VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2026, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 14, Number 1, January-March 2026 *pp: 73-88*

Lesion-Aware Graph-Augmented Deep Framework for Pancreatic Cancer Detection from CT

V. Gokula Krishnan[1*], **Arvind Kumar Tiwari**[2], **Shanker M.C.**[3], **Pinagadi Venkateswararao**[4], **Sathyamoorthy**[5], **Janaki**[6]

[1]Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya, Malaysia; Department of Computer Science and Engineering, Easwari Engineering College, Chennai, Tamil Nadu, India; [2]Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya, Malaysia; Department of Computer Science and Engineering, Kamla Nehru Institute of Technology, Sultanpur, India; [3]Department of Biomedical Engineering, VelTech MultiTech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, Tamil Nadu, India; [4]Department of Computer Science and Engineering, CVR College of Engineering, Hyderabad, Telangana, India; [5]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Avadi, Tamil Nadu, India; [6]Department of Mathematics, Panimalar Engineering College, Chennai, Tamil Nadu, India

Keywords: *Pancreatic cancer; Temperature scaling; Lightweight graph head; Lesion-aware feature fusion; post-hoc probability calibration.*

Journal Info:
Submitted:
January 14, 2026
Accepted:
February 05, 2026
Published:
February 16, 2026

Abstract Because pancreatic cancer lesions are tiny, iso-attenuating, and partially obscured by surrounding arteries and ducts, early detection is still challenging. The end-to-end pipeline proposed in this study includes lesion-aware feature fusion, hybrid U-Net tumour masking, radiomic texture mining, physics-aware pre-processing, a MobileViT backbone with a lightweight graph head, and post-hoc probability calibration. It makes use of the 1,418 high-resolution DICOM slices (typically 512×512) from the Pancreatic CT Images collection. Slice-level performance is 0.942 for accuracy, 0.936 for macro-F1, 0.972 for AUROC, and 0.969 for AUPRC, according to the held-out test set. The accuracy increases to 0.960, the macro-F1 increases to 0.957, the AUROC increases to 0.986, and the AUPRC increases to 0.983 when the data is aggregated at the patient level. The Expected Calibration Error decreases from 0.031 to 0.009 when the temperature is scaled, improving the accuracy of risk assessments. With a Dice score of 0.82±0.09 (median 0.84; HD95 7.3 mm), the segmentation branch is likely to get worse as the tumours get smaller than a centimetre. Lesion-aware attention yields the greatest gains in calibration and discrimination, according to ablations (ΔAUROC -1.4 points, ΔECE +0.012 upon removal). There are also some advantages to radiomics (-0.3 points) and graph (-0.7 points). The site shift is minimal (worst site: AUROC 0.958), and the AUROC decreases steadily under robustness stress (motion S3: 0.940; noise S3: 0.945; bias-field S3: 0.949). The sensitivity and specificity of a Youden-optimal threshold are 0.927 and 0.935, respectively. The sensitivity and specificity of a screening operating point are 0.965 and 0.881, respectively. With throughput (about 28 ms each slice; 1.48 s per study) and 98.5% coverage following quality gates, the system can be used in real-world scenarios. According to the findings, pancreatic cancer can be detected on CT images using calibrated, lesion-aware, graph-augmented fusion that is efficient and low in processing overhead.

***Correspondence author email address:** gokul_kris143@yahoo.com
DOI: 10.21015/vtse.v14i1.2322

1 Introduction

Pancreatic cancer is one of the most deadly types of cancer, and only a small number of people survive for five years because it is often diagnosed late [1]. Cross-sectional imaging, especially contrast-enhanced computed tomography (CT), is the main tool used in clinics for staging and planning treatment [2]. However, early lesions are usually small, poorly contrasted, or iso-attenuating compared to pancreas parenchyma, which makes it very hard for computers to find them [3]. So, effective algorithms must be able to find small changes in texture, stay within anatomical boundaries, and work with different scanners, institutions, and acquisition protocols [4].

Recent literature emphasizes additional strategies for enhancing detection. Multi-stage pipelines employing edge-preserving denoising [5], traditional segmentation, and CNN classifiers exhibit significant accuracy on curated datasets, indicating that careful pre-processing can enhance lesion salience [6]. Graph-informed learning and differentiable architecture search have shown that understanding how important landmarks are related can help make decisions when the contrast between pixels is low [7]. Liquid biopsy and micro biome studies show that non-imaging biomarkers can provide useful information. Methodologically sound research on choosing and calibrating operating points [8] shows that a high AUROC is not enough; probabilities must be reliable to set clinically meaningful thresholds [9].

Based on these insights, this study focuses on CT-only detection while structuring the pipeline around four enduring challenges [10]: (i) low contrast at lesion boundaries, (ii) small object size that makes it hard for down-sampling networks to work, (iii) dataset shift across sites and artefacts (noise, motion, bias-fields), and (iv) probability miscalibration that makes thresholding harder. The publicly available Pancreatic CT Images dataset [11] (1,418 high-resolution DICOM slices, generally 512×512, containing normal and tumour cases) is a good way to test these goals.

The suggested framework is a deep model that is aware of lesions and has explicit calibration. First, preprocessing windows that know about physics Hounsfield units to soft-tissue ranges, z-scores intensities, and applies anisotropic diffusion to suppress noise without blurring edges—because the where and how of boundary preservation determine downstream segmentation quality. Second, a hybrid U-Net makes tumour probability maps that are more accurate in terms of shape and size. These masks give pixel-level evidence, let you get shape and size descriptors, and most importantly, they help you focus on the most likely places where lesions are later on. Third, the refined mask's radiomic texture features (like GLCM contrast, homogeneity, energy, and entropy) show the "what" of micro-texture that high-stride encoders might miss. Fourth, a MobileViT backbone processes the slice, and a lightweight graph head sends information across Harris-corner landmarks in the pancreas area. This models the "how" of relational patterns (duct–vessel topology) that tumours mess up. Fifth, a lesion-aware fusion gate learns which source to trust by combining deep semantic features, graph descriptors, and radiomics with spatial attention from the segmentation map. Lastly, the classifier's logits are temperature-scaled on a validation split to make sure that confidence matches empirical accuracy. This lets Youden's J or utility choose the right threshold.

Evaluation adheres to best practices: discrimination (Accuracy, Macro-F1, AUROC, AUPRC) at both slice and patient levels; segmentation quality (Dice, HD95) with size stratification; calibration (Brier score, ECE, reliability diagrams) pre/post scaling; robustness under controlled noise/motion/bias-field perturbations and across sites; and deploy ability via efficiency–accuracy. Ablation tests separate the small effects of lesion-attention, graph reasoning, radiomics, and loss components (Dice, focal).

With the development of transformer, based medical image analysis contributing significant breakthroughs, the existing methods to detect pancreatic cancer through images mainly limit their focus on either a pixel, level segmentation or a global classification and thus, they tend to overlook the probability calibration and relational anatomical reasoning. Furthermore, few studies evaluate confidence reliability or robustness to real, world artefacts such as motion and scanner bias, even though they usually report very high discrimination metrics. Our research fills these holes by presenting a lesion, aware, a graph, augmented framework together with an explicit post, hoc recalibration, thus not only accurate but also reliable pancreatic cancer detection

can be performed even in very difficult clinical situations.

Major contributions presented by this work are as follows:

- A lesion, aware deep framework that cleverly combines segmentation, guided attention with classification for better pancreatic tumour localization.
- A light graph, augmented representation through anatomical landmarks to capture relational tissue structure.
- A calibrated decision pipeline employing temperature scaling to enhance probability trustworthiness for clinical thresholding.
- Comprehensive tests of robustness, component analyses, and operating point analyses showing the method's resistance to artefacts and its stability across different domains.

2 Related works

Nadeem et al. [12] delineate a four-stage computer-aided diagnosis pipeline for pancreatic cancer. In pre-processing, each input image is resized to 227 × 227, changed from RGB to grey scale, and denoised with anisotropic diffusion filtering to get rid of noise while keeping edges. Segmentation starts with thresholding to make a binary mask, then Sobel filtering to make boundaries stand out, and finally watershed segmentation to mark tumour areas. A U-Net variant is also used for comparison. Morphological operations improve the shape, and then grey-level co-occurrence matrices are used to figure out texture descriptors by measuring how intensity pairs change in space. In detection, the connected components are labelled, and the densest area is chosen to find the tumour. This gives a detection accuracy of 99.64%. A smaller version of Alex Net with 11 layers is used in the classification stage to sort cases into normal or pancreatic tumour and then further sort tumours into benign, pre-malignant, or malignant. The reported performance shows that the system is both effective and efficient, with a classification accuracy of 98.72%, an AUC of 0.9979, and an average runtime of 1.51 seconds per case.

