical procedures. Given its flexibility and scalability, ML can be an efficient tool for cancer diagnosis, in particular, the early disease stages (Young et al. 2020), (Granata et al. 2023). (Hayward et al. 2010) assessed multi-layer perceptron neural networks, Bayesian nets, and locally weighted naive Bayesian nets for clinical performance in a pancreatic cancer database. They found that ML algorithms provide a substantial improvement in predicting clinical outcomes. The most commonly used ML-based algorithms in image semantic segmentation, in particular in segmentation of different tumors, are Convolutional Neural Networks (CNNs) (Badrinarayanan et al. 2017), particularly the U-Net (Li et al. 2020), and its variation U-Net++ (Zhou et al. 2018), R2UNet (Alom et al. 2019). (Ren et al. 2023) examined the application of Attention U-Net to the prediction of solid pancreatic tumors. The proposed approach was based on feature fusion to assess the extent and boundaries of solid tumors. This approach was particularly helpful in generating early-stage images of tumors, thereby augmenting the training and testing sets. In (Janssen et al. 2021) U-nets with a DenseNet161 encoder were applied to the segmentation of residual tumors in pancreatic cancer histopathology samples. The results obtained show that AI-based algorithms can help the assessment of residual tumor burden. Another application of ML was shown by (Lee et al. 2021), i.e. the evaluation of the survival prediction of pancreas cancer based on Random Forest and the Cox proportional-hazards model.

In this paper, we also propose a so-the virtual shared experience paradigm, in which the primary HMD is worn by physicians with a secondary HMD worn by patients. Mixed Reality and Artificial Intelligence were applied to prepare 3D visualizations of preoperative surgical oncology patients' imaging.

2 Research Methodology

The study "Holographic MedAssistant" was conducted following the Declaration of Helsinki, and approved by the Medical Ethical Committee of the Jagiellonian University Medical College (JUMC) in Krakow, Poland No: 1072.6120.27.2020 and 1072.6120.92.2022.

2.1 Image Segmentation Based on Opened Database Source

Pancreatic segmentation from abdominal computer tomography (CT) images is a challenge because the morphology of the pancreas, i.e., its shape and position, has significant variability and may result in blurred boundaries (Dai et al. 2023). Also, the pancreas makes up a small amount of an entire abdominal CT scan. Moreover, the often the border between the pancreas and background is very subtle making it especially challenging to precisely define its boundaries. Segmentation was usually done manually by an expert moving a pointer. This was highly inefficient and often subject to error depending on the experience of the expert, especially when considering such challenging organs as the pancreas. Thus, accurate and automated segmentation will be important in the future of surgical planning. However, automated pancreas segmentation will also face the challenges described in segmenting a pancreas manually (Bagheri et al. 2020). In this paper, the data were obtained and delivered as de-identified DICOM files from open datasets. We have selected from a large collection (–100 3D CT images –) of healthy patients, namely 1) 281 patients with pancreatic tumors, (Simpson et al. 2019) and 2) 80 healthy patients (Roth et al. 2015), and the relevant anatomical structures were labeled using a series of Machine Learning-based segmentation algorithms by You Only Look Once (YOLO) version 7, which is open-source software to use under GNU General Public License v3.0 license (see, Fig. 1).

In this study, CT dicom files were processed in an open-source medical image analysis software, 3D Slicer (version 5.2.2, `functiongrow` from `seeds` function) to segment key structures. Anatomical landmarks including the pancreas, superior mesenteric vein (SMV), splenic vein (SV), portal vein (PV), abdominal aorta, and celiac trunk along with its branches were isolated and converted into a 3D mask using a combination of region-based segmentation to define organs and blood vessels as a volume rendering. An output surface matrix, or model, was obtained that identifies an object or instance to which each pixel belongs using input tags, or seeds. Object detection, i.e. localizing and identifying structures in a digital image or video was done first (Amit et al. 2020), followed by the bounding box, and ultimately object classification. Figure 2 shows an object detection example in a CT scan to show the liver and spine, i.e. it shows the bounding boxes around the detected regions. This is a different procedure than semantic segmentation, as in semantic segmentation the algorithm will classify each pixel (or voxel in 3D images) in a class (Thoma 2016). The semantic segmentation on a CT slice was visualized in Fig. 3 as an example.

The detection of pancreatic lesions consists of three stages. First, the region of interest (ROI) containing the pancreas, then the detection of a lesion in a pre-selected area of interest, and clinical validation consisting in comparing the results of the pancreatic lesion detection, including algorithms alone and combining the results of the algorithms with radiological verification to the standard approach, i.e. manual radiological assessment. The automated algorithm relies on a pipeline with two major steps, first is to detect the ROI i.e. the area around the pancreas, and then feed this detected ROI to another step that would perform the final segmentation. From a computer science point of view, the first step is called object detection and it relies on the YOLOv7 algorithm, while the second step is called semantic segmentation and relies on the 3D-UNET algorithm.

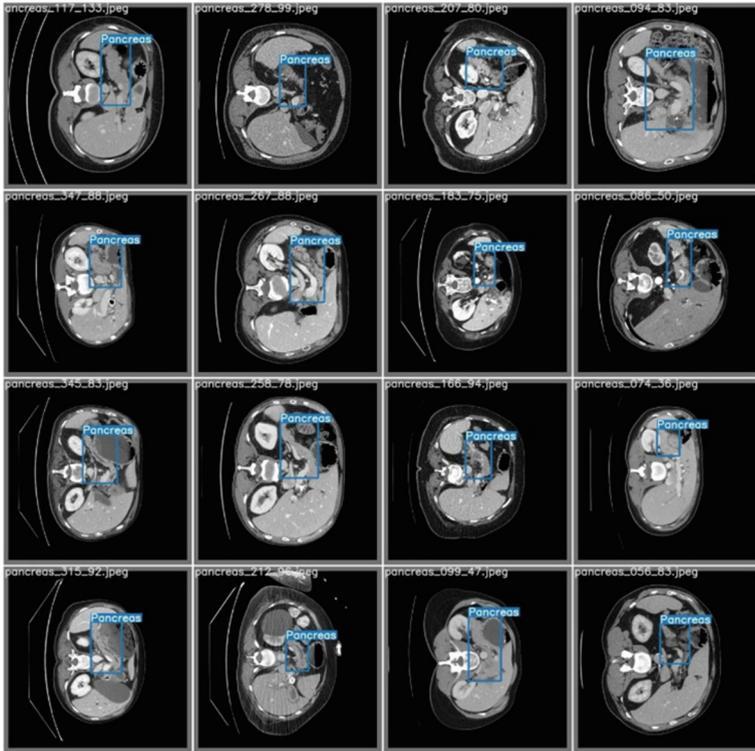


Fig. 1. Automatically generated region of interest by the YOLO training script



Fig. 2. Object detection in CT scan image. Red box around the pancreatic region. (Color figure online)



