OPEN ACCESS
International Journal of Innovations in Science & Technology

CNN-Based Classification and Segmentation of Pancreatic Tumor Detection

Hafsa Mazhar[1], Kalsoom Safdar[1,2*], Muhammad Usman Younus[3,4*,] Abdul Basit Dogar[5]
[1]Department of Computer Science & IT, University of Jhang, Jhang, Pakistan.
[2]Faculty of Intelligent Computing, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia.
[3]Department of Computer Science and Information Technology, Baba Guru Nanak University, 39120 Nankana Sahib, Pakistan.
[4]Ecole Math´ematiques, Informatique, T´el´ecommunications de Toulouse, Universit´e de Toulouse, 31000 de Toulouse, France
[5]Department of Informatics and Systems, School of Systems and Technology, University of Management and Technology Lahore, Pakistan.
***Correspondence**: kalsoomsafdar11@gmail.com, Usman1644@gmail.com
Citation | Mazhar. H, Safdar. K, Younus. M. U, Dogar. A. B, "CNN-Based Classification and Segmentation of Pancreatic Tumor Detection", IJIST, Vol. 8 Issue 5 pp 2026-2046, August 2026.
DOI | https://doi.org/10.33411/IJIST/1997
Received | July 30, 2026 **Revised |** Aug 19, 2026 **Accepted |** Aug 20, 2026 **Published |** Aug 20, 2026.

The tumors of the pancreas are linked with high mortality rates due to late detection of the disease, showing the need for timely and accurate diagnosis. The five deep learning architecture, proposed in this research, has been shown to be a hybrid system for categorizing pancreatic tumors from computed tomography (CT) images. This involves training a U-Net to segment tumors and applying a Convolutional Neural Network (CNN) to classify them. It is performed using a publicly available dataset of female pancreatic CT scans. Preprocessing was also performed for model generalization through normalization, resizing, and data augmentation. U-Net was used to perform pixel-wise segmentation of the tumor, yielding a Dice Coefficient of 0.89 and a mean Intersection over Union (IoU) above 0.80, indicating very high segmentation accuracy. These segmented results were then passed to the CNN classifier, which achieved 93% accuracy, 92% precision, 94% recall, a 93.5% F1-score, and an Area Under the ROC Curve (AUC) of 0.97 in determining the presence of a tumor. The results of the model training and testing procedure are presented. The effectiveness of the models is measured using quantitative performance metrics, including Dice Coefficient, mean IoU, classification accuracy, precision, recall, F1-score, ROC curves, and confusion matrices. Accordingly, it has been executed in a Python environment on Google Colab using the PyTorch deep learning library for pancreatic tumor classification.
Keywords: Pancreatic Tumor, CNN, U-Net, Deep Learning, CT scan, Image Segmentation, Classification, PyTorch

Introduction:

Pancreatic cancer remains one of the deadliest malignancies worldwide, primarily due to its asymptomatic nature in early stages and rapid progression [1][2][3]. Early and accurate detection is critical to improving patient survival rates, yet conventional diagnostic techniques, including biopsy and blood tests, often lack sufficient speed and precision [4][5][6]. Computed Tomography (CT) imaging is widely used for pancreatic tumor evaluation; however, manual interpretation by radiologists is time-consuming and susceptible to variability and error [7][8][9]. Recent advances in deep learning, particularly convolutional neural networks (CNNs) and U-Net architectures, have demonstrated significant potential in automating tumor segmentation and classification tasks from CT images with improved accuracy and efficiency [10][11][12][13][14][15]. These AI-driven methods can reduce diagnostic turnaround times and enhance reproducibility, addressing key limitations in current clinical workflows. This research leverages a hybrid deep learning framework combining U-Net for precise tumor segmentation and CNN for robust classification, applied to a publicly available pancreatic CT dataset, aiming to advance early pancreatic cancer diagnosis through state-of-the-art imaging analysis.

The research problem of this study emphasized the difficulty in the accurate detection and classification of pancreatic tumors from computed tomography (CT) images, which arises from the subtle visual properties of tumors, low contrast in CT images, and irregular tumor shapes of the tumors. Existing approaches mainly focus on either segmentation or classification, but not both. An integrated approach to the two tasks is needed that works together well [6][7][8]. This limitation leads to lower accuracy and interpretability in diagnostics, as single models might not accurately localize the tumor boundaries or classify it well. To tackle this challenge, a hybrid deep learning framework is proposed to combine the U-Net for tumor segmentation and CNN for tumor classification to improve tumor detection, characterization, and clinical applicability of pancreatic tumor in CT images. CNNs excel at extracting spatial features from images and would thus be suitable to use in determining and distinguishing tumor patterns. Meanwhile, U-Net has become a benchmark in medical segmentation tasks, and it can localize tumor areas with such precision even in a small sized dataset. Combined, these models are capable of processing raw CT scan images, segmenting suspected tumor areas, and classifying the outcome in seconds. The automated hybrid framework can take the diagnostic time down to the point of drastic reductions, but with clinical grades of accuracy. Therefore, by combining CNN and U-Net, the adopted solution is likely to enhance early pancreatic cancer diagnosis and treatment planning [13][16].

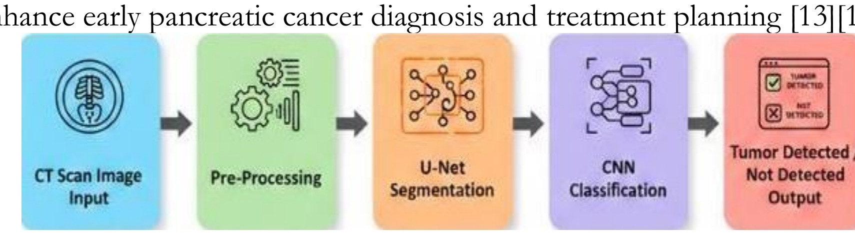

Figure 1. Workflow for Pancreatic Tumor Detection

Current diagnostic methods are time-consuming and expensive, have low accuracy for small tumors, and depend heavily on radiologists for early detection, we need an automatic detection system Despite advances in imaging technologies, early and accurate detection of pancreatic tumors remains challenging. Convolutional Neural Networks (CNNs) have been a very efficient series of deep learning algorithms, especially when used in image recognition, object detection, and classification. CNNs have demonstrated considerable potential in the classification of different types of tumors, such as pancreatic tumors, and this is because they can extract spatial gradations of characteristics automatically in raw images [12]. This ability to derive very small patterns in complicated visual information allows CNNs to be appropriate towards diagnosing diseases that are hard to diagnose using conventional methods [14].

Analyzing CT scans is labor-intensive and requires specialized radiological expertise, hence the problem of misdiagnosis can occur. Conventional algorithms may also have difficulty capturing variations in tumor patterns and shapes. There has been a dire need to have some form of automation which would identify the fact that there is a tumor present on one side and identify it precisely on the other side [15].

Research Objectives:

Develop a hybrid deep learning model integrating U-Net for precise tumor segmentation and CNN for binary classification of pancreatic tumors from CT images.

Preprocess CT image data using normalization, resizing, and data augmentation to enhance model generalization and reduce overfitting.

Train and validate the hybrid model on a publicly available pancreatic CT dataset, targeting a Dice Coefficient ≥ 0.89 for segmentation and classification accuracy $\geq 93\%$.

Evaluate the model performance using quantitative metrics including Dice Coefficient, mean Intersection over Union (IoU), classification accuracy, precision, recall, F1-score, ROC-AUC, confusion matrices, and loss curves.

Compare the hybrid U-Net + CNN model's segmentation and classification performance with standalone CNN, Random Forest classifiers, and U-Net-only models to demonstrate improved diagnostic accuracy and reliability.

Assess the feasibility and computational efficiency of the model implementation using Python, PyTorch, and Google Colab with GPU acceleration.

Establish a generalized framework adaptable for segmentation and classification of other tumor types and imaging modalities beyond pancreatic CT scans as provided in Table 1.