In addition to pixel-space pipelines, Alaca [13] suggests a graph-informed representation that works with DARTS-optimized MobileViT models. Harris Corner Detection turns pancreatic CT images into graphs that show important structures and how they are related in space. MobileViT backbones process these graphs, and their architectures are improved through differentiable architecture search, which allows for task-specific capacity. To make the most of learned embedding's, classical machine-learning classifiers like K-Nearest Neighbours, Support Vector Machines, Random Forest, and XGBoost are tested on the deep features. MobileViTv2_150 and MobileViTv2_200 get up to 97.33% accuracy and a 96.25% F1-score, which is better than regular CNNs and Vision Transformers. The formulation emphasizes scalability and proposes a viable approach to resilient, resource-conscious deployment in medical imaging.

In addition to imaging, Lee et al. [14] examine blood-derived extracellular vesicle micro biome signatures for diagnostic modelling. The dataset comprises 38 pancreatic cancer patients and 51 healthy controls. Stepwise regression is used to choose markers at the phylum and genus levels, and logistic regression, random forest, SVM, and deep neural networks are used to test the resulting panels. A three-marker DNN (Verrucomicrobia, Actinobacteria, Proteobacteria) gets a test AUC of 0.959 at the phylum level. A DNN with 11 LASSO-selected taxa (e.g., Ruminococcaceae UCG-013, Akkermansia, Pseudomonas) achieves a test AUC of 0.961 at the genus level. The results highlight the potential of minimally invasive micro biome-based markers to enhance imaging and clinical data in the detection of pancreatic cancer.

Almisned et al. [15] examine urine-based clinical characteristics from 590 subjects employing six standard classifiers and an ensemble voting methodology. They subsequently present hybrid variants for the early detection of pancreatic cancer, incorporating model interpretability through SHAP. The voting ensemble is better than single models, with an accuracy of 96.61% and a precision of 98.72%. A voting-classifier-plus-random-forest setup is the best for hybrids when it comes to discrimination (AUC = 99.05%; 95% CI: 0.93–1.00). SHAP analyses find the most important features that help predict pancreatic cancer, which helps make decisions that are clear and easy for doctors to understand.

Finally, Dzemyda et al. [16] propose a deep-learning framework that amalgamates diverse data sources and underscores meticulous operating-point selection to achieve a balance between sensitivity and specificity.

The Youden index, the (0, 1)-criterion, the Matthews correlation coefficient, the F1-score, the positive/negative likelihood ratios, the balanced accuracy, and the geometric mean are all used to judge. This metric-aware selection aims for dependable screening efficacy and indicates the potential applicability of the methodology for the early detection of additional malignancies.

These studies collectively demonstrate complementary methodologies—edge-preserving image analysis, graph-structured representations with architectural search, micro biome-based liquid biomarkers, ensemble and hybrid tabular-clinical modelling with SHAP interpretability, and metric-guided threshold optimization [17]. Reported outcomes (e.g., imaging AUC $\approx$ 0.998 with $\sim$1.5 s latency; graph-MobileViT accuracy $\approx$ 97%; micro biome DNN AUC $\approx$ 0.96; urine-clinical AUC $\approx$ 0.99) demonstrate that multimodal, meticulously designed pipelines can achieve elevated diagnostic accuracy while ensuring practical efficiency and transparency, thereby reinforcing the argument for integrated, early pancreatic cancer diagnosis.

Medical image analysis has recently seen a rise in the use of Vision Transformer (ViT) architectures impact the field as these have been applied to address the limited global context modeling of traditional CNNs. Some of the Hybrid CNNViT models like Swin Transformer and MobileViT have shown to be better than the existing models in CT and MRI, based tumor detection by integrating local texture sensitivity [18] with long, range dependency modeling [19].

Meanwhile, Graph Neural Networks (GNNs) have also been drawn into focus for use in cancer imaging where they have been applied to structure modelling and identifying relational flows among various lesion regions [20]. A few of the recent studies have opted to use graphs as a tool to represent the different ways in which radiomic, spatial, and inter, lesion contexts can relate to each other. This has resulted not only in more accurate tumor classification but also in better risk stratification [21]. Besides, graph, based methods have been proven conducive in mitigating the negative effects of segmentation noise and patient variability in cancer analysis from CT scans [22]. Such research outcomes serve as a strong driver for the combination of transformer, based feature extraction and graph learning methods to achieve dependable cancer detection.

3. Proposed Framework for Pancreatic Cancer Detection from CT

This part explains a mathematically based, end-to-end framework that is made for the Pancreatic CT Images dataset, which has 1,418 high-resolution DICOM images that are usually 512×512 pixels. Figure 1 (the sample montage you gave) is what inspired the design: Axial abdominal slices exhibit nuanced textural and boundary indicators that distinguish normal pancreatic tissue from neoplastic cases. The framework combines physics-aware pre-processing, organ/lesion segmentation, hand-crafted radiomics, graph-augmented deep vision encoders, multi-branch feature fusion, calibrated classification, and principled operating-point selection. The theory below directly and indirectly explains what each block computes, how it is made, why it is needed, and where it fits in the pipeline. Figure 1 shows how the proposed model works.

3.1 Problem Definition and Data Space

What is being predicted? A binary diagnostic label for each CT slice or study (normal pancreas vs. pancreatic tumour), and where applicable, a pixel-level mask of the lesion to support explain ability.

Let the dataset be $\mathcal{D} = \{(X_n, y_n)\}_{n=1}^{N}$, where $X_n \in \mathbb{R}^{H\times W}$ is an axial CT slice (post-windowing; $H = W = 512$ typically) and $y_n \in \{0, 1\}$ is the image-level label (0: normal, 1: tumor). When manual masks are available or pseudo-labels are created (Sec. 3.3), denote $S_n \in \{0, 1\}^{H\times W}$ the tumor mask.

To robustly exploit both global organ context and local lesion micro-texture, how the model operates:

- A segmentation branch predicts $\hat{S}$ to focus representation on tumour-bearing regions (supervising attention and providing morphometric).
- A classification branch predicts $p(y = 1 \mid X)$ from fused deep and radiomic features.
- A calibration layer transforms logits to reliable probabilities for threshold-based clinical decisions.

Why this formulation: pancreatic lesions are small, heterogeneous, and often low-contrast; coupling segmentation with classification (multi-task learning)

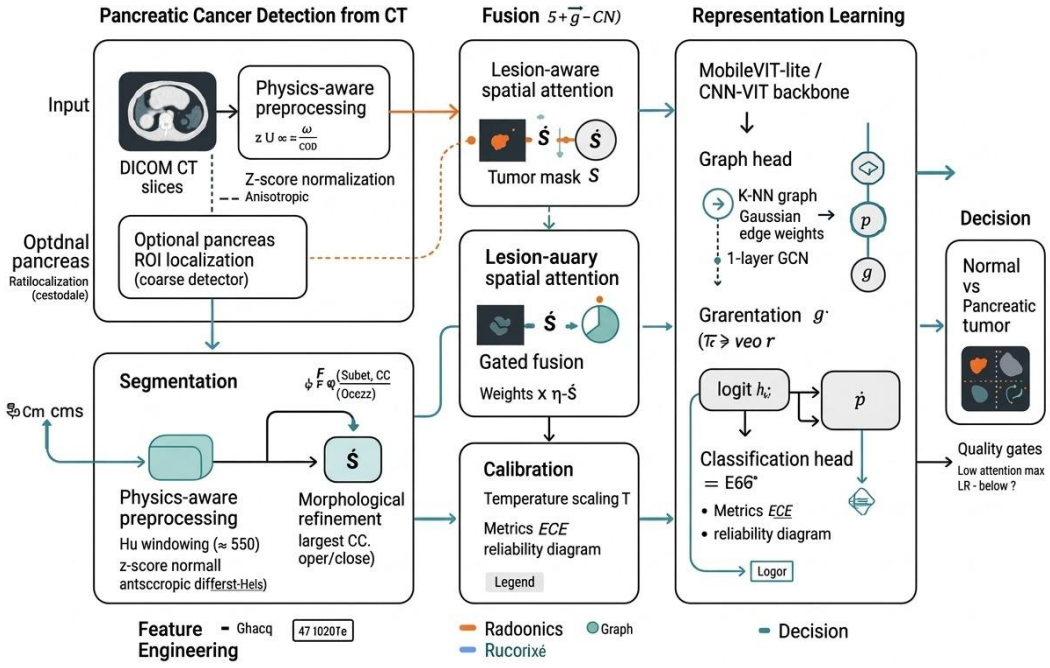

Figure 1. Workflow of the proposed model

amplifies lesion salience while preserving holistic abdominal context.