Fig. 3. Semantic segmentation in CT scan image. Note that the pancreas was marked by yellow color. (Color figure online)

The YOLO algorithm depends on deep learning for object detection (Redmon et al. 2016). It depends on the idea of the images passing only once through the neural network, hence the name, this is done by dividing the input image into a grid and predicting for each grid cell the bounding box and the probability of that class. The algorithm predicts different values about the object, i.e. the coordinates of the center of the bounding box around the object, the height and width of the bounding box, the class of the object, and the probability, or the confidence of the prediction. This way of working may cause the algorithm to detect the object multiple times, to avoid duplicate detections of the same object the algorithm uses non-maximum suppression (NMS), which works by calculating a metric called Intersection over Union or (IOU) between the boxes according to the following formula.

$$IOU = \frac{\text{Area of overlap}}{\text{Area of Union}}. \quad (1)$$

If the IOU between two boxes is larger than a threshold, the box with a higher confidence score is chosen and the other is ignored. There have been many improvements to the YOLO algorithm (Redmon & Farhadi 2017), (Bochkovskiy et al. 2020), (Zhu et al. 2021). In this study, YOLO version 7 (YOLOv7), which was released in 2022, is applied (Wang et al. 2022). It has several structural modifications that provide higher accuracy, faster performance, improved scalability, and greater flexibility for customization. YOLOv7 operates in pixel space, i.e., 2D space. Thus, there is a need to extract the three-dimensional ROI for the pancreas from different anatomical planes (Axial, Coronal, Sagittal) and combine this information to get the 3D ROI. Figure 4 shows the general idea of the YOLOv7 algorithm. The first row shows the pancreas detection in the three planes, combining this information would produce the coordinates in the voxel space.

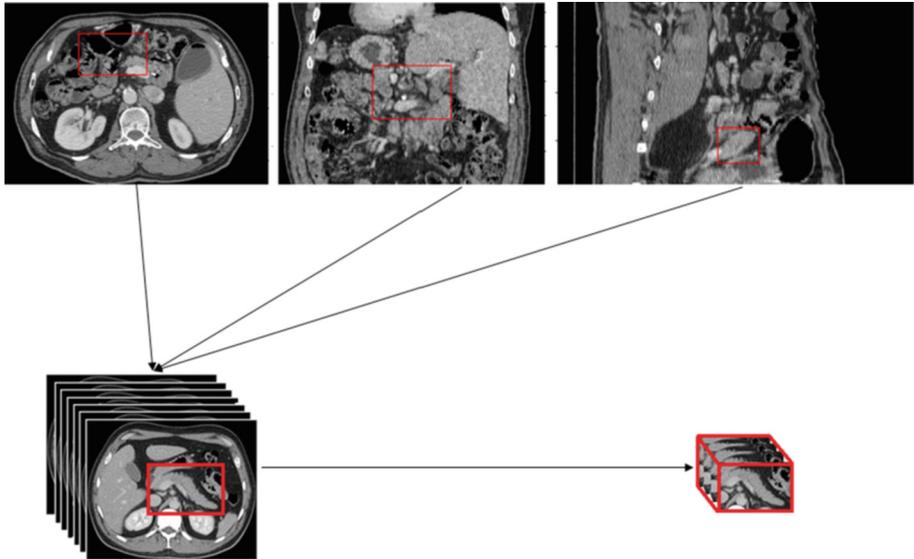


Fig. 4. The general functional structure of the YOLOv7 algorithm to detect the pancreas 3D ROI.

The next step, i.e., training the 3D UNET usually requires a substantial amount of computational power. However, since the image size will be significantly reduced, as shown in Fig. 4, the computational power will likewise be reduced allowing more freedom to optimize network parameters.

UNET is a convolutional neural network architecture, which was introduced by Ronneberger et al. (2015) (Ronneberger et al. 2015), it was built upon a well-established architecture called a fully connected network (FCN) (Shelhamer et al. 2014), to perform semantic segmentation on biomedical images, in the original paper, which was cited more than 62,000 times by the time this article was written, described the architecture as having two paths contracting path (also known as encoder) and an expansive path (also known as a decoder). The contracting path consists of 3×3 convolutional layers followed by 2×2 max-pooling layers with strides equal to two, after each max-pooling the number of features channels is doubled, while the expansive path is like a mirror for the contracting path with upsampling layers instead of max pooling, and a final 1×1 convolutional layer that should have a number of channels equal to the required number of classes, UNET also have skip connections between the two paths, which copies the images from steps of the contracting path to the mirrored stages of the expansive path.

The architecture showed promising results when applied to biomedical images especially when applying data augmentation with elastic deformation and trained with a weighted loss function with reasonable training time.

Although the original work considered 2D images only, the 3D UNET appeared later in 2016 to perform volumetric segmentation, using $3 \times 3 \times 3$ window convolutional layers followed by $2 \times 2 \times 2$ max-pooling layers with strides equal to two, This had many uses in segmenting CT images (Wu et al. 2020) and MRI images (Tomassini et al. 2023).

3D UNET would require more parameters which would apply limitations on the complexity of the model, also 3D data needs more memory space to handle, which might affect the size of the batches, we tend to deal with this problem by performing object detection step first to significantly reduce the size of the input data.

These steps were implemented using Python in the Google Colab Pro cloud service environment. NVIDIA T4 GPU was used for training of YOLO v7 algorithm whereas Tesla A100 was used for training of U-Net, using a loss function based on a combination of dice loss and Cross entropy. The object detection part (using the YOLO v7 algorithm) was evaluated for precision and recall. The semantic while the segmentation part (U-Net) was evaluated with the dice score, IOU as well as voxel-level ROC curve with AUC.

For YOLO a total of 100 epochs were used, with batch size 40, learning rate = 0.01, and momentum = 0.937, whereas U-Net was trained for 50 epochs, using ADAM optimizer, learning rate of 0.01, momentum = 0.99, beta_1 = 0.9 and beta_2 = 0.999 with 544,970 trainable parameters. The total training time took approximately 4 h (1.5 h for YOLO and 2.5 h for U-Net).

2.1.1 Input Data for Neural Network

The input data should be standardized to be inserted into a neural network using YOLOv7. The proper format of data is to have pairs of images and text files, where the text files contain information about the objects in the image, in it the class of the object should be specified, as well as the coordinates of the center of the bounding box with the high and width of that box, Fig. 4 shows the proper encoding of the data. Coordination and dimensional information should be normalized to have values between zero and one, this can be done by assuming that the top left corner of the image has the coordination of (0, 0) and the bottom right corner has the coordination of (1, 1), see Fig. 5.