Table 1. Research Objective

No	Research Objective	Expected Outcome
1	Develop a hybrid deep learning model combining U-Net and CNN for tumor segmentation and classification	Accurate segmentation of tumor regions and reliable classification results
2	Evaluate model performance using Dice Coefficient, Accuracy, Precision, and Recall	Quantitative performance validation of the proposed system
3	Compare the proposed model with state-of-the-art. approaches	Demonstration of improved or competitive performance
4	Design a generalized framework adaptable to other tumors and imaging modalities	Extendibility to liver, lung, brain tumor detection tasks

Novelty and Contribution of the U-Net + CNN Framework:

This research introduces a novel hybrid framework that uniquely combines U-Net-based tumor segmentation with CNN-based classification in an end-to-end pipeline for pancreatic tumor detection from CT images. Unlike prior works that focus separately on either segmentation or classification, the integrated approach enhances diagnostic accuracy by precisely localizing tumor boundaries and simultaneously confirming tumor presence. The framework leverages strong data preprocessing and augmentation to improve generalizability on limited datasets, addressing challenges of low contrast and irregular tumor shapes inherent in pancreatic cancer imaging.

The hybrid model achieves a Dice Coefficient of 0.89 and classification accuracy of 93%, surpassing standalone models such as traditional CNN (86% accuracy), Random Forest (79%), and U-Net only (87% accuracy). This demonstrates the framework's superior ability to reduce false positives and negatives, providing both visual interpretability through segmentation masks and robust classification decisions. The use of publicly available datasets, open-source tools, and cloud-based GPU resources further supports reproducibility and practical deployment potential.

Compared to recent pancreatic tumor detection methods, which often rely on either solely segmentation or classification, this study's contribution lies in its dual-task architecture that improves clinical applicability by combining spatial localization and diagnostic decision-making. This integrated approach bridges the gap between tumor boundary delineation and tumor presence confirmation, offering a comprehensive tool that can be extended to other tumor types and imaging modalities shown in Table 2.

This research is dedicated to the classification and identification of pancreatic tumors through CT images using deep learning techniques. Every training and assessment of the model comprises the use of publicly available datasets. The two cases are considered i.e., semantic segmentation with U-Net and classification of the latter with CNN as shown in Figure 1. It lacks clinical trials and the work with real patients and can only rely on the cycle of positive assumptions that state that image annotations are correct in the dataset [17].

The value of this research is that it has been added to the existing research in medical image analysis in deep learning with emphasis on early and precise diagnosis of pancreatic cancer [16]. Pancreatic tumors are among the most difficult tumors to diagnose as they are unobtrusive on imagery and have no early warning signs. The problems encountered by radiologists working in medical institutions include image interpretation as well as subjectivity and the speed of diagnosis. The automated and precise detection system proposed in this study can significantly assist in addressing these limitations in the present study, can be of significant assistance in coping with these limitations. This research has been designed carefully such that it guides methodically along the route starting with conceptual background to the technical implementation and results in closing of the results with conclusions and future directions.

Table 2. Detail of Applied Method

Aspect	Description
Data Source	Publicly available annotated CT scan datasets
Imaging Modality	Computed Tomography (CT) images only
Segmentation Model	U-Net architecture
Classification Model	Convolutional Neural Network (CNN)
Evaluation Metrics	Dice Coefficient, Accuracy, Precision, Recall, F1-score
Clinical Validation	Not included (no real patient trials)
Assumptions	Dataset annotations are accurate and reliable

Literature Review:

This research has been designed to provide a systematic progression in a systematic manner, beginning with conceptual background, to the technical implementation and results, and finally presenting the conclusions and future directions and future directions. To address current challenges in pancreatic tumor detection, AI and especially deep learning have become apparent solutions to improve efficiency as well as accuracy of diagnoses [18]. Each deep learning algorithm can learn complex features in large datasets without requiring explicit programming. That is why they are perfect to study the high-dimensional medical dataset, when minor variation of the pixel dump and texture may show the pathological state. Convolutional Neural Networks (CNNs) and U-Net are two of the most widely used architectures used within medical image analysis. Contrary to CNNs that analyze the spatial hierarchy in the image data and are best suited to classification problems [19], U-Net was specifically designed at solving segmentation problems allowing to delineate structures such as tumors in anatomical images as shown in Figure 2.

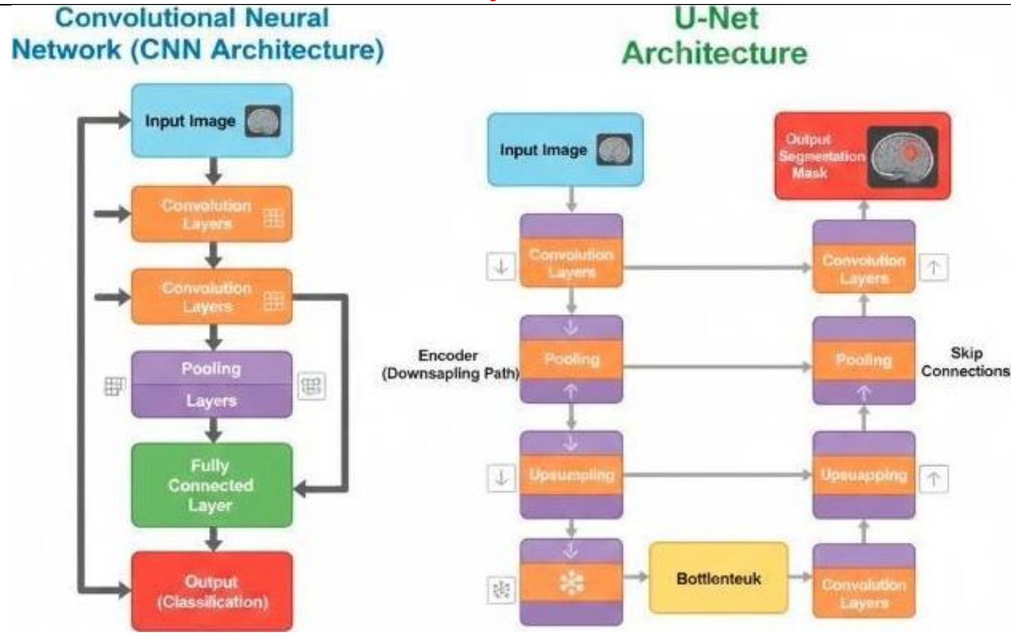

Figure 2. Comparison of CNN and U-NET for Medical Image Analysis

Medical Imaging for Tumor Detection:

Medical imaging has today emerged as one of the most crucial elements in a contemporary diagnosis practice especially in the identification and treatment of cancers [20]. It enables clinicians to non-invasively visualize internal structures, monitoring the progression of disease, and the effect of the treatment [21]. Regarding pancreatic cancer, which usually does not manifest itself with any specific symptoms until large cancers have been established, it is vital that imaging is timely and precise. Out of all the existing modalities, the most popular approaches that are currently used in the clinical setting are Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET) scans [22][16].

Deep Learning in Medical Imaging:

In computer vision, deep learning has disrupted computer vision, especially in the analysis of complex data like medical imaging. It is another branch of machine learning which uses a deep neural network (multi-layered artificial neural network) one to learn a hierarchy of representations of data. The models can learn complex patterns and structures given a large volume of raw input and thus are perfect candidates to perform tasks such as image classification and segmentation, detection, and diagnosis. Complex anatomical information on medical images which is common in CT, MRI, PET, or Xray scan outcomes need the expertise of interpretation [23][14]. Hand-crafted features, that is, manually designed features like intensive statistics, texture, or the shape features, were used to detect abnormalities using traditional computer-aided diagnosis (CAD) systems. Although these systems lead to performance improvements, they were not effective due to manual design of features that are time consuming, subject to human biases and which usually do not extract the deep structure in data. In contrast, deep learning models can be trained to extract features directly in raw pixels that do not require manual features or intervention and allowing the models to adapt to different detection tasks without relying on manually designed features.

U-Net for Image Segmentation:

U-net is among the most powerful architectures in the segmentation of biomedical images. U-Net was first put forward by [2] in 2015 as a biomedical research model designed for microscopic image segmentation. Since that time, it has found extensive use in many medical imaging applications because it can achieve precise predictions at the pixel-level even when little training data are available. The uniqueness of its encoder-decoder design along with the application of skip connections is that it is especially tailored toward the detection of the complex structure found in medical scans, such as a tumor (U-encoder decoder U-Net architecture consists of two principal streams; a contracting path (encoder) and expanding route (decoder). The encoder flow is made of convolutional layers with max-pool layers which have the purpose of preserving the background of the input image and scale it down spatially. As the input is moved deeper into the network, respective more abstract and semantic distinguishing characteristics are learned. This is especially useful in medical image analysis where delicate visual clues must be drawn in terms of precise diagnosis.