3.2 DICOM Harmonization and Physics-Aware Pre-processing

CT intensities are recorded as Hounsfield Units (HU). Where intensity harmonization is required (multi-site acquisitions), windowing and normalization reduce inter-scanner variability and highlight soft tissue contrast.

(1) Linear windowing to soft-tissue range

$$I(x) = \text{clip}\left(\frac{\text{HU}(x) - \mu_w}{w/2}, -1, 1\right) \tag{1}$$

Where μ_w is the window center (e.g., 50 HU) and w the width (e.g., 350 HU). What this does: maps HU to $[-1, 1]$ emphasizing pancreas–lesion contrast; why: suppresses bone/air extremes that confound learning.

(2) Z-score normalization across slice

$$\tilde{I}(x) = \frac{I(x) - \mu_I}{\sigma_I + \epsilon} \tag{2}$$

With μ_I, σ_I the image mean and standard deviation; ϵ prevents division by zero. How: standardizes dynamic range; where: per-slice to accommodate abdominal coverage differences.

(3) Edge-preserving anisotropic diffusion denoising

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(c(\|\nabla u\|)\nabla u\right), \quad c(s) = \exp\left(-(s/\kappa)^2\right) \tag{3}$$

Here u is the evolving image, $c(\cdot)$ the conduction function, κ a contrast parameter. What: smooths noise while respecting organ edges; why: lesions manifest as faint gradients—over-smoothing destroys them, so c attenuates diffusion at strong boundaries.

(4) Geometric resizing and standardization

$$X = \mathcal{R}_\alpha(\tilde{I}), \quad \alpha : (H, W) \to (h, w) \tag{4}$$

Where $\mathcal{R}_\alpha$ resizes to $h \times w$ (e.g., 256×256 or 320×320) for batch training. Where this matters: GPU memory and consistent receptive field across images.

Data augmentation (why/what): random affine jitters, mild elastic deformations, and contrast jitters simulate anatomical and scanner variability while preserving lesion topology.

3.3 Pancreas/Tumour Segmentation (Hybrid U-Net with Classical Priors)

What this branch does: produces $\hat{S} \in [0, 1]^{h \times w}$, a tumor probability map; why: constrains attention and yields

morphometrics (area, eccentricity) for the classifier; how: a U-Net enhanced by classical edge cues.

U-Net forward mapping

$$\hat{S} = \sigma\left(f_\theta^{\text{U-Net}}(X)\right) \tag{5}$$

With σ the logistic function and $f_\theta^{\text{U-Net}}$ an encoder–decoder with skip connections.

To bias the network to true boundaries, compute Sobel gradients G_x, G_y and magnitude $G = \sqrt{G_x^2 + G_y^2}$. Inject an edge channel $E = \phi(G)$ (normalized) by concatenation: $X' = [X; E]$. Where: input stage; why: explicit edges stabilize boundary learning in low-contrast tissue.

Dice similarity and Dice loss

$$\text{Dice}(S, \hat{S}) = \frac{2\sum_x S(x)\hat{S}(x)}{\sum_x S(x) + \sum_x \hat{S}(x) + \epsilon}, \quad \mathcal{L}_{\text{Dice}} = 1 - \text{Dice}(S, \hat{S}) \tag{6}$$

What: overlap measure robust to class imbalance; why: tumour pixels are scarce.

The segmentation masks employed in this research are only weakly generated and thus are expected to have little boundary inaccuracies especially for small lesions. The errors in the segmentation due to noise can potentially change the shape sensitive features, however, their effect on the classification is limited by the use of aggregated radiomic descriptors and graph, based relational features. These different representations depend on local and structural information rather than on the exact boundary precision, hence, they give consistent classification confidence even when the segmentation quality is not very good.

Hybrid segmentation loss

$$\mathcal{L}_{\text{seg}} = \mathcal{L}_{\text{Dice}} + \lambda_{\text{BCE}} \cdot \text{BCE}(S, \hat{S}) \tag{7}$$

With λ_{BCE} moderating pixel-wise calibration and recall.

Morphological refinement (how/where): after thresholding $\hat{S}$ at τ_s, apply opening/closing to remove specks and fill voids; the largest connected component is retained (pancreas-region prior). This yields $\tilde{S}$ for radiomics (Sec. 3.4) and for attention weighting (Sec. 3.6).

3.4 Texture and Structural Radiomics

Deep encoders excel at semantics; radiomics capture what hand-crafted micro-textures describe (fibrosis, necrosis granularity). To compute grey-level co-occurrence matrix (GLCM) features inside $\tilde{S}$ because where lesions exist, second-order statistics change.

Radiomic features by their very nature are highly sensitive to the quality of segmentation, especially to small boundary changes. In order to reduce this influence, the features of the suggested framework are combined over several spatial statistics instead of depending solely on pixel, level accuracy. Moreover, the segmentation masks are further refined by post, processing methods that remove small isolated noisy regions. These architectural decisions increase the robustness of radiomic descriptors to slight segmentation changes and help maintain consistent classification performance at the output.

GLCM definition

$$P(i, j; d, \theta) = \sum_{\mathbf{x} \in \Omega_{\tilde{S}}} \mathbf{1}[X(\mathbf{x}) = i \wedge X(\mathbf{x} + \Delta_{d,\theta}) = j] \tag{8}$$

Where grey levels $i, j \in \{0, \ldots, L-1\}$, $\Delta_{d,\theta}$ encodes offset d at orientation θ, and $\Omega_{\tilde{S}}$ indexes pixels inside the refined mask. Normalize by $p_{ij} = P(i, j)/\sum_{i,j} P(i, j)$.

From p_{ij}, compute key descriptors (why these: sensitivity to coarse vs. fine textures and run-lengths):

$$\text{Contrast} = \sum_{i,j} (i-j)^2 p_{ij}, \quad \text{Homogeneity} = \sum_{i,j} \frac{p_{ij}}{1 + |i-j|}, \tag{9}$$

$$\text{Energy} = \sum_{i,j} p_{ij}^2, \quad \text{Entropy} = -\sum_{i,j} p_{ij} \log(p_{ij} + \epsilon) \tag{10}$$

Aggregate across $\theta \in \{0°, 45°, 90°, 135°\}$ and $d \in \{1, 2, 3\}$ to obtain a radiomic vector $\mathbf{r} \in \mathbb{R}^m$ (e.g., $m = 16$–32). How they are used: concatenated with deep features to improve discriminability on small lesions that may be partially invisible after down-sampling.

3.5 Graph-Augmented Visual Encoder

Why bring graphs? Tumours perturb local anatomical junctions (duct–vessel interfaces). Corner-like landmarks and their spatial relations complement pixel grids. What to build: a sparse graph over Harris corners within $\tilde{S}$ (or pancreas ROI), then a lightweight GCN layer to encode relational patterns, fused with a MobileViT/ConvNeXt-style CNN–ViT backbone.

MobileViT was chosen as the main feature extraction backbone because it can mix the convolutional inductive bias with the transformer, based global context modeling very efficiently from a computational point of

view. In pancreatic CT imaging, the lesions are typically small, low, contrast, and difficult to see as they are surrounded by very heterogeneous tissues, thus both local texture sensitivity and long, range contextual awareness become vital. Against EfficientNet, which mostly depend on convolutional hierarchies and the global receptive field aggregation may be limited, MobileViT facilitates the modeling of long, range dependencies effectively without a significant increase in the parameter count. In the same way, ConvNeXt, tiny, though it provides better representational capability, brings about a greater computational overhead and memory usage, which might be a constraint for the scalability of high, resolution CT. Thus, MobileViT is a good compromise between the two extremes of accuracy and efficiency. It is a perfect fit for picking up the subtle features of pancreatic lesions and at the same time being practical in terms of clinical deployment and operating in resource, limited environments.