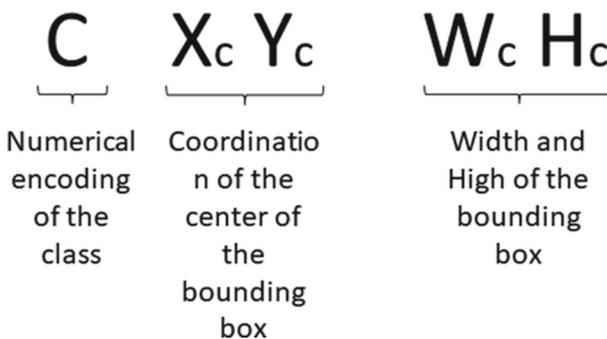


Fig. 5. Data encoding inside text files so that they can be used as input to the YOLOv7 algorithm.

Thus, even though the masks provided by experts have only the semantic segmentation information i.e., the class of each voxel, it is relatively easy to generate information about the bounding box, since the algorithm needs only to get the largest bounding box

in each plane to build a good 3D box. We chose only to train on the 2D images that have bounding boxes with sizes greater than 95.00% of the largest bounding box in that image. In practice, the steps to generate the data for YOLOv7 are following repeat step two for the other planes.

- Step 1: Load the image.
- Step 2: Iterate through the slices in one plane.
- Step 3: Check if the image has a pancreas.
- Step 4: if not skip this slice.
- if yes:
 - Detect the bounding box.
 - Register its coordinates and its size,
 - Register the number of the slice.
- Step 5: When done from all slices, calculate the max bounding box size.
- Step 6: iterate through the registered sizes.
- Step 7: if the size of the bounding box is greater or equal to the max size:
 - Adjust the window/level of the image to the abdominal preset.
 - Export the image as a jpeg file.
 - Normalize the coordinates.
 - Add numerical encode for the plane (0 for Axial, 1 for Coronal, and 2 for Sagittal)
 - Export this information as a text file that has the same name as the image.

After generating all training data, it should be divided into three sets, i.e., training 60.00%, validation 10.00%, and testing 30.00%. The training set is the largest portion of the dataset and is used to train the model. It is used to optimize the model's parameters and learn patterns from the data. The model is exposed to both input data and corresponding output labels during training. The validation set is used to fine-tune the model and select the best hyperparameters. It is typically used for model selection, hyperparameter tuning, and performance monitoring. The validation set helps to evaluate the model's performance on unseen data and can be used to make adjustments to improve the model's generalization. The test set is a completely independent dataset that is used to evaluate the final performance of the model. It is used to simulate the model's performance on new, unseen data. The test set should be representative of the real-world data the model is expected to encounter. The model should not be exposed to the test set during the training or validation phases to prevent any potential bias or overfitting.

Training time took approximately 4 h (1.5 h for YOLO v7 algorithm and 2.5 h for U-Net). The YOLO v7 algorithm model performance expressed as a relation between the precision and recall is visualized on the precision-recall (PR) plot in Fig. 6.

The U-Net model performance expressed as a dice score in the test set (independent from model development) was 0.73 (95.00%CI: 0.68–0.77) for healthy subjects and 0.72 (95.00%CI: 0.7–0.74) for patients with cancer. The IOU was 0.58 (95%CI: 0.53–0.63) and 0.57 (95.00%CI: 0.54–0.6), respectively. Voxel-level AUCs were 0.9865 (95.00%CI: 0.98646–0.98658) and 0.9898 (95.00%CI: 0.98976–0.98989), respectively. High values

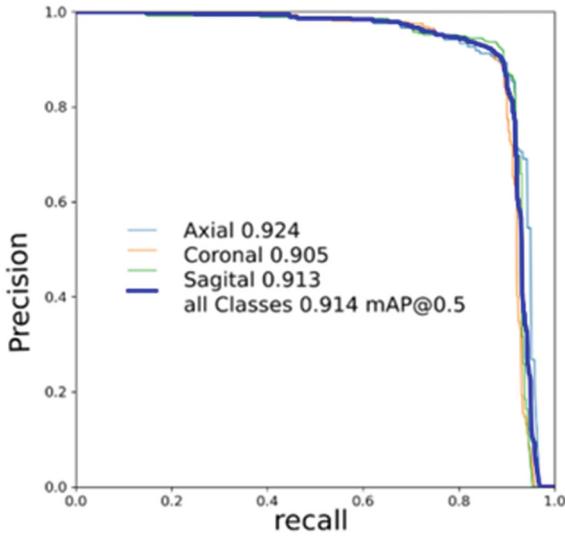


Fig. 6. Precision-Recall graph for YOLO algorithm.

of AUCs compared to the values of dice/IOU resulted in the good identification of background voxels, which are not taken into account in the dice/IOU metrics. The AUCs are voxel-level hence the number of observations is very large resulting in very narrow confidence intervals Fig. 7 and Fig. 8.

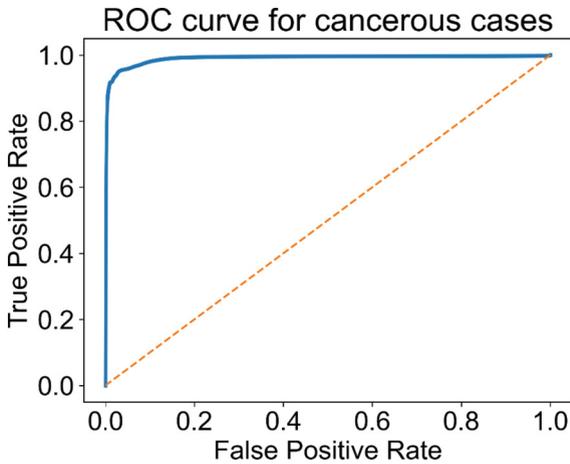


Fig. 7. ROC curve for cancerous cases.

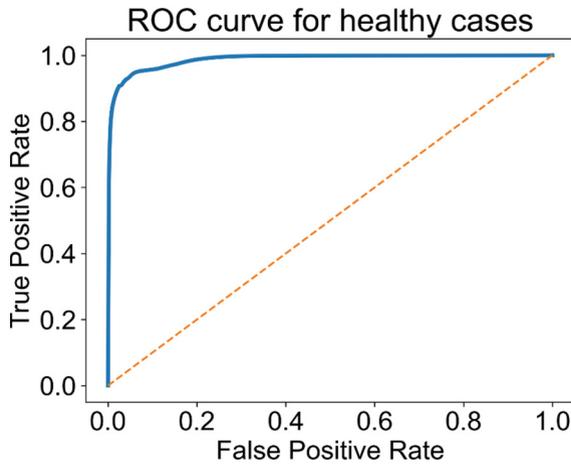


Fig. 8. ROC curve for healthy cases.