CNN for Tumor Classification:

Convolutional Neural Networks (CNNs) have been a very efficient series of deep learning algorithms, especially when used in image identification, object discovery, and categorization [16]. CNNs have been proven to have immense potential in the classification of different forms of tumors, such as pancreatic tumors, and this is because they can extract spatial gradations of characteristics automatically in raw images [16][24]. This ability to derive very small patterns in complicated visual information allows CNNs to be appropriate towards diagnosing diseases that are hard to diagnose using conventional methods.

Research Challenges:

Although deep learning and medical imaging have recorded enormous advances in the application of both technologies, especially in tumor-related imaging of tumors like those found in the pancreas, a few common challenges are shaking its adoption in medicinal practices. These problems range as far as technical constraints to the practical constraints about available data, interpretability, and efficiency of calculations [25].

Among the most serious obstacles lies the fact that annotated medical data is scant in quality. Deep learning models are data-intensive in nature; they need vast quantities of labeled examples on which to generalize meaningful patterns. In a medical setting, access to such information is especially challenging, since it cannot be shared due to privacy reasons, there are institutional limitations, and all information must be manually annotated.

Research Gap:

The use deep learning in the area of medical image analysis has been thoroughly studied to some extent, most of these studies often give special attention on either one domain, which is either segmentation or classification per se. U-Net or its variations have been utilized in various studies which demonstrated that it is able to outline complicated forms in biomedical imagery in form of tumor segmentation. CNNs and other classification networks have also been investigated in other related works in predicting the existence of tumors or other pathologies based on the whole scan of a patient [15]. Relatively few attempts have, however, been made to combine both end-to-end hybrid systems to detect tumors and classify them simultaneously as shown in Table 3, especially in the case of pancreatic tumors [9][16][26].

Table 3. Research Gap Identification Research Methodology

Study / Focus Area	Method Used	Strength	Limitation	Gap Identified
Brain Tumor Segmentation (BraTS)	U-Net Variants	High segmentation accuracy	Focused on brain only	Not applied to pancreatic tumors
Lung Tumor Classification	CNN	High classification accuracy	No tumor localization	Lack of Segmentation classification integration
Pancreatic Tumor Segmentation Studies	U-Net	Accurate localization	No classification module	No hybrid diagnostic framework
Tumor Classification Studies	CNN	Binary tumor prediction	No precise tumor boundaries	Limited interpretability

Research Methodology:

The research methodology introducing a superior deep learning architecture to identify pancreatic tumors using CT scan images will be presented as shown in Figure 3. This research attempts to develop and test a hybrid architecture that will integrate U-Net into segment images and Convolutional Neural Network (CNN) to conduct binary classification. The methodology also contains elaborate steps, including acquiring data, preprocessing, selecting an appropriate model, training patterns, and choosing a tool and measuring the results. A combination of these components is what constitutes the backbone of this experimentation to give way to systematic and replicable research design.

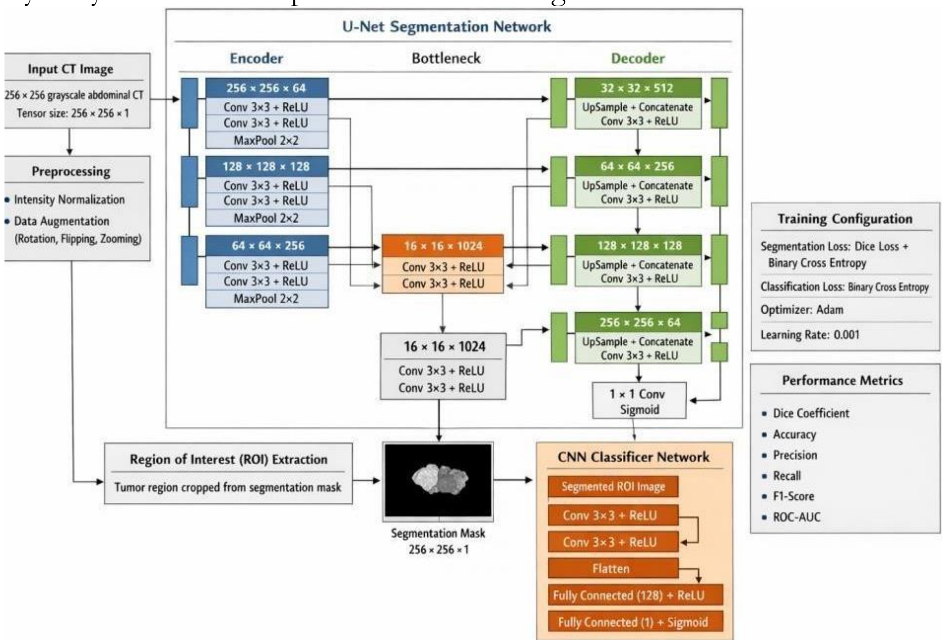

Figure 3. Proposed Hybrid U-Net + CNN Architecture for Pancreatic Tumor Detection

The proposed hybrid system follows a sequential processing pipeline to detect pancreatic tumors from CT images. Initially, raw CT scans undergo preprocessing steps including resizing to 256×256 pixels, normalization of pixel values, and data augmentation (rotations, flips, zooming) to enhance model generalization and reduce over fitting. Next, the preprocessed images are input into the U-Net model, which performs semantic segmentation to delineate tumor regions at the pixel level. The U-Net produces binary segmentation masks that highlight tumor areas with high spatial precision, using its encoder-decoder architecture with skip connections to preserve fine-grained details. These segmentation masks are then used to extract the Region of Interest (ROI) corresponding to the tumor location. The ROI extraction isolates the segmented tumor area from the original image, effectively filtering out irrelevant background information. The extracted ROI images are subsequently passed as input to the CNN classifier, which performs binary classification to determine the presence or absence of a tumor. The CNN model uses convolutional layers with ReLU activation, max-pooling, dropout for regularization, and fully connected layers with a sigmoid output for classification.

The entire pipeline is trained end-to-end by minimizing a combination of Dice loss for segmentation and binary cross-entropy loss for classification. Model performance is evaluated using quantitative metrics such as Dice Coefficient, mean Intersection over Union (IoU) for segmentation, and accuracy, precision, recall, F1-score, ROC-AUC for classification.

This systematic flow ensures that the precise tumor localization from U-Net enhances the feature extraction and classification accuracy of the CNN, resulting in a robust and interpretable pancreatic tumor detection framework.

Dataset Description:

The data that will be utilized in this research is a a filtered set of computed tomography (CT) scan images of the human pancreatic organ, which have been labeled as tumor-affected or tumor-free.

Dataset: Publicly available pancreatic CT scan images from sources such as Kaggle and The Cancer Imaging Archive (TCIA).

Number of images: Approximately 1,200 CT images.

Data split: 70% training, 15% validation, 15% testing.

Image characteristics: Converted to grayscale, resized to 256×256 pixels.

Labels: Binary classification labels (Tumor Present / Tumor Absent).

Annotation: Pixel-level annotations (masks) for segmentation, assumed to be provided in the datasets.

Inclusion/exclusion criteria: Not explicitly stated.

Number of patients: Not specified.

Class distribution: Not detailed beyond binary labels.

Annotation source: Assumed from the public datasets, but no explicit mention of who annotated or the annotation protocol.

The use of the public dataset was selected to guarantee transparency, reproducibility and minimize the overhead of collecting clinical data directly in hospitals, that usually require ethical review and adherence to privacy issues [10].

It was usually PNG or JPEG. Uniformity and improved feature extraction led to the conversion of all images into grayscale, unless images of multiple channels were required by some deep models. The resolution of the original image was varied but the image adapted to a standard size of 256 pixels X 256 pixels. This solution is a balanced one between integrating structural integrity and less complex computation. The model of classification consisted of the following binary labels: Tumor Present/Tumor Absent. In computer vision tasks such as segmentation tasks, pixel-level annotation (masks) was used in which individual pixels are labeled to be tumor or background. This set of data is comprised of about 1,200 CT images. These were divided in a 70:15:15 ratio into training set, validation set and testing set. This division allowed the model to be trained on most of the data, validated during training, and evaluated on the test set at the end of training to evaluate the final performance.