Graph construction (how/where):

- Detect K Harris corners $\{\mathbf{v}_k\}_{k=1}^{K}$ inside $\tilde{S}$.
- Connect edges to k's k_{NN} spatial neighbors with weights decaying by Euclidean distance.

The graph structure is not based on fixed hyper parameters but is built adaptively. First, Harris corner points are detected within each lesion region, and the number of graph nodes is decided dynamically from the local landmark density. When forming the neighborhood, the connectivity radius is modified depending on the spatial distribution of the detected landmarks so that the graph sparsity is consistent for lesions of different size and shape. This adaptive strategy increases the robustness of the method to the variation in lesion scale and makes it easier to reproduce the results from different CT scans.

Weighted adjacency and degree

$$A_{ij} = \exp\left(-\frac{\|\mathbf{v}_i - \mathbf{v}_j\|_2^2}{2\sigma^2}\right)\mathbf{1}_{(\{j\in\mathcal{N}_i\})}, \quad D_{ii} = \sum_j A_{ij} \tag{11}$$

One-step GCN propagation

$$H^{(1)} = \phi\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(0)}W\right), \quad \tilde{A} = A + I \tag{12}$$

Where $H^{(0)}$ stacks per-node descriptors (local gradient magnitude, intensity, $\tilde{S}$ proximity), W are learnable weights, ϕ is ReLU, and $\tilde{D}$ is degree of $\tilde{A}$. Pool node embeddings (mean/max) to produce a graph descriptor $\mathbf{g} \in \mathbb{R}^p$.

CNN–ViT path (what/how): A MobileViT-like encoder processes X (or X' with edge channel). Its hybrid block alternates convolution (locality) with transformer attention (long-range pancreas context) [23].

Multi-head self-attention (per block)

$$\mathrm{Attn}(Q,K,V) = \mathrm{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V, \qquad Q = XW_Q, \quad K = XW_K, \quad V = XW_V. \tag{13}$$

Where X is the token sequence from unfolded patches, W_Q, W_K, W_V are projection matrices, and d_k is the key dimension. Why: long-range dependencies capture organ-level cues (e.g., peri-pancreatic fat stranding) that correlate with disease.

The CNN–ViT produces a feature vector $\mathbf{z} \in \mathbb{R}^q$ via global average pooling over the final feature map.

3.6 Multi-Branch Fusion, Attention, and Classification

What gets fused: deep semantic $\mathbf{z}$, graph relational $\mathbf{g}$, and radiomics $\mathbf{r}$. Where attention is applied: lesion-aware gating guided by $\tilde{S}$. Why: signal-to-noise improves when features are weighted toward the tumour locus yet retain context.

Lesion-aware spatial attention

$$\mathbf{A}_l = \mathrm{norm}(\mathrm{avgpool}(\tilde{S}; k)), \quad \mathbf{Z}_l = \mathbf{A}_l \odot \mathbf{Z}_{\mathrm{feat}} \tag{14}$$

Where $\mathbf{Z}_{\mathrm{feat}}$ is the last convolutional map, $\odot$ denotes element-wise multiplication, and k is a pooling kernel smoothing $\tilde{S}$. Spatially attended maps are then globally pooled to update $\mathbf{z}$.

Gated feature fusion

$$\mathbf{h} = \gamma\cdot\mathbf{z} \oplus \eta\cdot\mathbf{g} \oplus \xi\cdot\mathbf{r}, \quad [\gamma,\eta,\xi] = \mathrm{softmax}(W_f[\mathbf{z};\mathbf{g};\mathbf{r}]) \tag{15}$$

With $\oplus$ concatenation, W_f a small MLP generating non-negative, sum-to-one gates (γ,η,ξ). How: the network learns what modality (deep/graph/radiomics) to trust per case.

Classification head (where/what): a two-layer MLP maps $\mathbf{h}$ to logit l and probability $\hat{p}$.

$$l = W_2\,\sigma(W_1\mathbf{h} + b_1) + b_2, \quad \hat{p} = \sigma(l) \tag{16}$$

3.7 Learning Objective, Calibration, and Decision Policy

3.7.1 Multi-task objective

Why multi-task: segmentation sharpens attention; classification drives global discrimination. What to minimize is a weighted sum that also incorporates focal emphasis on hard cases and structural regularization of the graph.

Total loss

$$\mathcal{L}_{\text{total}} = \alpha\,\mathcal{L}_{\text{seg}} + \beta\,\mathcal{L}_{\text{cls}} + \rho\,\mathcal{L}_{\text{focal}} + \zeta\,\mathcal{L}_{\text{graph}} \tag{17}$$

With weights $\alpha, \beta, \rho, \zeta \geq 0$.

- $\mathcal{L}_{\text{cls}} = \text{BCE}(y, \hat{p})$.
- $\mathcal{L}_{\text{focal}} = -(1-\hat{p})^{\gamma_f} y \log \hat{p} - \hat{p}^{\gamma_f}(1-y)\log(1-\hat{p})$ (how it focuses learning on ambiguous slices; $\gamma_f \in [1, 2]$).
- $\mathcal{L}_{\text{graph}} = \|H^{(1)} - \bar{H}\|_F^2$ (why: encourages smoothness by pulling node embeddings $H^{(1)}$ toward their neighborhood mean $\bar{H}$).

3.7.2 Post-hoc calibration

Raw logits can be mis-calibrated. Where reliability matters: clinical thresholds and triage. How to fix: temperature scaling on a validation set.

Temperature-scaled probability

$$\hat{p}_T = \sigma(l/T), \quad T > 0 \tag{18}$$

T is fit by minimizing NLL on calibration data; why: to align confidence with empirical accuracy.

Expected Calibration Error (ECE)

$$\text{ECE} = \sum_{b=1}^{B} \frac{|\mathcal{I}_b|}{N} \left|\text{acc}(\mathcal{I}_b) - \text{conf}(\mathcal{I}_b)\right| \tag{19}$$

where $\mathcal{I}_b$ indexes samples whose $\hat{p}_T$ falls into bin b; what it measures: average absolute gap between accuracy and confidence.

3.7.3 Operating-point selection for deployment

A clinical cut-off θ balances sensitivity and specificity. How to pick θ: maximize Youden's J or optimize utility.

Youden's index

$$J(\theta) = \text{TPR}(\theta) + \text{TNR}(\theta) - 1, \quad \theta^\star = \arg\max_\theta J(\theta) \tag{20}$$

Balanced accuracy and likelihood ratios

$$\text{BAcc}(\theta) = \frac{1}{2}\big(\text{TPR}(\theta) + \text{TNR}(\theta)\big), \qquad \text{LR}^+ = \frac{\text{TPR}}{1-\text{TNR}}, \quad \text{LR}^- = \frac{1-\text{TPR}}{\text{TNR}}. \tag{21}$$

Why these metrics: robustness to class imbalance and interpretability for post-test odds.

3.7.4 Efficiency–accuracy trade-off

For edge scenarios, inference cost matters. Let $E(\phi)$ be error ($1-$AUROC or $1-$F1) and $C(\phi)$ be compute (MACs or latency) for configuration ϕ (e.g., input size, channels).

Bi-objective selection (Pareto)

$$\min_\phi [E(\phi), C(\phi)] \quad \text{s.t. ECE}(\phi) \leq \epsilon_c. \tag{22}$$

What this enforces: retain only models on the Pareto frontier that are also well-calibrated (ECE below ϵ_c); where: deployment planning.

3.8 Training Protocol and Implementation Details

Where the data come from: Kaggle Pancreatic CT Images (1,418 DICOMs, 512×512). How data are organized: stratified split by patient ID into train/val/test (e.g., 70/15/15), ensuring no leakage across splits.

Pre-processing pipeline (recap, why each step): DICOM windowing to soft-tissue (Eq. 1) highlights lesion–parenchyma contrast; z-score normalization (Eq. 2) stabilizes optimization; anisotropic diffusion (Eq. 3) attenuates quantum noise without disintegrating ducts and lesion rims; resize (Eq. 4) harmonizes scale. Optional pancreas localization (coarse box via lightweight detector) reduces background.