2.2 Future Manual Segmentation Approach to Machine Learning Concept

The next stage of this project is to gather data from local university hospitals and anonymized them, then data can be uploaded to the open-source medical imaging software, 3D Slicer (version 5.2.2). The resulting structures transfer into color-coded segments visible on CT scans and convert into a merged 3D file in STL format. Bearing in mind patient experience, additional structures also visualize to provide a more robust image that while largely not surgically relevant, helped them see and understand the relationship between local and overall anatomy, the extent of their surgery, and precisely what findings on a CT warranted them being scheduled for surgery to begin with. A standardized approach is to adopted to always visualize the same structures namely, bone, arterial, venous (portal and systemic), liver, gall bladder, stomach, spleen, kidneys, and pancreas, together with masses described by radiology along with separately marked areas of resection and potential anastomosis. According to this approach, several stages of manual segmentation have been provide:

- Step 1: independent radiologists make the characteristic points and the pancreas on each cross-section of each image according to a developed and strictly defined protocol, which served as partly a training set and partly as a bounding box.
- Step 2: detection models for each point and pancreas are constructed. The diagnostic capabilities of the constructed models, calculated based on the test set, are used to calculate the sample size required for validation, which will be performed on an independent, new dataset obtained from other centers. The new data set is evaluated by a panel of three independent experts, which enabled the selection of the so-called Ground Truth, based on which the diagnostic ability to detect characteristic points and the pancreas will be assessed. For each point, a binary decision will be determined, as well as the probabilities of the given class. Parameters like sensitivity, and specificity, the area under the ROC curve, and parameters related to the classical (JA) FROC methodology for MRMC-type problems are used as diagnostic endpoints. For each

point, a binary decision will be determined as well as the probability of cancer, i.e. semantic segmentation.

- Step 3: segmentation of the pancreas and tumor. A set of pancreatic cancer patient data is obtained. The CT images are cropped to contain only an area that included the pancreas. Segmentation is performed by radiologists on each cross-section of each image according to a developed and strictly defined protocol, which is treated as partly a training and partly a test set.
- Step 4: validation results from the previous two steps (Step 2 and Step 3) are used to calculate the required number of cases to adequately assess the diagnostic capacity of the created models in comparison to radiologists. A panel of three experts segmented the pancreas, with cancer confirmed by histopathology, forming the so-called Ground Truth – a reference for assessing the diagnostic capacity of both radiologists and algorithms. A validation dataset was collected from pancreatic cancer patients and non-pancreatic patients who had a CT.
- Step 5: radiologists segment images of the pancreas and tumors. An algorithm-based segmentation was performed independently (first detect the area of interest, then segment the pancreas and tumor).
- Step 6: The cases segmented by the algorithm are reviewed by radiologists. This allowed the radiologists' diagnostic capacity to be compared to the algorithm and radiologists assisted by the algorithm. The time to diagnose, cognitive load, and fatigue of the radiologist is also measured. Sensitivity, specificity, the area under the ROC curve, and parameters related to the classical (JA)FROC methodology for MRMC-type problems were diagnostic endpoints. The ability to segment was assessed using the Dice coefficient. Moreover, the number of observations in the validation set is sufficient to perform analogous analyses with subgroups defined by confounding factors.

2.3 Mixed Reality Application with the Possibility of Artificial Intelligence Enhancement as a Future Approach

Development of a mixed reality application that has the potential for enhancement through the integration of Artificial Intelligence (AI) technology. The mention of AI enhancement implies that the application could utilize AI algorithms, techniques, or functionalities to enhance its capabilities. AI can bring various benefits to mixed reality applications, such as improved object recognition, intelligent interaction, natural language processing, or personalized experiences. The specific AI enhancements would depend on the goals and requirements of the application, but the overarching idea is to leverage AI to augment and optimize the mixed reality experience for users. Based on our knowledge of manual individual case studies preprocedural planning, as described below, we can easily transfer this application to an AI approach based on dedicated algorithms like YOLO.

The creation of the Mixed Reality-based application, which will be tailored to run on HMDs, e.g., Microsoft HoloLens 2, to display Dicom data as 3D models can be described in the two steps. The first is to create the model, using data in the Dicom format. The second is the creation of the application to visualize the model. Before proceeding to step one, it is necessary to understand what kind of data will be processed in step one.

When performing an MRI or CT scan the human body is split into slices of varying millimeter thickness. Each slice is saved next as a grey-scale Dicom file similar to the picture. The quality of a model resulting from this data depends on the size of the pixel used to save the slice as a picture and the thickness between each slice. Typical thickness can be between 0.5–6 mm, depending on the purpose and quality needed. Thinner slices minimize the signal-to-noise ratio and produce optimal 3D segmentation models. The first step in segmentation utilizes the open-source software 3DSlicer. Model creation (step 1) is as follows:

- Load Dicom files into 3DSlicer. Data is displayed as grayscale pictures where the user can manipulate the slider to move between slices.
- Using the module “Segment Editor” the user creates layers on visible image slices signifying the anatomical structure to be segmented.
- The 3DSlicer function “Threshold”, allows users to set a range of greyscale that will automatically label structures fitting within that range across all available slices. This involves specifying the minimum and maximum intensities based on the histogram of a region of interest in the slice view, i.e., where a high-intensity, white region such as bone is selected while omitting soft tissues. This allows users to preview what will be included. After confirming that the selected range captures the desired structures, the resulting mesh can be previewed in 3D.
- This mesh is referred to as a mask and provides the limits of structures that will be segmented.
- Additional algorithms applied to this mesh include seeding and filling slices. Seeding involves tagging two or more specific regions of interest in sagittal, coronal, and axial views of a CT scan. For example, an object and background, where the object corresponds to the desired organ and the background corresponds to all undesired regions. After tagging these respective seeds, a customized GrowCut algorithm is applied to grow the region of seeded interest while excluding the region of undesired structures. This expands the binary “object” and “background” tags voxel by voxel along with a statistical weight as to the probability a given voxel still corresponds to the object. Filling slices is similar except it seeds only an “object” in one view, i.e., axial predominantly for vertical structures such as the inferior vena cava. Upon manually tagging the object in multiple levels of the CT scan, the Fill between Slices algorithm, based on ND morphological contour interpolation, is applied to form a 3D mask corresponding to the object. These algorithms are automated, however, the resulting models require varying degrees of manual correction which is done before accepting the finalized 3D mask they generate.
- Exporting the finalized model will be done via the function “Export segment to file” with OBJ format. This will then be accepted by Unreal Engine.

Creating the MR-based application (step 2) with the Unreal Engine is as follows

- Install the Visual Studio package.
- Start Unreal Engine and create a simple new blank mobile application.
- Add the plugin “Microsoft OpenXR” which adds tools useful for Extended Reality devices.
- Create a new object which will use the HoloLens 2 positional data to create a virtual space.

This allows the application to understand the position and perspective from which a user wearing the HoloLens 2 will be watching the 3D model (see Figure 9).