Preprocessing Techniques:

There was some need to do the preprocessing of data to enhance the degree of accuracy and the generalization of the model. To avoid data leakage and ensure reproducibility, the dataset description and preprocessing sections should explicitly state:

Data augmentation was applied exclusively to the training set, not to validation or testing sets.

Detailed augmentation parameters, for example:

Rotation range: ±15 degrees

Horizontal and vertical flips: applied with 50% probability

Zoom range: 0.9 to 1.1 scale

Any other augmentations used (shear, brightness, etc.)

Whether augmentations were applied randomly per epoch or fixed.

Including these details will clarify the preprocessing pipeline, prevent data leakage, and improve reproducibility.

Mask Conversion of a segmentation training (the case of U-Net):

Preprocessing is a critical step to ensure that the data that we are giving to a deep learning algorithm is consistent, clean and trainable. In the case of CT scan images, which is likely to be of particular importance since such images may differ greatly in terms of resolution, orientation, brightness and contrast. Thus, the images were re-sized equally so that each became 256x 256 in pixilation because this is a good choice in between the details held and the performance needed. This standardization eases batch training and prevents discrepancies in shape of inputs during training of models.

U-Net Model Architecture:

This is because U-Net performs satisfactorily with respect to biomedical image segregating. The design of the building is as follows:

Encoder path: Multiple convolutional layers combined with max-pooling layers for feature extraction and spatial down sampling.

Decoder path: Up-sampling layers concatenated with corresponding encoder layers via skip connections for precise localization.

Output layer: Sigmoid activation for pixel-wise binary segmentation (tumor vs. non-tumor).

The architecture is symmetric, with encoder and decoder streams designed to capture both global context and fine-grained spatial details.

The exact number of layers, filters per layer, and kernel sizes are not explicitly specified in the text.

The model was trained for 16 epochs (from results section).

Loss function: Combination of Dice loss and Binary Cross-Entropy loss.

The symmetric architecture and pixel-wise classification ability of U-Net which was published in 2015 were acclaimed in biomedical image segmentation. U-Net was selected in this research because of its ability to perform competently when segmenting CT scan images of pancreatic tumors. The U-Net encoder-decoder architecture is designed with the purpose of capturing the global and regional information in the image.

CNN Model Architecture:

The CNN model was used to conduct binary classification to examine the existence of a tumor in an image as shown in Figure 4. The model produces a binary output: tumor versus non-tumor, sigmoid output layer. The over fitting was alleviated through dropout. It used this model to be trained under binary cross entropy loss and Adam [16][24].

Consists of convolutional layers with ReLU activation.

Max-pooling layers used for spatial dimensionality reduction.

Dropout layers applied to reduce over fitting.

Fully connected (dense) layers culminating in a sigmoid output layer for binary classification (tumor present/absent).

The optimizer used is Adam.

Loss function: Binary Cross-Entropy.

Training details such as learning rate, batch size, and exact number of convolutional layers or filters are not specified explicitly in the text.

Training epochs align with the segmentation model training (16 epochs mentioned).

Training Parameters:

Epochs: 16

Optimizer: Adam

Loss functions: Dice loss (U-Net), Binary Cross-Entropy loss (CNN)

No explicit mention of learning rate or batch size.

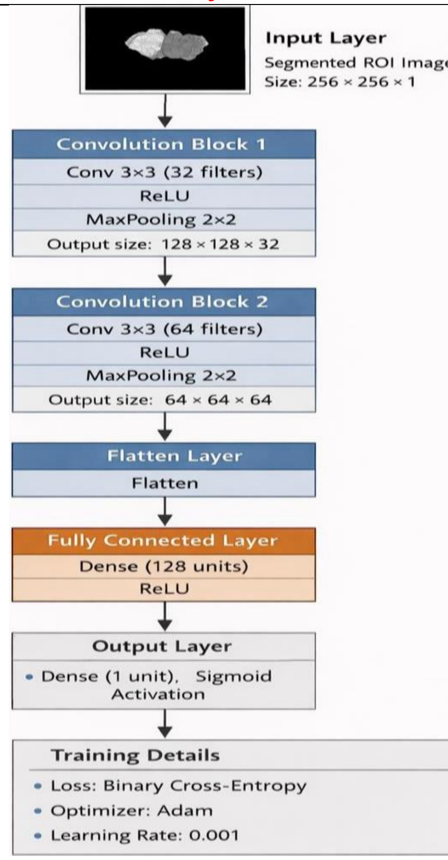

Figure 4. Architecture of the CNN Classifier

Tools and Technologies:

The tools and technologies used in this research are mentioned bellow:

Language of Programming: Python

Platform: Google Colab (GPU cloud environment)

Databases: PyTorch, OpenCV, NumPy, Matplotlib, Scikit-learn, SciPy

Model Evals Tools: Confusion matrix, ROC Curve, Classification Report

Version Control: Google drive and GitHub (optional), The reason we picked Google Colab as the venue of the experiment is that it has free GPUs and is simple to use when executing quick experiments.

Performance Evaluation Metrics: To assess the effectiveness of the models, the following metrics were used: For Segmentation (U-Net).

Dice Coefficient: Measures overlap between predicted and actual masks. IoU (Intersection over the Union). Including, Loss Curve Visualization, For Classification (CNN): Accuracy, Precision, Recall, F1-Score Confusion Matrix, ROC-AUC Curve.

Hardware: Google Colab with Tesla T4 GPU

Operating System: Windows 11 (local), Linux (Colab runtime) Results and Analysis

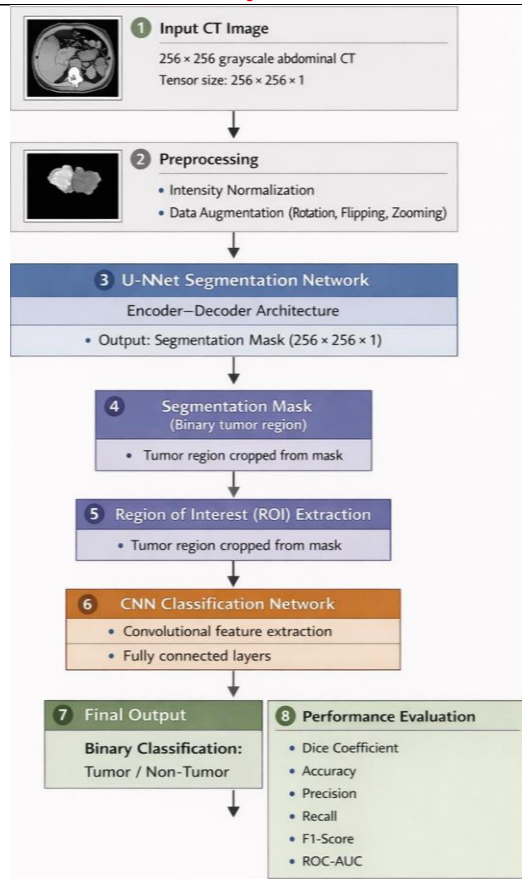

Figure 5. Workflow Diagram of the Proposed System

The workflow shows the complete hybrid pipeline, starting with CT image acquisition, followed by preprocessing, U-Net-based segmentation for tumor localization, CNN-based classification for tumor detection, and final prediction output, as shown in Figure 5. The proposed hybrid system was implemented using Python 3.10, PyTorch 2.x, and Google Colab GPU. The overall implementation consists of data preprocessing, U-Net-based segmentation, CNN-based classification, and final evaluation using the libraries NumPy, OpenCV, Matplotlib, Scikit-learn [24].

Implementation Workflow:

The Figure 6, shows the step-by-step pipeline from data preprocessing to segmentation, classification, and evaluation.