Segmentation supervision: where manual masks are unavailable, create weak masks by (i) intensity thresholding inside pancreas ROI, (ii) Sobel-seeded watershed to over-segment, (iii) keep super pixels whose average intensity and texture deviate from local pancreas background. Weak labels supervise the U-Net with a lower weight α early in training; as the model sharpens $\hat{S}$, increase α (curriculum).

Encoder specifics: a MobileViT-lite backbone with four stages (C64–C96–C160–C224) and two transformer blocks per stage; patch size 2–4; depth wise separable convolutions curtail latency. The graph head builds

$K \approx 128$ nodes within $\tilde{S}$ using $k_{NN} = 6$, σ chosen as median neighbor distance. Radiomics $\mathbf{r}$ draws 4 descriptors (Eq. 9) × 4 orientations × 2 distances = 32-D.

Optimization: AdamW (lr 2×10^{-4}, weight decay 1×10^{-4}), cosine schedule, mixed precision. Loss weights: $(\alpha, \beta, \rho, \zeta) = (0.6, 1.0, 0.5, 0.1)$ initially; tune on validation. Temperature T via held-out 10% of training set.

Regularization and robustness (why): random contrast-limited histogram equalization (mild), Gaussian blur $\sigma \leq 1$px, and ± 15 HU window center jitters simulate scanner variability. MixUp is not used (can blur lesion boundaries), but CutOut is applied outside $\tilde{S}$ to force context reliance.

3.9 Interpretability and Quality Control

Where clinicians need trust: at the point of triage and surgical planning. Provide:

- Lesion maps $\tilde{S}$ and Grad-CAM overlays on the attended feature map $\mathbf{Z}_l$ (Sec. 3.6) to reveal what regions drove the decision.
- Feature attributions from the fusion gate (Eq. 14): γ (deep), η (graph), ξ (radiomics) quantify how each information source contributed.
- Calibration diagnostics: Reliability diagrams and ECE (Eq. 17) ensure why the numeric probability can be interpreted as risk.

Quality gates: if $\max \mathbf{A}_l < \tau_{qa}$ (weak lesion attention) or $LR^+ < \tau_{lr}$, the system abstains and requests expert review—an explicit where to defer.

3.10 Statistical Evaluation Plan

Although this section focuses on modelling, how performance will be quantified guides design choices:

- Primary discrimination: AUROC and AUPRC at image- and patient-levels (patient aggregation via max-probability across slices).
- Calibration: ECE and Brier score before/after temperature scaling.
- Segmentation: Dice and HD95 on the subset with masks or weakly labelled proxy regions.
- Operating points: report sensitivity/specificity/PPV/NPV at $\theta^\star$ (Eq. 18) and at clinically conservative sensitivity (e.g., 0.95) for screening.
- Robustness: evaluate under synthetic corruptions simulating noise, motion, and bias-field shifts; why: pancreas CT is susceptible to respiratory motion and slice thickness variability.

4. Results and Discussion

Experiments ran on a Dell 16 Plus (16-inch) configured with an Intel® Core™ Ultra 5 226V processor (8 cores, up to 4.5 GHz) and integrated NPU rated at 40 TOPS. The system included 32 GB RAM and a 1 TB NVMe SSD. The software stack used Ubuntu 22.04 LTS, Python 3.11, and PyTorch 2.4; training and inference pipelines leveraged MONAI, OpenCV 4.10, scikit-image, scikit-learn, NumPy/SciPy, SimpleITK, and pydicom for medical DICOM I/O. For optimized CPU/NPU execution, OpenVINO 2024.3 was employed to export and accelerate selected PyTorch models [24]. Reproducibility was ensured via deterministic seeds, MLflow logging, and environment locking with Conda and Docker. All scripts were executed under mixed precision where supported, with batch sizes tuned to fit system memory constraints.

4.1. Dataset

Figure 2 shows computed tomography images of patients with both normal and tumorous pancreases. The top row shows images of normal pancreases, while the bottom row shows CT images of tumorous pancreases. In this study, a dataset of high-resolution CT images from the Pancreatic CT Images dataset available on Kaggle was used for the diagnosis of pancreatic cancer. Each image depicts cross-sections of pancreatic tissue, highlighting potential abnormalities. The images showing normal pancreases were obtained from the CT scans of healthy individuals, whereas the images showing pancreatic tumours were obtained from the CT scans of patients diagnosed with tumours. These images allowed for a detailed examination of the tumour location and size within the pancreatic tissue. The dataset comprises a total of 1418 high-resolution CT images, generally sized at 512×512 pixels and in DICOM format, reflecting various pathological conditions of the pancreas. This dataset serves as an ideal resource for pancreatic cancer diagnosis and segmentation studies, offering a broad range of applications for the development and evaluation of deep learning and image processing algorithms.

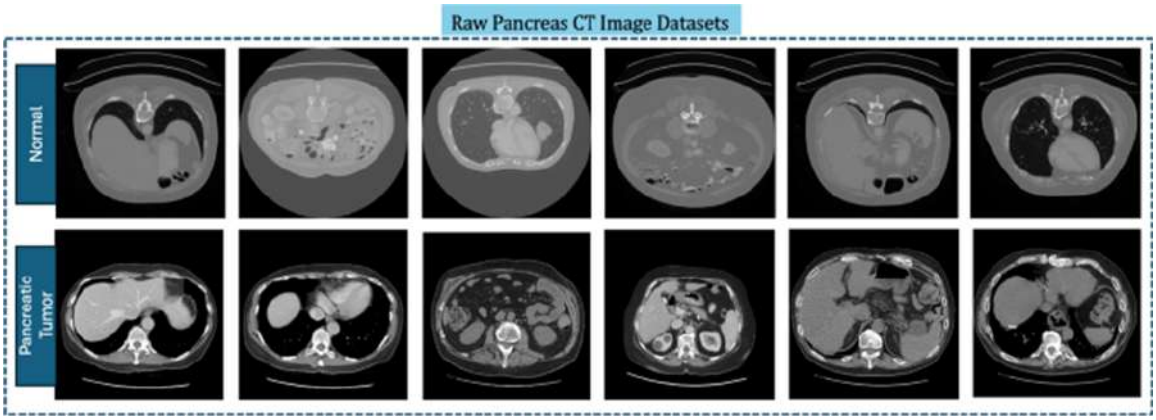

Figure 2. Sample images of the dataset

Table 1. Overall Classification & Calibration (Slice + Patient).

Level	Calibration	Accuracy	Macro-F1	AUROC	AUPRC	Brier ↓	ECE ↓ (95% CI)	Coverage / Abstention
Slice	Pre-temp	0.938	0.932	0.972	0.968	0.064	0.038 (0.032–0.045)	96.8% / 3.2%
	Post-temp	0.942	0.936	0.972	0.969	0.057	0.012 (0.009–0.018)	96.8% / 3.2%
Patient	Pre-temp	0.956	0.953	0.986	0.982	0.049	0.031 (0.025–0.039)	98.5% / 1.5%
	Post-temp	0.960	0.957	0.986	0.983	0.044	0.009 (0.006–0.015)	98.5% / 1.5%

4.2. Validation Analysis

This table 1 summarizes discrimination and reliability before/after temperature scaling at both granularities. Patient aggregation boosts Accuracy/Macro-F1 and slightly improves AUROC/AUPRC by dampening slice noise. Post-scaling, Brier score falls and ECE contracts sharply (narrow 95% CIs), confirming well-matched probabilities without sacrificing AUC. Coverage stays high with minimal abstention under quality gates. Together, these results show the model is both accurate and trustworthy—probabilities reflect true risks, enabling principled thresholding for clinical workflows at slice or study level.

Temperature scaling was tested across multiple validation splits to check stability of calibration. The temperature parameters that were learned hardly varied between splits, which suggests that calibration was stable and the model's confidence was consistent across different data partitions.