- Create an AR session to allow applications to link with Augmented Reality devices.
- AR session configuration.
- A final plugin to the project, “Mixed Reality UX Tools” will control interactions between the user, environment, and the displayed model.
- Import the created 3D model and interact with it in a virtual space
- Finally, this bundle is packaged as a file that can be directly imported to any HoloLens 2 device (see Fig. 10).

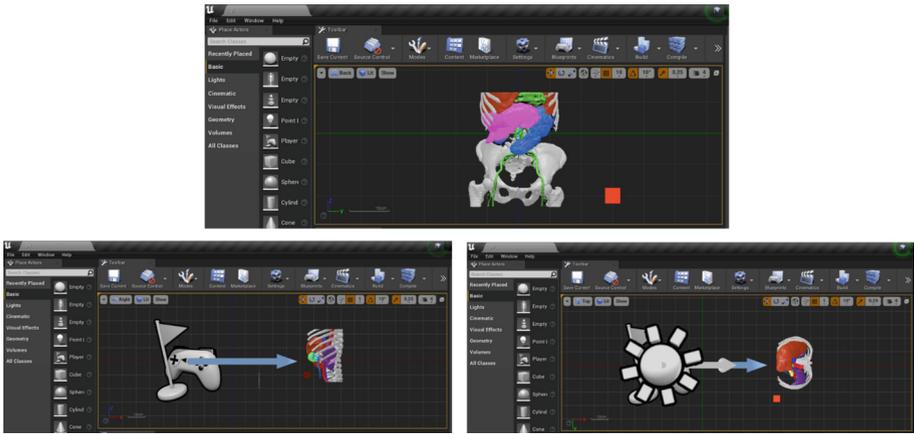


Fig. 9. A final plugin to the project, “Mixed Reality UX Tools” will control interactions between the user, environment, and the displayed model.

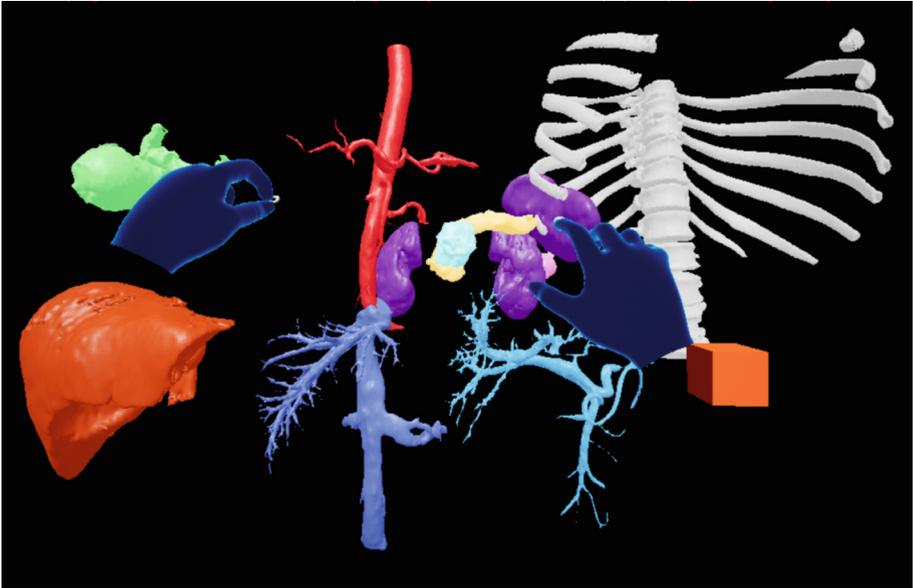


Fig. 10. Manual individual case studies preprocedural planning as a future approach for AI algorithms. 3D visualization of the pancreas and related organs as a digital file that can be directly imported to any HoloLens 2 device.

3 Case Study Description to Present the Concept of Automatization for Preprocedural Planning with 3D Visualization

The presented case study refers to a 74-year-old male with cancer involving the head of the pancreas who was scheduled for a pancreaticoduodenectomy. Upon review of the organ segmentation model, a region at the SMV/PV junction measuring approximately 2.5 cm was identified as likely requiring intraoperative reconstruction to avert changing the procedure into a prophylactic bypass anastomosis, see Fig. 11 and Fig. 12.

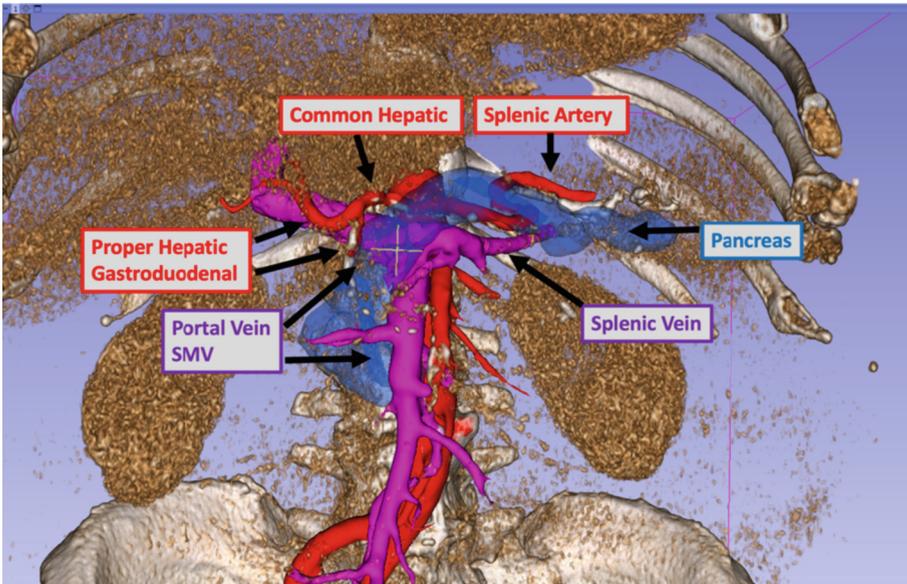


Fig. 11. Segmentation and volume rendering of key anatomical landmarks including arteries, veins, and pancreas. The arteries are red, the veins are purple, and the pancreas was marked by translucent blue color. The yellow crosshair indicates the region overall between the SMV/PV and pancreas, i.e., the area potentially requiring vessel reconstruction.

Based on the source data, is it possible to visualize images in the form of 3D images in the pre-procedural process (see Fig. 13) as well as interprocedural (see Fig. 14).

The results can be confirmed intraoperatively where a patch was successfully applied following vessel dissection and pancreatic head resection (see Fig. 15).