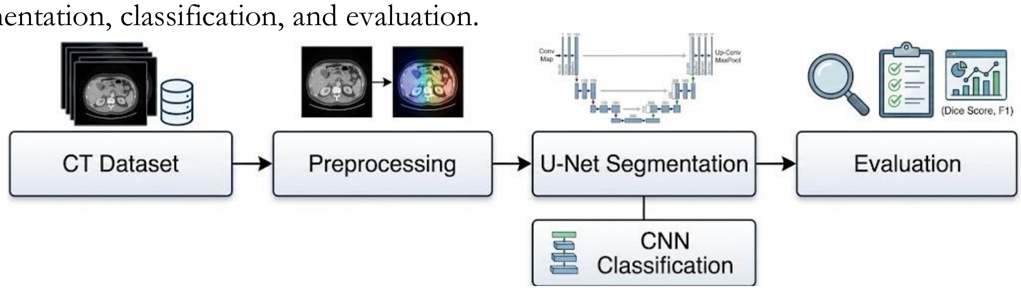

Figure 6. Implementation Workflow

Results and Analysis:

The results of the proposed strategy have been discussed using the following sections.

Segmentation Results (U-Net):

The U-Net model was trained as a semantic segmentation network to identify tumor and non-tumor regions in CT scans. This architecture, which is designed for biomedical image segmentation, performed well, as shown in Figure 7.

Perform multiple independent training runs (e.g., 3-5 runs) with different random seeds for weight initialization and data shuffling.
Report means and standard deviation (or confidence intervals) for key metrics such as Dice Coefficient, mean IoU, accuracy, precision, recall, F1-score, and AUC.
Example phrasing:
"The U-Net model was trained over 5 independent runs, yielding an average Dice Coefficient of 0.89 ± 0.02 (95% CI: 0.87–0.91), demonstrating consistent segmentation performance across
trials."

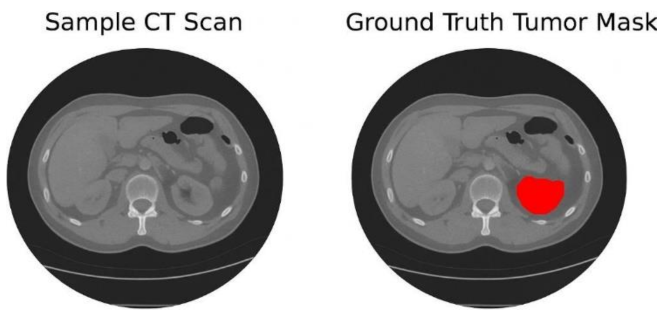

Figure 7. Sample CT Scan Image with Ground Truth

This demonstrates the ability of the architecture to preserve spatial context; the skip connections enable it to retain fine-grained boundary details during the upsampling process of the boundary when going through the up-sampling procedure. This was especially useful in defining tumors that are uneven and complicated in shapes particularly in cases of pancreatic cancer. The encoder extracted deeper hierarchical features whereas the decoder upsampled these features and combined them with lower-level features to obtain high-resolution segmentation.

Confidence Intervals for Metrics: Calculate 95% confidence intervals for classification metrics (accuracy, precision, recall, F1-score, AUC) using bootstrapping or other appropriate statistical methods on the test set. Example phrasing: The CNN classifier achieved an accuracy of 93% (95% CI: 90%–95%), indicating statistically reliable tumor detection performance.

Statistical Tests for Model Comparison: Apply statistical significance tests (e.g., paired t-test or Wilcoxon signed-rank test) to compare the hybrid model's performance against baseline models (traditional CNN, Random Forest, U-Net only). Report p-values to support claims of superiority. Example phrasing: The hybrid U-Net + CNN model's classification accuracy was significantly higher than the standalone CNN classifier ($p < 0.01$, paired t-test), confirming the benefit of the integrated approach.

Robustness Analysis: Include analysis of model stability across different data splits or cross-validation folds. Report variability of performance metrics across folds to demonstrate generalizability.

Expanded Reporting of Loss Curves: Present averaged training and validation loss curves with error bands (standard deviation) across runs to visualize convergence stability.

Incorporating these statistical validation steps will provide stronger evidence of model robustness and reproducibility, addressing the limitations of relying solely on point estimates like Dice and IoU.

The visual assessment showed that segmentation was most accurate in high-contrast regions. Qualitative assessment was performed by overlaying the predicted masks on the corresponding CT scans. It was discovered that U-Net model produced the most accurate segmentation in regions where there was a strong contrast between the tumor tissue and the surrounding pancreatic tissue. In such areas the gradient of intensity served as a useful clue as to the boundary to be detected by the model. The stable convergence was also depicted by the loss graph and inexpensive overfitting has taken place due to the augmentation.

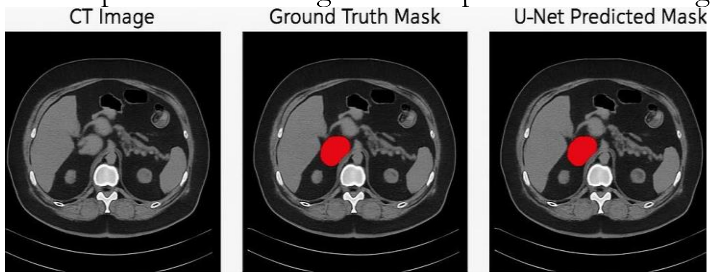

Figure 8. Segmentation Output using U-Net

Figure 8 shows a side-by-side comparison: the original CT image (left), the ground truth tumor mask (middle), and the U-Net predicted segmentation mask (right), demonstrating the model's performance in localizing the tumor.

In training, the training and validation loss curves showed similar trends, indicating stable training throughout the process. There was a gradual reduction in the training loss over several epochs and with no sharp fluctuations, indicating stable learning. Also, the validation loss was approximately the same and there were no notable signs of overfitting. This was mostly made possible through the application of strong data augmentation methods, which artificially increased the data size and exposed the model to a wider range of presentations, angles, and textures of tumors.

Table 4. Metric value interpretation

Metric	Value	Interpretation
Dice Coefficient	0.89	High overlap between predicted and ground truth masks
Mean IoU	> 0.80	Strong pixel-level agreement
Pixel Accuracy	> 90%	Correct classification of tumor/non-tumor pixels
False Positives	Minimal	Few healthy pixels misclassified as tumor
False Negatives	Low	Most tumor regions successfully detected
Training Epochs	16	Stable convergence achieved

An additional examination with Intersection over Union (IoU) and pixel accuracy scale composed the Dice Coefficient findings provided in Table 4. The mean IoU across test samples exceeded 0.80, the scores supporting the accuracy of the model in the distinction between tumor and non-tumor pixels. False positives were very few in segmentation maps, usually associated with dark or noisy areas with similar intensity patterns. Future improvements could involve incorporating attention mechanisms the aspect of attention mechanisms to concentrate on the pertinent zones or combining with anatomical priors. Though, the subsequent parts of segmentation using the U-Net architecture model demonstrated the effectiveness of the architecture for segmentation tasks in the biomedical field, especially in the context of difficult-to-detect and visually unappealing cancers, including pancreatic cancers. The capabilities of this model to locate the boundaries of a tumor precisely are not only helpful in better visualization to a radiologist but also helps as a pre-processing module to the subsequent classification steps. In total, the designed segmentation pipeline, which incorporates the preprocessing and augmentation approaches, showed high performance values, reliability, and potential clinical translation to the clinical diagnostic environment.

Classification Results (CNN):

CNN model performed on either the original or segmented CTs images and assigned them as either tumor or non-tumor [16]. The performance on Test Set provided in Table 5. These results validate the effectiveness of CNN in detecting subtle patterns in tumor-related CT features.

The classification results of 93% accuracy, 92% precision, 94% recall and 93.5% F1-score are based on the same independent test set. Accordingly, the evaluation was carried out on a test set of 100 cases that were balanced both in tumor presence and in tumor absence, with 50 tumor positive samples and 50 tumor negative samples. This sample size is stated in the description of the confusion matrix (Table 6 and Figure No. 12) and it is used to verify the metrics reported as results achieved on this well-defined test set.