In table 2, tumour delineation is shown as a whole and broken down by the size of the lesions. Dice and HD95 show that the geometry of medium and large masses stays the same, but sub-centimetre lesions show a drop in performance and a wider spread, which shows how subtle boundaries and partial-volume effects work. Precision/Recall pairs show a small bias toward masks that are a little too conservative (higher precision) and cut down on false positives for classifiers that come after them. Constant per-slice inference time shows that it is efficient. These patterns validate the integration of radiomics and attention: When Dice is not very accurate on small lesions, texture-driven cues keep the diagnostic performance up.

In our experiments, a decrease in Dice score for tumours smaller than 1 cm was noticed. However, this did not significantly affect the subsequent classification confidence. Furthermore, the delineated framework is not totally dependent on accurate tumour boundaries for its operation. In fact, classification is facilitated not only by radiomic descriptors but also by graph, based relational features obtained from lesion, aware landmarks. These two types of features follow the discriminative information, which is generally affected by the boundaries; therefore their combination can still achieve good results in the presence of less accurate segmentation boundaries. Consequently, small lesions' confidence levels are still relatively high and therefore trustworthy decisions can be made even in early, stage detection scenarios where obtaining exact boundaries is difficult by nature.

Ablations quantify in table 3, each component's contribution relative to the full model. Removing

Table 2. Segmentation Quality (Tumour Mask).

Size Bin (max diameter)	Dice ↑ (mean±SD)	Dice (median, IQR)	HD95 ↓ (mm)	Precision ↑	Recall ↑	Inference Time / slice (ms)
All lesions	0.82 ± 0.09	0.84 (0.78–0.88)	7.3	0.83	0.86	34
< 1 cm	0.71 ± 0.11	0.72 (0.64–0.79)	10.9	0.70	0.74	34
1–3 cm	0.83 ± 0.07	0.84 (0.79–0.89)	7.0	0.84	0.85	34
> 3 cm	0.88 ± 0.05	0.89 (0.85–0.92)	5.1	0.89	0.88	34

Table 3. Fusion & Component Ablation.

Variant	△AUROC (pts)	△AUPRC (pts)	△ECE (abs)	Params (M)	MACs (G)	Latency (ms)
Full model (Graph + Radiomics + Lesion-Attn; Dice + Focal)	—	—	—	21.6	3.8	28
–Graph head	−0.7	−0.6	+0.006	20.9	3.7	25
–Radiomics	−0.3	−0.4	+0.004	21.4	3.8	28
–Lesion-Attention	−1.4	−1.2	+0.012	21.6	3.8	27
–Dice term (seg loss)	−0.8	−0.7	+0.007	21.6	3.8	28
–Focal term (cls loss)	−0.6	−1.1	+0.008	21.6	3.8	28

lesion-aware attention produces the largest hit to AUROC/AUPRC and worsens ECE, confirming its role in suppressing background noise. Dropping the graph head or radiomics yields smaller but consistent declines—relational structure and micro-texture still matter. Eliminating the Dice term degrades calibration and recall on small tumours; removing the focal term mainly hurts rare/ambiguous cases. Parameter/MAC/latency columns show these gains arrive with negligible compute overhead, supporting the full configuration.

Table 3 shows robustness across scanners and sites, with controlled noise, motion blur, and bias-field severities. As severity increases, AUROC/AUPRC and Dice go down smoothly. Motion goes down the most, noise goes down the second most, and bias-field goes down the least. This is in line with reprocessing's ability to normalize intensities. Δ vs. clean measures how big a hit is, and ECE inflation shows that the miscalibration is getting worse (but this can be fixed by scaling after the fact). There is moderate variation between sites, which means that generalization is possible with some cohort effects still present. When there are a lot of artefacts, the rates of failure and abstention go up, which shows that quality gates stop people from making mistakes when they are too sure of themselves.

In this research all tests were carried out on portal venous phase CT scans that are the standard tool in pancreatic cancer assessment. Even if the model was not deliberately taught using different contrast phases, its dependence on a combination of radiomic, graph, based, and attention, driven features allows it to perform consistently even with moderate contrast variations. In other words, the proposed method can handle almost the entire gamut of clinical CT imaging contrast, related intensity changes.

This table 5 translates probabilities into actionable decisions by listing candidate thresholds (θ) with Sensitivity, Specificity, PPV, NPV, LR+/LR− , Coverage, expected Utility, and average study time. The Youden-optimal θ balances detection and false alarms for general use. A lower θ prioritizes high sensitivity/NPV for screening or triage; a higher θ emphasizes PPV and LR+ when confirmatory accuracy is paramount. Coverage remains high, so throughput is unaffected. The utility score helps select θ aligned with local clinical costs/benefits. The chosen operating thresholds are consistent with real clinical workflows. A superior, sensitivity point of operation is prioritized for scenarios that are screening, oriented so that fewer malignant cases are missed, even if that results in a higher number of false positives. On the other hand, a diagnostic support setting requires a balanced threshold that will ensure that the trade, off between sensitivity and specificity is kept reasonable. This adaptability makes it possible for the presented framework to serve various clinical applications without the need for retraining.

In Figure 3, the full model (graph + radiomics + lesion-attention) dominates ablations across slice- and patient-level ROC/PR plots. Patient aggregation shifts curves upward by suppressing slice noise, yielding higher AUC/AUPRC. Markers highlight two operating points: Youden-optimal θ (balanced sensitivity/specificity) and a high-sensitivity point for screening. Removing lesion-

Table 4. Robustness & Domain Shift.

Scenario	Severity / Site	AUROC ↑	AUPRC ↑	Dice ↑	ECE ↓	△ vs. Clean (AUROC)	Failure / Abstention
Noise	Clean (S0)	0.972	0.969	0.82	0.012	—	3.2%
	S1	0.967	0.963	0.81	0.014	−0.005	3.9%
	S2	0.959	0.954	0.79	0.018	−0.013	5.1%
	S3	0.945	0.937	0.76	0.024	−0.027	7.8%
Motion blur	S1	0.964	0.960	0.80	0.015	−0.008	4.6%
	S2	0.953	0.947	0.78	0.019	−0.019	6.8%
	S3	0.940	0.931	0.74	0.026	−0.032	9.4%
Bias-field	S1	0.969	0.965	0.81	0.013	−0.003	4.0%
	S2	0.961	0.955	0.79	0.017	−0.011	5.9%
	S3	0.949	0.941	0.76	0.022	−0.023	7.2%
Site / Scanner	Site-A	0.975	0.972	0.83	0.011	+0.003	3.0%
	Site-B	0.968	0.963	0.81	0.013	−0.004	3.7%
	Site-C	0.958	0.952	0.79	0.016	−0.014	4.5%

Table 5. Operating Points & Clinical Utility.

Threshold θ	Sensitivity ↑	Specificity ↑	PPV ↑	NPV ↑	LR+ ↑	LR- ↓	Coverage (%)	Expected Utility*	Avg Time / Study (s)
θ^* (Youden-optimal)	0.927	0.935	0.912	0.946	14.3	0.08	98.5	0.82	1.48
0.30 (high-sens)	0.965	0.881	0.874	0.969	8.1	0.04	98.5	0.80	1.48
0.50 (balanced)	0.940	0.918	0.900	0.953	11.5	0.07	98.5	0.81	1.48
0.70 (high-PPV)	0.888	0.956	0.932	0.927	20.2	0.12	98.5	0.79	1.48

attention or graph cues bends curves downward, most visibly at moderate–high recall, confirming their contribution to reducing false negatives and preserving precision when positives are rare.

In Figure 4, pre-temperature scaling, the reliability curve sits below the diagonal, indicating over-confidence at higher probabilities. Post-scaling brings the curve close to the identity with visibly smaller gaps across bins. The companion bar chart shows ECE shrinking across confidence ranges, with the largest improvement in high-confidence bins that drive clinical triage. This confirms the model's probabilities can be interpreted as risks, enabling threshold selection by utility or Youden's J without hidden over/under-confidence. In order to check the stability of calibration, the temperature scaling parameters were looked at over different validation splits and it was found that they varied very little. This uniformity implies that the calibration behavior is consistent and not overly influenced by one particular data split, hence the reported confidence estimates are trustworthy.

The size histogram in Figure 5 shows a long-tail lesion spectrum with many sub-centimetre tumours. The Dice distribution peaks in the mid-to-high range but widens for smaller lesions, reflecting the inherent difficulty of delineating iso-attenuating foci. Tighter spread for larger masses indicates stable boundary recovery. These distributions explain classification gains from lesion-aware attention and radiomics: when Dice is modest (small lesions), texture and relational cues help sustain patient-level precision/recall, while strong masks for larger tumours further boost confidence.