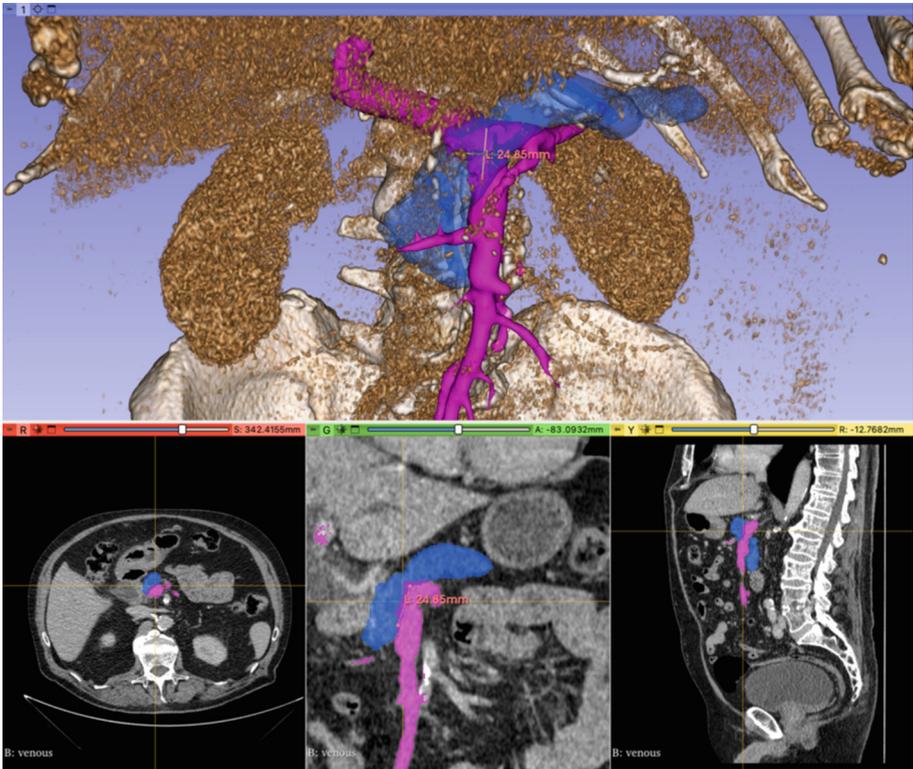


Fig. 12. Segmentation and volume rendering including CT scan image views of the approximately 2.5 cm region of potential vessel reconstruction. The arteries were marked by red color, veins are marked by purple color, while the pancreas was marked by translucent blue color. The yellow crosshair indicates the region overall between the SMV/PV and pancreas, i.e., the area potentially requiring vessel reconstruction (lower image: views from left to right – horizontal, coronal, sagittal). (Color figure online)



Fig. 13. Distal pancreatectomy sample. Superior view angled caudal to rostral showing a model rendering that is to be uploaded onto the HoloLens 2. Arteries are in red, systemic veins in blue, portal veins in purple, and lungs in pink, with a yellow lesion proximal to the pancreatic tail section in orange. (Color figure online)



Fig. 14. Intraoperative use of the HoloLens2 depicting surgeons discussing a case in a shared experience virtual space.

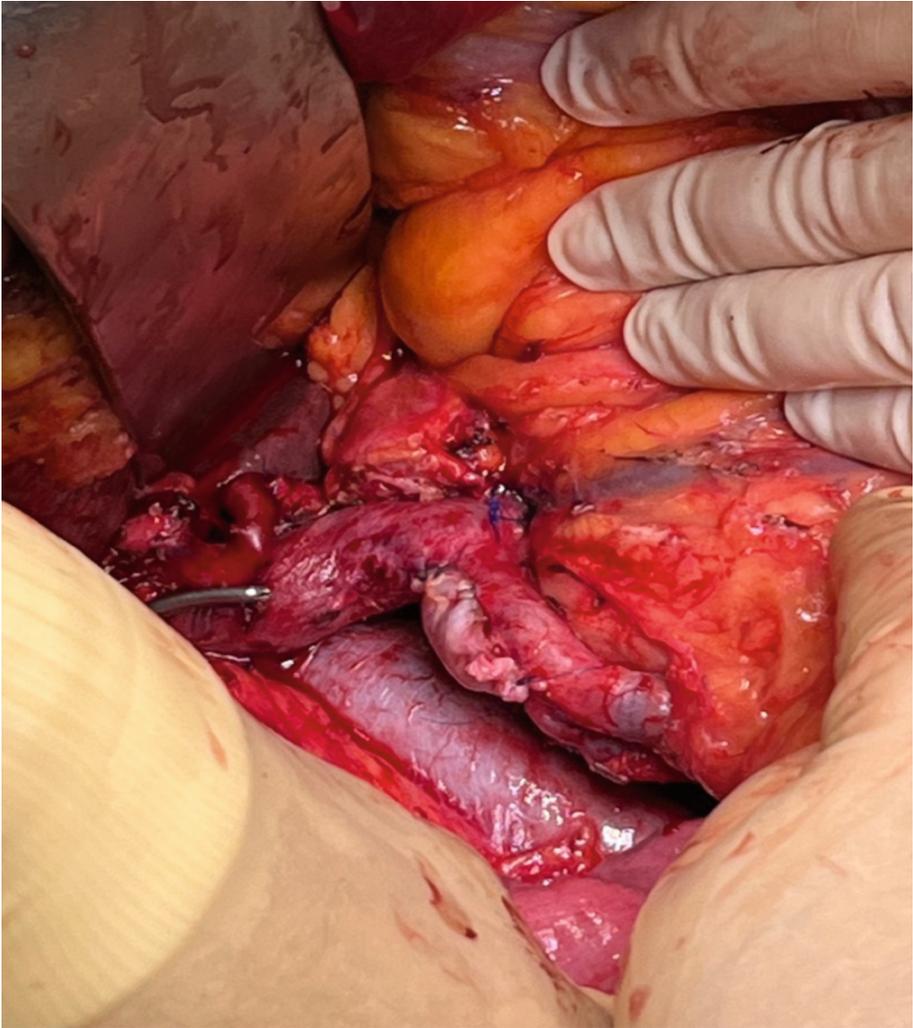


Fig. 15. Completed blood patch at the region predicted by 3D visualization.

4 Conclusion

The availability of an automated and non-invasive tool for visualization of the pancreas and its abnormalities can greatly improve the pre and intraoperative process along with providing previously impossible benefits to patient experience. Organ and vessel segmentation with subsequent 3D visualization with Artificial Intelligence and Mixed Reality enhancement can substantially improve planning, reduce the incidence of unforeseen intraoperative decision-making, and improve clinical outcomes. Thus, the proposed approach can be considered as both an accurate and efficient image segmentation tool

of the pancreas and its abnormalities, which is crucial for diagnostics and delivery of targeted therapy.

In the future, the application of preoperative image-guided 3D visualization supported by Machine Learning to the prediction of organ reconstruction during pancreaticoduodenectomy via Head-Mounted Displays holds tremendous potential. This combination of technologies can revolutionize surgical planning and execution, leading to improved outcomes and enhanced patient care.