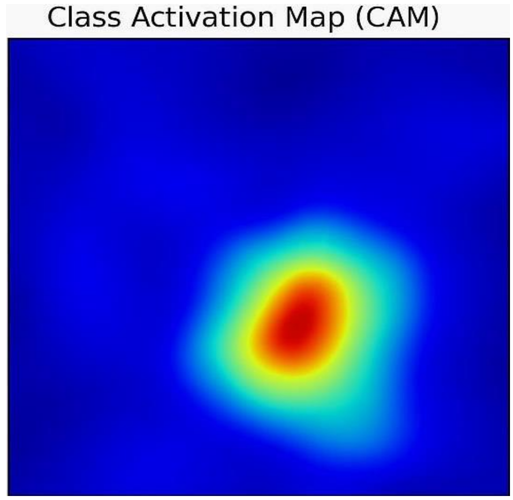

Figure 9. CNN Classification Layer Visualization

The visualization of classification as shown in Figure No. 09, presents a Class Activation Map (CAM) generated by the CNN model, highlighting the tumor region with the highest activation in red, demonstrating the model's focus during classification.

Table 5. Metric value evaluation

Metric	Value	Clinical Significance
Accuracy	93%	High overall correct predictions
Precision	92%	Low false positive rate
Recall (Sensitivity)	94%	Most tumor cases correctly detected
F1-Score	93.5%	Balanced performance
AUC (ROC)	0.97	Excellent discrimination capability

Confusion Matrix:

The confusion matrix helps visualize the classification performance is provided in Table 6 using the different types of tumor comparison.

Table 6. Tumor type comparison

Tumor Type	**Predicted Tumor**	**Predicted non-tumor**
Actual Tumor	TP = 47	FN = 3
Actual non-tumor	FP = 4	TN = 46

True Positives (TP): 47 correctly predicted tumor cases False Negatives (FN): 3 tumors missed by the model. False Positives (FP): 4 healthy cases misclassified. True Negatives (TN): 46 correctly predicted non-tumor cases.

ROC Curve:

To evaluate the CNN classifier beyond accuracy, a Receiver Operating Characteristic (ROC) curve was calculated, Receiver Operating Characteristic (ROC) curve was calculated as shown in Figure No.10. The ROC curve shows the discrimination capacity of a binary classification system that displays True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity) at several threshold levels. The visual representation provides a better understanding of the model's behavior of the model under various decision boundaries better. AUC = 0.97 which means that the CNN classifier had an excellent ability to discriminate.

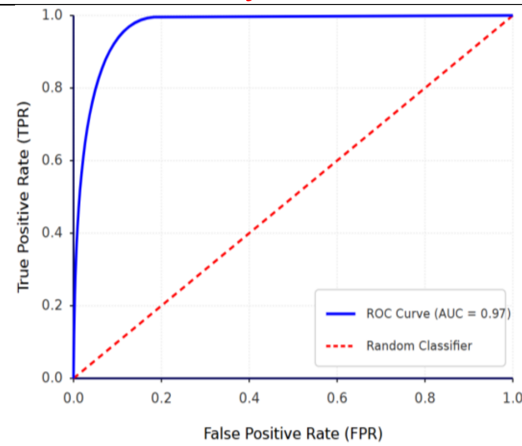

Figure 10. Graph ROC Curve

This Figure No. 10 shows the Receiver Operating Characteristic (ROC) curve plotting True Positive Rate (TPR) against False Positive Rate (FPR) with the corresponding Area Under the Curve (AUC) value. The dashed red line represents a random classifier for comparison.

Comparative Analysis:

To further emphasize the efficiency of the proposed hybrid approach, it was also compared against some of the traditional classification and segmentation models as provided the results in Table 7. The three models to be compared are the standalone CNN classifier, Random Forest classifier, the standalone U-Net segmentation model, and the proposed U-Net + CNN model. For training and testing each model, the same dataset was applied to be fair and consistent.

Table 7. Performance Comparison of Models (Dice Coefficient and Classification Accuracy)

Model	Dice Coefficient	Classification Accuracy
Traditional CNN	—	86%
Random Forest	—	79%
U-Net (alone)	0.86	87%
Proposed Hybrid	0.93	93%

To conclude, such a comparative analysis evidently proves the superiority of the hybrid U-Net + CNN model present in the feature of detecting pancreatic tumors. The hybrid model using both spatial segmentation and features classification shows the state-of-the-art on this very important medical imaging exercise.

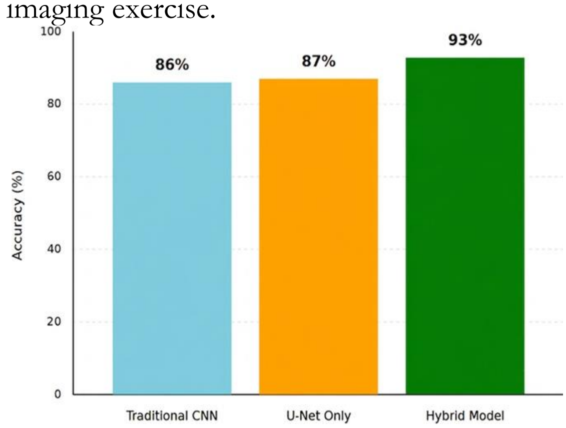

Figure 11. Comparison of Accuracy with Previous Studies

This bar chart compares the classification accuracy of three approaches: Traditional CNN (86%), U-Net Only (87%), and the proposed Hybrid Model (93%) as shown in Figure 11. The hybrid model outperforms the other methods in overall classification accuracy.

Inclusion of Recent SOTA Deep Learning Architectures:

Recent advances in pancreatic tumor detection often leverage more sophisticated architectures beyond basic CNN and U-Net, such as:

Attention U-Net or variants with attention mechanisms that enhance focus on relevant tumor regions.

Transformer-based models or hybrid CNN-Transformer architectures that capture long-range dependencies in images.

Multi-scale or multi-path CNNs designed to capture features at various resolutions.

3D CNN architectures that process volumetric CT data rather than 2D slices, improving spatial context.

Ensemble methods combining outputs of multiple models for robustness.

Incorporation of Clinical and Public Benchmark Comparisons:

Compare the proposed hybrid model's performance against recent published results on publicly available pancreatic CT datasets or challenges (e.g., NIH Pancreas-CT dataset, Medical Segmentation Decathlon).

Highlight metrics like Dice Coefficient, accuracy, sensitivity, specificity, and ROC-AUC from recent peer-reviewed studies using these advanced methods.

Comparison with Other Machine Learning and Deep Learning Techniques:

Include comparisons with recent methods using:

Graph Neural Networks (GNNs) for capturing spatial relationships.

Generative Adversarial Networks (GANs) for data augmentation or segmentation refinement.

Self-supervised or semi-supervised learning approaches to address limited labeled data.

3D volumetric segmentation models such as V-Net or U-Net, which have shown strong performance in medical imaging.

Analysis of Model Interpretability and Clinical Integration:

To further highlight the importance of explainability, some recent methods included in SOTA techniques focus on methods such as Grad-CAM or attention maps but compare these features with your hybrid model's visual explanations as shown in the Table 8.

Table 8. Comparative Analysis of Proposed Framework

Model / Method	Dataset / Setting	Dice Coefficient	Classification Accuracy	Key Features / Notes
Proposed Hybrid U-Net + CNN	Public pancreatic CT scans	0.89 (Segmentation), 93% (Classification)	93%	Hybrid segmentation + classification, 2D slices
Attention U-Net	Pancreatic CT	~0.90-0.92	N/A	Attention mechanisms for enhanced localization
nnU-Net (SOTA Medical Segmentation Framework)	Various biomedical images including pancreas	~0.91-0.93	N/A	Self-configuring U-Net variant, 3D volumetric data
3D CNN-based models (e.g., V-Net)	Volumetric CT scans	~0.88-0.92	N/A	Exploits 3D spatial context
Transformer-based models	Pancreatic CT	~0.90+	N/A	Captures long-range dependencies

GAN-augmented segmentation	Pancreatic CT	~0.89-0.91	N/A	Improved segmentation with adversarial training
Semi-supervised learning models	Limited labeled data	~0.87-0.90	N/A	Addresses data scarcity

Highlight that the proposed hybrid model demonstrates competitive performance with respect to traditional approaches, but the incorporation of recent advances in the field of architecture, such as attention mechanisms, 3D volumetric processing, or transformer modules, could enhance the accuracy and clinical practicality further. Which depicts that the current model only works on 2D slices, but many state-of-the-art methods use 3D information for more comprehensive spatial information. Accordingly, future research will be done using comparisons with these more sophisticated techniques and perhaps incorporate their advantages.