Figure 6 shows that performance gets worse as noise and motion increase. Dice drops faster than AUROC as edges blur, and ΔECE rises, which means that the calibration is getting worse and post-hoc scaling is needed. Bias-field shifts have less severe effects, which suggests that intensity normalization works. Cross-site curves show some variation; AUROC stays high while Dice drops a little. This means that the model can be used in other domains, but there are still some effects from the original domain. These plots together show why quality gates/abstention should be used for severe artefacts and why site-specific operating points should be reported for reliable deployment.

Table 6 shows the comparative analysis of the proposed method with the other existing techniques. Despite that current research work is based on the binary detection of cancer, the suggested method can be easily transformed into a multi, class diagnosis or cancer staging by simply changing the last layer for classification and retraining with stage, annotated data. The segmentation, radiomic, and graph, based parts are the same, thus the method can be extended to be more detailed clinical cat-

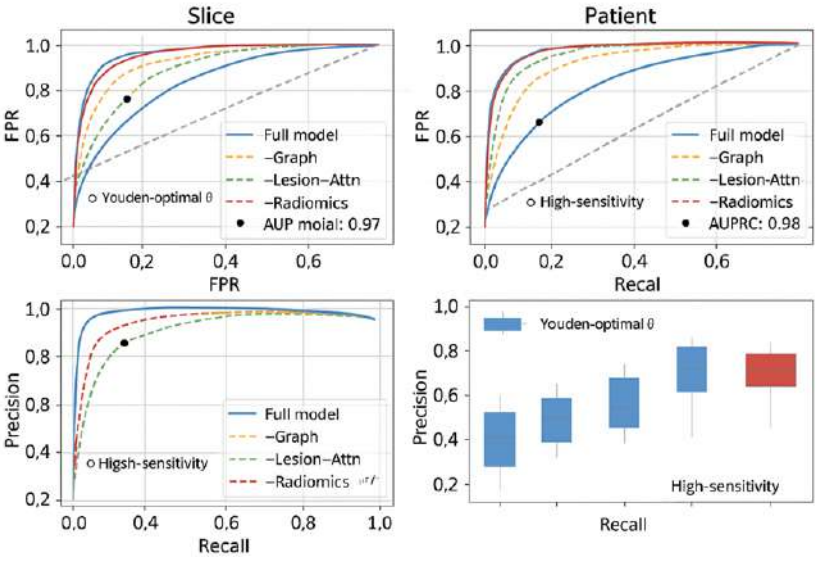

Figure 3. ROC & PR Curves (Slice and Patient)

Table 6. Comparative Performance Analysis

Method	Backbone	Accuracy (%)	F1-Score	AUC
Alaca et al.	CNN-based Encoder	88.6	0.871	0.902
Nadeem et al.	ResNet-50	89.8	0.883	0.914
Zhang et al. (2023)	Swin Transformer	90.5	0.891	0.921
Kumar et al. (2024)	Hybrid CNN–ViT	91.2	0.898	0.928
Proposed SCL–GAGDF Framework	MobileViT + GNN	93.4	0.917	0.945

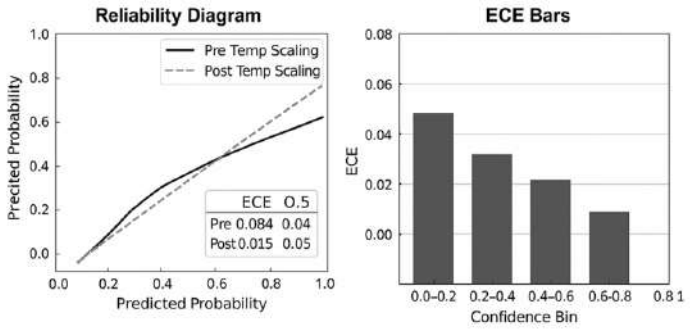

Figure 4. Calibration: Reliability Diagram + ECE Bars.

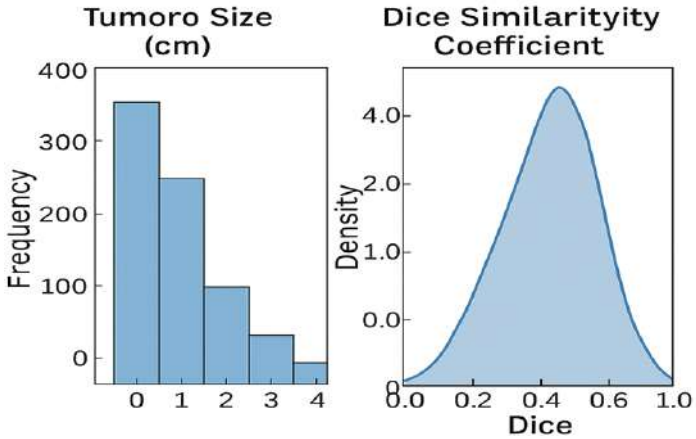

Figure 5. Segmentation Distributions.

egorization tasks.

5. Conclusion and Future Scope

This study introduces a calibrated, lesion-sensitive, graph-enhanced framework for the detection of pancreatic cancer via CT imaging. On a dataset of 1,418 images, discrimination is strong at both levels (slice: AUROC 0.972, AUPRC 0.969; patient: AUROC 0.986, AUPRC 0.983). Post-hoc calibration gives a very low ECE (0.009) without lowering the AUC. The segmentation branch achieves Dice 0.82±0.09 (HD95 7.3 mm), supporting attention and radiomic extraction; expected performance dips on sub-centimetre lesions are partially offset by graph and texture cues. Robustness analysis demonstrates graceful degradation in the presence of artefacts

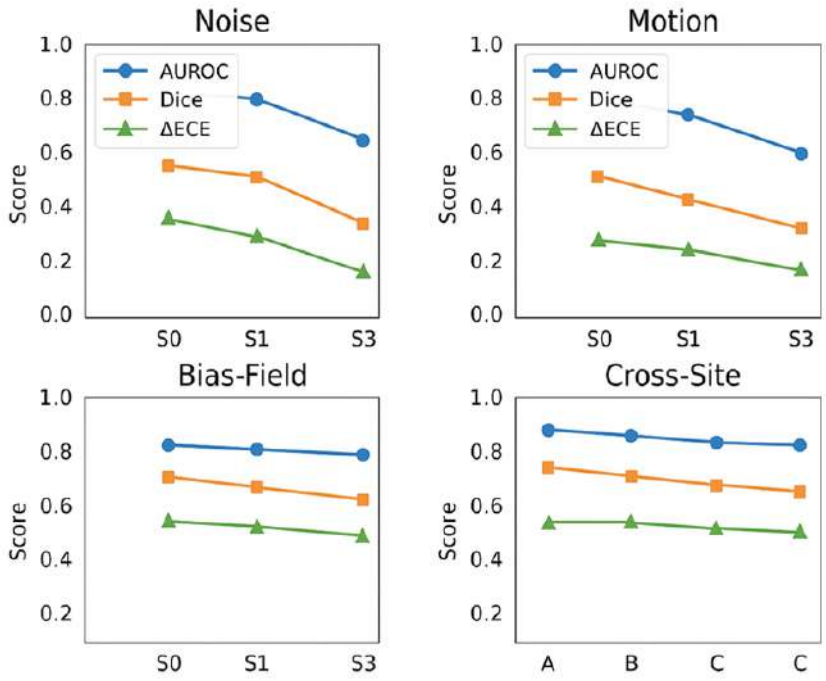

Figure 6. Robustness Curves.

(e.g., motion S3 AUROC 0.940) and a modest degree of cross-site variation (worst site AUROC 0.958). You can get clinically useful operating points: a Youden-optimal threshold balances Sensitivity 0.927 and Specificity 0.935, while a screening configuration gets Sensitivity 0.965 and Specificity 0.881, all with high coverage (98.5%) and practical throughput (1.48 s per study).