With advances in medical imaging techniques, including computed tomography and magnetic resonance imaging, surgeons have access to detailed 3D representations of the patient's anatomy. By utilizing ML algorithms, these preoperative images can be analyzed and processed to provide accurate predictions and simulations of organ reconstruction outcomes.

Head-Mounted Displays, such as augmented reality glasses, enable surgeons to visualize the patient's anatomy in a more immersive and intuitive manner. By wearing HMDs, surgeons can overlay the 3D visualizations onto the patient's body during surgery, enhancing their spatial awareness and precision.

Machine Learning algorithms play a crucial role in this process. They can be trained on large datasets of preoperative images, incorporating factors such as patient demographics, disease characteristics, and surgical outcomes. This training enables the ML models to learn patterns and make accurate predictions about the success of different organ reconstruction approaches.

By leveraging ML-supported 3D visualization and HMDs, surgeons can benefit from the following future perspectives:

- **Enhanced Surgical Planning:** Surgeons can thoroughly examine the patient's anatomy in 3D and visualize potential complications before the actual surgery. This allows them to devise personalized surgical plans, considering factors like the location of blood vessels, tumor margins, and other critical structures.
- **Real-time Intraoperative Guidance:** During the surgery, HMDs can overlay the preoperative 3D visualizations onto the surgical field, providing real-time guidance to the surgeon. ML algorithms can help identify and highlight key anatomical structures, assisting the surgeon in making informed decisions and optimizing the reconstruction process.
- **Reduced Complications:** The combination of ML predictions and visual guidance can help minimize the risk of complications during pancreaticoduodenectomy. Surgeons can anticipate potential issues, such as vascular involvement or inadequate margins, and adjust their approach accordingly, thereby reducing postoperative complications and improving patient outcomes.
- **Training and Education:** ML-supported 3D visualization via HMDs can be a valuable tool for training new surgeons. Trainees can benefit from interactive virtual simulations, allowing them to practice complex procedures and refine their skills in a controlled environment before performing surgeries on actual patients.
- **Remote Collaboration and Expertise:** The integration of ML and HMDs also enables remote collaboration and consultation with experts. Surgeons can share live surgical views and 3D visualizations with off-site specialists, allowing for real-time guidance

and support during complex cases. This can be particularly beneficial in regions with limited access to specialized surgical expertise.

While these future perspectives hold great promise, their realization requires further advancements in medical imaging, ML algorithms, and HMD technologies. Additionally, comprehensive validation through clinical trials and regulatory approvals will be essential to ensure the safety and efficacy of these applications.

In summary, the application of preoperative image-guided 3D visualization supported by Machine Learning to the prediction of organ reconstruction during pancreaticoduodenectomy via Head-Mounted Displays has the potential to transform surgical practices. By leveraging ML algorithms and immersive visualization, surgeons can enhance surgical planning, and intraoperative guidance, and ultimately improve patient outcomes in this complex surgical procedure.

Acknowledgments. This study was supported by the National Centre for Research and Development (Grant Lider No. LIDER/17/0064/L-11/19/NCBR/ 2020).

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations. Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acidi, B., Ghallab, M., Cotin, S., Vibert, E., Golsé, N.: Augmented reality in liver surgery. *J. Visc. Surg.* **160**(2), 118–126 (2023). <https://doi.org/10.1016/j.jvisurg.2023.01.008>
- Alom, M.Z., Yakopcic, C., Hasan, M., Taha, T.M., Asari, V.K.: Recurrent residual U-Net for medical image segmentation. *J. Med. Imaging (Bellingham)* **6**(1), 014006 (2019). <https://doi.org/10.1117/1.Jmi.6.1.014006>
- Amit, Y., Felzenszwalb, P., Girshick, R.: Object detection. In *Computer Vision: A Reference Guide*, pp. 1–9. Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-03243-2_660-1
- Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017). <https://doi.org/10.1109/tpami.2016.2644615>
- Bagheri, M.H., et al.: Technical and Clinical factors affecting success rate of a deep learning method for pancreas segmentation on CT. *Acad. Radiol.* **27**(5), 689–695 (2020). <https://doi.org/10.1016/j.acra.2019.08.014>
- Bochkovskiy, A., Wang, C.Y., Liao, H.-Y.M.: YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv, abs/2004.10934* (2020)

- Brockmeyer, P., Wiechens, B., Schliephake, H.: The role of augmented reality in the advancement of minimally invasive surgery procedures: a scoping review. *Bioengineering (Basel)*, **10**(4), 501 (2023). <https://doi.org/10.3390/bioengineering10040501>
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(4), 834–848 (2018). <https://doi.org/10.1109/TPAMI.2017.2699184>
- Chen, P.T., et al.: Pancreatic cancer detection on CT scans with deep learning: a nationwide population-based study. *Radiology* **306**(1), 172–182 (2023). <https://doi.org/10.1148/radiol.220152>
- Dai, S., Zhu, Y., Jiang, X., Yu, F., Lin, J., Yang, D.: TD-Net: trans-deformer network for automatic pancreas segmentation. *Neurocomputing* **517**, 279–293 (2023). <https://doi.org/10.1016/j.neucom.2022.10.060>
- England, A., et al.: A comparison of perceived image quality between computer display monitors and augmented reality smart glasses. *Radiography* **29**(3), 641–646 (2023). <https://doi.org/10.1016/j.radi.2023.04.010>
- Garlinska, M., Osial, M., Proniewska, K., Pregowska, A.: The influence of emerging technologies on distance education. *Electronics* **12**(7), 1550 (2023). <https://www.mdpi.com/2079-9292/12/7/1550>
- Gheorghie, G., et al.: Early Diagnosis of pancreatic cancer: the key for survival. *Diagnostics (Basel)*, **10**(11), 869 (2020). <https://doi.org/10.3390/diagnostics10110869>
- Granata, V., et al.: Risk assessment and pancreatic cancer: diagnostic management and artificial intelligence. *Cancers (Basel)*, **15**(2), 351 (2023). <https://doi.org/10.3390/cancers15020351>
- Hasan, M.M., Islam, M.U., Sadeq, M.J., Fung, W.K., Uddin, J.: Review on the evaluation and development of artificial intelligence for COVID-19 containment. *Sensors (Basel)*, **23**(1), 527 (2023). <https://doi.org/10.3390/s23010527>
- Hayward, J., Alvarez, S.A., Ruiz, C., Sullivan, M., Tseng, J., Whalen, G.: Machine learning of clinical performance in a pancreatic cancer database. *Artif. Intell. Med.* **49**(3), 187–195 (2010). <https://doi.org/10.1016/j.artmed.2010.04.009>
- Hosny, K.M., Kassem, M.A., Foad, M.M.: Classification of skin lesions using transfer learning and augmentation with Alex-net. *PLoS ONE* **14**(5), e0217293 (2019). <https://doi.org/10.1371/journal.pone.0217293>
- Janssen, B.V., et al.: Artificial intelligence-based segmentation of residual tumor in histopathology of pancreatic cancer after neoadjuvant treatment. *Cancers (Basel)*, **13**(20), 5089 (2021). <https://doi.org/10.3390/cancers13205089>
- Kenner, B., et al.: Artificial intelligence and early detection of pancreatic cancer: 2020 summative review. *Pancreas* **50**(3), 251–279 (2021). <https://doi.org/10.1097/MPA.0000000000001762>
- Klinker, K., Wiesche, M., Krcmar, H.: Digital transformation in health care: augmented reality for hands-free service innovation. *Inf. Syst. Front.* **22**(6), 1419–1431 (2020). <https://doi.org/10.1007/s10796-019-09937-7>
- Kochanski, R.B., Lombardi, J.M., Laratta, J.L., Lehman, R.A., O’Toole, J.E.: Image-guided navigation and robotics in spine surgery. *Neurosurgery* **84**(6), 1179–1189 (2019). <https://doi.org/10.1093/neuros/nyy630>
- Lee, K.-S., et al.: Usefulness of artificial intelligence for predicting recurrence following surgery for pancreatic cancer: retrospective cohort study. *Int. J. Surg.* **93**, 106050 (2021). <https://doi.org/10.1016/j.ijsu.2021.106050>
- Li, X., et al.: Multi-task refined boundary-supervision U-Net (MRBSU-Net) for gastrointestinal stromal tumor segmentation in endoscopic ultrasound (EUS) images. *IEEE Access* **8**, 5805–5816 (2020). <https://doi.org/10.1109/ACCESS.2019.2963472>