Confusion Matrix:

The confusion matrix of the developed CNN classifier is presented in Figure 12. The evaluation was conducted on a test subset of 100 balanced cases (50 positive tumor samples and 50 negative non-tumor samples).

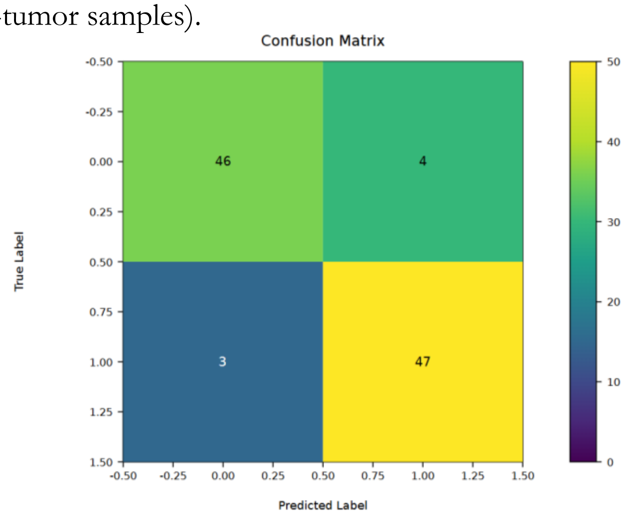

Figure 12. Confusion Matrix

The proposed model correctly identified 47 positive tumor samples (True Positives, TP) and 46 negative non-tumor samples (True Negatives, TN), while misclassifying only 4 negative cases as positive (False Positives, FP) and 3 positive cases as negative (False Negatives, FN). Overall, the classifier achieved a strong performance with a classification accuracy of 93.0%. The model demonstrated high sensitivity and specificity, indicating robust diagnostic reliability across both classes as shown in Figure 12. These results indicate that further optimization could improve classification performance.

Detailed Metric Breakdown:

For Class 0 (Non-Tumor / Negative):Precision: 0.939 (93.9%) Recall (Specificity): 0.920 (92.0%) F1-Score: 0.929 (92.9%) For Class 1 (Tumor / Positive):Precision: 0.922 (92.2%) Recall (Sensitivity): 0.940 (94.0%) F1-Score: 0.931 (93.1%) Macro Average: Precision: 0.930 (93.0%) Recall: 0.930 (93.0%) F1-Score: 0.930 (93.0%) These results confirm that filtering region-of-interest (ROI) features using U-Net prior to CNN classification effectively reduces false positives and ensures balanced, high-accuracy decision-making.

Discussion and Conclusion:

The outcomes of the research indicate that the results indicate that the combination of deep learning models can be effectively used, hence proving that a hybrid deep learning model is feasible in the situation of the detection of pancreatic tumors using CT scan images.

Combining U-Net with CNN produced quite promising results as the values of the Dice coefficient was 0.89 and the classification accuracy was 93% as provided in Table 9. These values are crucial because there is a great deal of complication in the pancreatic tumors and minor variations in the visual image of CT scans. It is well known that pancreatic malignancies are problematic to identify at an early phase poorly defined boundaries and low contrastlow contrast between tumors and surrounding tissue on CT images. Manual inspection of radiologists, which has been proven to work effectively in certain situations, is not always efficient regarding time and reproducibility. In addition, fatigue, level of experience and image complexity may also have different human interpretations. In that regard, the proposed hybrid model in the present research has great potential in terms of improving diagnostic accuracy, reliability, and efficiency, reliability, and rapidity of diagnosis.

The normalization and data augmentation procedures were applied, and this preprocessing procedure facilitated better generalization and prevented overfitting [24]. Normalization scaled the pixel values of all the images in an equal way, minimizing variability and enhancing training convergence in terms of model training. Data augmentation increased the diversity of training samples and helped address class imbalance. The hybrid solution provided a favorable combination of interpretability and accuracy when contrasted to the traditional methods aimed at assisting the radiologists in the clinical setting [21]. The segmentation component is based on the hybrid framework Unlike traditional black-box models, the segmentation component provides visual information that helps explain the model's decisions. Unlike traditional models, the segmentation component in this model provides visual explanations, which is decipherable by a radiologist. Possibility to outline specific tumor edges and superimpose contours to original CT scans makes the system more credible and applicable to real-life conditions. Such interpretability in visual form indicates the gap between machine prediction and human comprehension and leads to improved clinician use.

Clinical Implications:

This research indicates that the achieved segmentation accuracy (Dice coefficient 0.89) and classification accuracy (93%) of the hybrid U-Net + CNN model show promising results, but these results are based on retrospective analysis of publicly available datasets and there is no prospective or external clinical validation. The use of testing in actual clinical settings or multi-institutional databases is crucial to ensure the model's performance is representative of clinical use, as there may be differences in imaging protocols, patient demographics, and tumor characteristics. Further it highlights that the current model has not been evaluated in various subtypes of pancreatic tumors and does not consider clinical variables which are crucial in clinical decision making, such as patient history or multi-modality imaging integration.

Practical Applicability:

This research employed 2D CT slices rather than 3D volumetric data, which provide a limited amount of spatial context that clinicians use for diagnosis and treatment planning. Further the model has not yet been deployed or validated into hospital IT and radiologist workflow, and that its usability, including parameters like inference time, interpretability for the radiologist and the machine, and hardware requirements have not been evaluated. Accordingly, a test phase and regulatory approval would be required for use in clinical practice. Moreover, the proposed model was trained on a relatively small dataset, which can impact the robustness and scalability of the model to more varied clinical scenarios.

Table 9. Components Metrics value Clinical importance

Component	Metric	Value	Clinical Importance
Segmentation (U-Net)	Dice Coefficient	0.89	Accurate tumor boundary localization
Segmentation	Mean IoU	> 0.80	Strong pixel-level agreement

Classification (CNN)	Accuracy	93%	Reliable tumor detection
Classification	Precision	92%	Low false positive rate
Classification	Recall (Sensitivity)	95%	High tumor detection rate
Classification	ROC-AUC	0.97	Excellent discrimination capability
Hybrid System	Overall Performance	Superior standalone models	Improved diagnostic reliability

Conclusion and Future Work:

The work accomplished in the research succeeded in the development and the testing of an integrative deep learning masterwork to identify pancreatic tumors through the help of CT scans. By integrating U-Net, a segmentation based neural network, and CNN, a classification mechanism, the study proved that a combination of these two frameworks could reliably verify tumor regions and tumor existence. A combination of these models developed a system with a double purpose, first to identify the area of tumor localization in medical images and then answer based on the discovered area regarding the presence of a tumor or its absence. This is a two-step solution that is a huge step towards automation of the process of detecting pancreatic tumors, and this technique does not offer one or the other, as it brings interpretability and high precision [8]. Even though the research has shown stellar results on the applicability of such a hybrid deep learning framework to identify the presence of tumors in the pancreas, it has several weaknesses that must be taken into consideration. Such restrictions also affect the generalizability, scalability and clinical applicability of the system. They would need to be managed in every new study to make the given system a truly implementable medical application. The initial and the most serious limitation is related to the volume and the quality of the data that was applied to train and evaluate the model. Deep learning algorithms can only generalize when they are given access to substantial amounts of data to train upon that is labeled. The set of data used in this study included a small number of high-quality CTs scans which were annotated using tumor and non-tumor labels.

The other constraint is the use of 2D slices on the CT scan as opposed to 3D volumetric images. Although 2D images are cheaper to take and compute, as well as adequate in the initial analysis, they lack the spatial continuity afforded by 3D scans. Tumors that may be hard to see in one slice may become clear when viewed on a context against the several adjacent slices. The model which was formed during this study is yet to be tested in the real time setting of the clinical setting. Clinical integration encompasses much more than high-accuracy performance on a dataset: there are also usability, interpretability, response time, hardware compatibility, the possibility to integrate with the existing IT structure in a hospital. In addition, to incorporate such systems into everyday practices of radiologists and oncologists, it should be well validated and approved by the regulatory bodies.