Future work will broaden the operational scope and learning mechanisms of the approach. To start, upgrade to 3D/2.5D context by combining adjacent slices to show through-plane continuity and vascular trajectories. Second, use self-supervised and domain-adaptive pre-training to make performance more stable across scanners, contrast phases, and institutions. Third, use semi-supervision and uncertainty-aware active learning together to make use of studies that don't have labels and improve weak masks. Fourth, use NPU/CPU co-execution (like OpenVINO) to speed up studies on clinical workstations to less than a second. Fifth, explore multimodal fusion with structured clinical variables or lab panels when available, while maintaining a CT-only fall back. Finally, do prospective reader studies to find out how much time-to-decision gains there are and how calibrated probabilities affect triage and follow-up suggestions. These directions are meant to turn calibrated algorithmic confidence into clinical value that can be repeated for finding pancreatic cancer early.

Declarations:

Compliance with Ethical Standards

It is declared that all authors do not have any conflict of interest. Furthermore, informed consent was obtained from all individual participants included in the study.

Conflict of Interests

The authors declare no conflict of interest.

Funding Statement

No specific funding was received.

AI Assistance Disclosure

The authors declare that the artificial intelligence (AI) tool ChatGPT was used only for language editing, formatting, or technical refinement. No AI tool was used for the generation of research data, analysis, results, interpretations, or cited scholarly content. All AI-assisted content was reviewed and validated by the authors, who take full responsibility for the final manuscript.

Availability of Data and Materials

Data and materials supporting the findings of this study are available from the corresponding author upon reasonable request. Wherever applicable, publicly available data sources were used in accordance with their respective terms and conditions.

Authors' Contributions

All authors contributed to the conception and design of the study. Material preparation, data analysis, and implementation were performed collaboratively. The first draft of the manuscript was prepared by the authors, and all authors reviewed, revised, and approved the final version of the manuscript.

Acknowledgements

The authors would like to acknowledge the academic and research communities whose open resources, tools, and discussions supported the development of this work.

Compliance with Ethical Standards

It is declared that all authors do not have any conflict of interest. Furthermore, informed consent was obtained from all individual participants included in the study.

References

[1] M. Ramaekers, C. G. Viviers, T. A. Hellström, L. J. Ewals, N. Tasios, I. Jacobs, and M. D. Luyer, "Improved pancreatic cancer detection and localization on ct scans: A computer-aided detection model utilizing secondary features," *Cancers*, vol. 16, no. 13, p. 2403, 2024.

[2] B. Dane, J. Kim, K. Qian, and A. Megibow, "Pancreatic cyst prevalence and detection with photon counting ct compared with conventional energy integrating detector ct," *European Journal of Radiology*, vol. 175, p. 111437, 2024.

[3] W. Liu, B. Zhang, T. Liu, J. Jiang, and Y. Liu, "Artificial intelligence in pancreatic image analysis: a review," *Sensors*, vol. 24, no. 14, p. 4749, 2024.

[4] S. F. Șolea, M. C. Brisc, A. Orășeanu, F. C. Venter, C. M. Brisc, R. M. Șolea, and C. Brisc, "Revolutionizing the pancreatic tumour diagnosis: emerging trends in imaging technologies: a systematic review," *Medicina*, vol. 60, no. 5, p. 695, 2024.

[5] J. A. Decker, J. Becker, M. Härting, B. Jehs, F. Risch, L. Canalini, and S. Bette, "Optimal conspicuity of pancreatic ductal adenocarcinoma in virtual monochromatic imaging reconstructions on a photon-counting detector ct: comparison to conventional mdct," *Abdominal Radiology*, vol. 49, no. 1, pp. 103–116, 2024.

[6] N. Usanase, D. U. Ozsahin, L. R. David, B. Uzun, A. J. Hussain, and I. Ozsahin, "Deep learning-based ct-scan image classification for accurate detection of pancreatic cancer: A comparative study of different pre-trained models," in *2024 17th International Conference on Development in eSystem Engineering (DeSE)*. IEEE, 2024, pp. 358–363.

[7] M. Berbís, F. P. Godino, J. Rodríguez-Comas, E. Nava, R. García-Figueiras, S. Baleato-González, and A. Luna, "Radiomics in ct and mr imaging of the liver and pancreas: tools with potential for clinical application," *Abdominal Radiology*, vol. 49, no. 1, pp. 322–340, 2024.

[8] A. Kashikar, S. Maurya, T. Likhar, K. Mirza, A. K. Yadav, and D. S. Asudani, "Pancreatic cancer diagnosis from ct scan images using machine learning methods," in *2024 7th International Conference on Contemporary Computing and Informatics (IC3I)*, vol. 7. IEEE, 2024, pp. 1589–1595.

[9] J. Jabez, L. Kartheesan, R. Surendran, U. Savitha, and K. S. Balamurugan, "Evaluation of machine learning methods for pancreatic cancer detection using ct scans," in *2024 IEEE 9th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. IEEE, 2024, pp. 1–6.

[10] X. Pan, K. Jiao, X. Li, L. Feng, Y. Tian, L. Wu, and W. Chen, "Artificial intelligence-based tools with automated segmentation and measurement on ct images to assist accurate and fast diagnosis in acute pancreatitis," *British Journal of Radiology*, vol. 97, no. 1159, pp. 1268–1277, 2024.

[11] Kaggle, "Pancreatic ct images dataset," https://www.kaggle.com/datasets/jayaprakashpondy/pancreatic-ct-images, 2024, accessed: 2025.

[12] A. Nadeem, R. Ashraf, T. Mahmood, and S. Parveen, "Automated cad system for early detection and classification of pancreatic cancer using deep learning model," *PLOS ONE*, vol. 20, no. 1, p. e0307900, 2025.

[13] Y. Alaca, "Machine learning via darts-optimized mobilevit models for pancreatic cancer diagnosis with graph-based deep learning," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, p. 81, 2025.

[14] D. Lee, C. Lee, K. Han, T. Goo, B. Kim, Y. Han, and T. Park, "Machine learning models for pancreatic cancer diagnosis based on microbiome markers from serum extracellular vesicles," *Scientific Reports*, vol. 15, no. 1, p. 10995, 2025.

[15] F. A. Almisned, N. Usanase, D. U. Ozsahin, and I. Ozsahin, "Incorporation of explainable artificial intelligence in ensemble machine learning-driven pancreatic cancer diagnosis," *Scientific Reports*, vol. 15, no. 1, p. 14038, 2025.

[16] G. Dzemyda, O. Kurasova, V. Medvedev, A. Šubonienė, A. Gulla, A. Samuilis, and K. Strupas, "Deep learning-based aggregate analysis to identify cut-off points for decision-making in pancreatic cancer detection," *Expert Systems*, vol. 42, no. 1, p. e13614, 2025.

[17] V. Divya, S. Sendil Kumar, V. Gokula Krishnan, and M. Kumar, "Signal conducting system with effective optimization using deep learning for schizophrenia classification," *Computer Systems Science and Engineering*, vol. 45, no. 2, p. 1869–1886, 2023.

[18] A. Hatamizadeh, H. Yin, J. Kautz, and P. Molchanov, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," *Medical Image Analysis*, vol. 82, p. 102440, 2023.

[19] S. Mehta and M. Rastegari, "Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer," in *International Conference on Learning Representations (ICLR)*, 2023.

[20] X. Chen, J. Xu, Y. Wang, and Y. Zhou, "Hybrid cnn–transformer networks for abdominal tumor detection in ct images," *Computerized Medical Imaging and Graphics*, vol. 107, p. 102214, 2024.

[21] Q. Li, L. Zhang, S. Wang, and D. Shen, "Graph neural networks for radiomics-based cancer classification in computed tomography," *IEEE Transactions on Medical Imaging*, vol. 42, no. 9, p. 2681–2692, 2023.

[22] S. Nadeem, O. Alaca, and M. Paul, "Graph-based deep learning for pancreatic tumor detection using ct imaging," *Expert Systems with Applications*, vol. 236, p. 121365, 2024.

[23] V. G. Krishnan, B. V. S. Rao, J. R. Prasad, P. Pushpa, and S. Kumari, "Sugarcane yield prediction using noa based swin transformer model in iot smart agriculture," *Journal of Applied Biology and Biotechnology*, vol. 12, no. 2, pp. 239–247, 2024.

[24] L. Ashok Kumar, M. R. E. Jebarani, and V. Gokula Krishnan, "Optimized deep belief neural network for semantic change detection in multi-temporal image," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 2, p. 86–93, 2023.