- Nguyen, T., Plishker, W., Matisoff, A., Sharma, K., Shekhar, R.: HoloUS: augmented reality visualization of live ultrasound images using hololens for ultrasound-guided procedures. *Int. J. Comput. Assist. Radiol. Surg.* **17**(2), 385–391 (2022). <https://doi.org/10.1007/s11548-021-02526-7>
- Pan, J., et al.: Real-time segmentation and tracking of excised corneal contour by deep neural networks for DALK surgical navigation. *Comput. Meth. Progr. Biomed* **197**, 105679 (2020). <https://doi.org/10.1016/j.cmpb.2020.105679>
- Patric, B., Baowei, F.: An Advanced System with Advanced User Interfaces for Image-Guided Intervention Applications. *SPIE, Proc* (2023)
- Polyviou, A., Zamani, E.D.: Are we nearly there yet? a desires & realities framework for Europe's AI strategy. *Inf. Syst. Front.* **25**(1), 143–159 (2023). <https://doi.org/10.1007/s10796-022-10285-2>
- Pregowska, A., Osial, M., Dolega-Dolegowski, D., Kolecki, R., Proniewska, K.: Information and communication technologies combined with mixed reality as supporting tools in medical education. *Electronics*, **11**(22), 3778 (2022). <https://www.mdpi.com/2079-9292/11/22/3778>
- Quero, G., et al.: Virtual and augmented reality in oncologic liver surgery. *Surg. Oncol. Clin. N. Am.* **28**(1), 31–44 (2019). <https://doi.org/10.1016/j.soc.2018.08.002>
- Rawla, P., Sunkara, T., Gaduputi, V.: Epidemiology of pancreatic cancer: global trends, etiology and risk factors. *World J. Oncol.* **10**(1), 10–27 (2019). <https://doi.org/10.14740/wjon1166>
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016, 27–30 June 2016)
- Redmon, J., Farhadi, A.: YOLO9000: Better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7263–7271 (2017, 21–26 July 2017)
- Ren, Y., Zou, D., Xu, W., Zhao, X., Lu, W., He, X.: Bimodal segmentation and classification of endoscopic ultrasonography images for solid pancreatic tumor. *Biomed. Sign. Process. Control* **83**, 104591 (2023). <https://doi.org/10.1016/j.bspc.2023.104591>
- Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv, abs/1505.04597* (2015)
- Neher, P.F., Götz, M., Norajitra, T., Weber, C., Maier-Hein, K.H.: A machine learning based approach to fiber tractography using classifier voting. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9349, pp. 45–52. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9_6
- Ruger, C., Feufel, M.A., Moosburner, S., Ozbek, C., Pratschke, J., Sauer, I.M.: Ultrasound in augmented reality: a mixed-methods evaluation of head-mounted displays in image-guided interventions. *Int. J. Comput. Assist. Radiol. Surg.* **15**(11), 1895–1905 (2020). <https://doi.org/10.1007/s11548-020-02236-6>
- Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440 (2014)
- Simpson, A.L., et al.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms (2019). [arXiv:1902.09063](https://arxiv.org/abs/1902.09063). <https://ui.adsabs.harvard.edu/abs/2019arXiv190209063S>. Accessed 01 Feb 2019
- Thoma, M.: A Survey of Semantic Segmentation. *ArXiv, abs/1602.06541* (2016)
- Tomassini, S., Anbar, H., Sbrollini, A., Mortada, M.J., Burattini, L., Morettini, M.: A double-stage 3D U-net for on-cloud brain extraction and multi-structure segmentation from 7T MR volumes. *Information*, **14**(5), 282 (2023). <https://www.mdpi.com/2078-2489/14/5/282>
- von Haxthausen, F., Rüger, C., Sieren, M.M., Kloeckner, R., Ernst, F.: Augmenting image-guided procedures through in situ visualization of 3d ultrasound via a head-mounted display. *Sensors* **23**(4), 2168 (2023). <https://www.mdpi.com/1424-8220/23/4/2168>

- Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y.M.: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. ArXiv, abs/2207.02696 (2022)
- Wang, K., et al.: Fluorescence image-guided tumour surgery. *Nature Rev. Bioeng.* **1**(3), 161–179 (2023). <https://doi.org/10.1038/s44222-022-00017-1>
- Wu, W., Gao, L., Duan, H., Huang, G., Ye, X., Nie, S.: Segmentation of pulmonary nodules in CT images based on 3D-UNET combined with three-dimensional conditional random field optimization. *Med. Phys.* **47**(9), 4054–4063 (2020). <https://doi.org/10.1002/mp.14248>
- Young, M.R., Abrams, N., Ghosh, S., Rinaudo, J.A.S., Marquez, G., Srivastava, S.: Prediagnostic image data, artificial intelligence, and pancreatic cancer: a tell-tale sign to early detection. *Pancreas* **49**(7), 882–886 (2020). <https://doi.org/10.1097/MPA.0000000000001603>
- Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested u-net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support Cham*, pp. 3–11 (2018). https://doi.org/10.1007/978-3-030-00889-5_1
- Zhu, X., Lyu, S., Wang, X., Zhao, Q.: TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In: *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 2778–2788 (2021)