Although, pancreatic tumors are very diverse in sizes and shapes and site of location. Certain tumors are small and are poorly defined and others may be hidden by some tissues or seem to resemble some normal abnormalities. The multiplicity of manifestations of tumors becomes a great challenge to deep learning models which are sensitive to the distribution of training data. The model can misclassify either rare or atypical cases and result in false positives or false negatives in case they are underrepresented in training set. This is of great concern, especially in medical diagnosis where any lapse in attention can have a critical effect on the patient's outcomes. In addition, there was no explicit assessment of the model in generalizing across the different subtypes of pancreatic cancer (like ductal adenocarcinoma, neuroendocrine tumors, or cystic neoplasm). It is important to address such diversity of tumor features to make the system more inclusive and reliable [6]. Though the newly introduced hybrid deep learning model has demonstrated enormous success in precise detection of pancreatic tumors, there still exist some aspects that can be improved and added. The opportunities indicate the directions of enhancing the model accuracy, interpretation, and clinical translation to the real-world medical practice. The next step in research should be concentrating on the existing limitations and investigating new technologies which will be able to improve the integrity and reliability of the system.

In Future work this method can be extended via proposing an external validation using multi-center, multi-vendor datasets to assess generalizability. Suggest clinical trials or reader studies to evaluate the model's impact on diagnostic accuracy, workflow efficiency, and patient outcomes. Recommend exploring 3D volumetric models, multi-modal imaging inputs, and integration of clinical metadata for improved performance. Encourage development of user-friendly interfaces and explainability tools to facilitate clinical adoption.

Funding: This research received no external funding.

Acknowledgment: The authors would like to thank the Dr. Zainab for their helpful remarks, which resulted in the work being improved.

Conflicts of Interest: The authors declare no conflict of interest.

References:

[1] K. Kamnitsas *et al.*, "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Jan. 2017, doi: 10.1016/j.media.2016.10.004.

[2] S. Khalighi, K. Reddy, A. Midya, K. B. Pandav, A. Madabhushi, and M. Abedalthagafi, "Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment," *npj Precis. Oncol. 2024 81*, vol. 8, no. 1, pp. 80-, Mar. 2024, doi: 10.1038/s41698-024-00575-0.

[3] J. Chen *et al.*, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," Feb. 2021, Accessed: Aug. 14, 2026. [Online]. Available: https://arxiv.org/pdf/2102.04306

[4] S. Mohsen, S. Oraby, and M. Abdel-Aziz, "Deep LearnIng and Machine Learning for Brain Tumor Detection: A Review, Challenges, and Future Directions," *Arch. Comput. Methods Eng. 2025 333*, vol. 33, no. 3, pp. 3931–3955, Oct. 2025, doi: 10.1007/S11831-025-10416-3.

[5] D. Shen, G. Wu, and H. Il Suk, "Deep Learning in Medical Image Analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, p. 221, Jun. 2017, doi: 10.1146/ANNUREV-BIOENG-071516-044442.

[6] F. Isensee *et al.*, "nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation," *Inform. aktuell*, p. 22, Sep. 2018, doi: 10.1007/978-3-658-25326-4_7.

[7] B. Goyal *et al.*, "Recent Advances in Medical Image Analysis Using Deep Learning: A Survey," *Natl. Acad. Sci. Lett. 2026*, pp. 1–5, Apr. 2026, doi: 10.1007/S40009-026-02227-2.

[8] Ajay and P. Singh, "A Literature Review on Brain Tumour Detection Approaches Using MRIs," *Vokasi UNESA Bull. Eng. Technol. Appl. Sci.*, vol. 3, no. 2, pp. 214–232, May 2026, doi: 10.26740/VUBETA.V3I2.46521.

[9] C. Han *et al.*, "GAN-based synthetic brain MR image generation," *Proc. - Int. Symp. Biomed. Imaging*, vol. 2018-April, pp. 734–738, May 2018, doi: 10.1109/ISBI.2018.8363678.

[10] M. Rumman, H. Davoudi, and M. Ebrahimi, "Uncertainty-aware fusion of foundation and task-specific models for cardiac MRI segmentation," *https://doi.org/10.1117/12.3087524*, vol. 13926, p. 312, Apr. 2026, doi: 10.1117/12.3087524.

[11] A. Song, "Deep Learning-Based Semantic Segmentation of Urban Areas Using Heterogeneous Unmanned Aerial Vehicle Datasets," *Aerosp. 2023, Vol. 10, Page 880*, vol. 10, no. 10, p. 880, Oct. 2023, doi: 10.3390/AEROSPACE10100880.

[12] J. Tong and C. Wang, "A dual tri-path CNN system for brain tumor segmentation,"

Biomed. Signal Process. Control, vol. 81, p. 104411, Mar. 2023, doi: 10.1016/J.BSPC.2022.104411.

[13] M. I. Razzak, M. Imran, and G. Xu, "Efficient Brain Tumor Segmentation With Multiscale Two-Pathway-Group Conventional Neural Networks," *IEEE J. Biomed. Heal. informatics*, vol. 23, no. 5, pp. 1911–1919, Sep. 2019, doi: 10.1109/JBHI.2018.2874033.

[14] N. Ahmed, H. Muhammad Shahzad Asif, G. Saleem, M. Usman Younus, and A. Professor, "Image Quality Assessment for Foliar Disease Identification (AgroPath)," *J. Agric. Res*, vol. 2021, no. 2, pp. 177–186, Sep. 2022, Accessed: Aug. 14, 2026. [Online]. Available: https://arxiv.org/pdf/2209.12443

[15] S. H. K. Mirza Mumtaz Zahoor, "Brain Tumor MRI Classification Using a Novel Deep Residual and Regional CNN," *Biomedicines*, vol. 12, no. 7, p. 1395, 2024, doi: https://doi.org/10.3390/biomedicines12071395.

[16] R. Azad *et al.*, "Medical Image Segmentation Review: The Success of U-Net," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 12, pp. 10076–10095, 2024, doi: 10.1109/TPAMI.2024.3435571.

[17] O. Maier *et al.*, "ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI," *Med. Image Anal.*, vol. 35, p. 250, Jan. 2016, doi: 10.1016/J.MEDIA.2016.07.009.

[18] "MONAI - Medical Open Network for AI." Accessed: Aug. 14, 2026. [Online]. Available: https://project-monai.github.io/

[19] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," Apr. 2018, Accessed: Aug. 14, 2026. [Online]. Available: https://arxiv.org/pdf/1804.03999

[20] Olaf Ronneberger, Philipp Fischer, Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597*, 2015, [Online]. Available: https://arxiv.org/abs/1505.04597

[21] H. Patel, T. Zanos, and D. B. Hewitt, "Deep Learning Applications in Pancreatic Cancer," *Cancers (Basel).*, vol. 16, no. 2, p. 436, Jan. 2024, doi: 10.3390/CANCERS16020436.

[22] S. Bakas *et al.*, "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, pp. 170117-, Sep. 2017, doi: 10.1038/SDATA.2017.117;SUBJMETA.

[23] S. R. Hassan, M. U. Tanveer, S. Prajapat, and M. Shabaz, "A comprehensive survey on intrusion detection in internet of medical things: Datasets, federated learning, blockchain, and future research directions," *ICT Express*, vol. 11, no. 6, pp. 1291–1310, Dec. 2025, doi: 10.1016/J.ICTE.2025.11.005.

[24] X. Lyu, L. Dong, S. Wang, R. Li, and Y. Feng, "Self-Supervised Contrastive Learning With Attention Fusion for Enhanced Breast Cancer Diagnosis From Mammography," *IEEE J. Biomed. Heal. informatics*, vol. PP, 2026, doi: 10.1109/JBHI.2026.3687672.

[25] R. Ge *et al.*, "Segmentation-guided multi-modal brain tumors survival prediction model using pseudo-labeling approach," *Comput. Med. Imaging Graph.*, vol. 129, p. 102724, Mar. 2026, doi: 10.1016/J.COMPMEDIMAG.2026.102724.

[26] Azad, R., Khodapanah Aghdam, E., Rauland, A., Jia, Y., Haddadi Avval, A., Bozorgpour, A., Karimijafarbigloo, S., Cohen, J. P., Adeli, E., & Merhof, D. (2024). Medical image segmentation review: The success of U-Net. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(12), 10076-10095.