A Multi-Model Deep Learning Framework FOR Pancreatic Tumor Detection

Mrs.D.Sasikala[1], Y.Vishnuvardhan Reddy[2], G. Jaswanth[3], M.Bhanu Prakash[4]
[1] Prof, Dept of cyber security
[2, 3, 4]Dept of cyber security
[1, 2, 3, 4] Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu – 621 112, India

Abstract- *The Pancreatic Tumor Detection Pipeline introduces a unified deep learning system for automated detection, localization, and segmentation of tumors in medical imaging. The system utilizes a three-stage architecture by combining complementary neural network models to provide robust and interpretable analysis. A Convolutional Neural Network (CNN) is used in the first stage for binary classification for detecting the presence of a tumor in medical scans. The second phase makes use of a Swin Transformer-based model for accurate bounding box location using hierarchical attention mechanisms to understand spatial context. The third phase uses the nnU-Net approach to attain pixel-level segmentation for clearly defining tumor boundaries with high accuracy. The pipeline is designed to work with various medical image formats such as PNG, JPG, DICOM, and NIfTI, thus increasing its flexibility to accommodate different clinical data. User-friendly web interface based on Flask offers visualization of raw images, masks, and overlay outputs, whereas a SQLite database stores history of analysis for traceability. API endpoints are also available for integration with external diagnostic tools. The pipeline is researched for clinical translation and is designed to have end-to-end automation, GPU acceleration for performance, and modular retraining across various datasets. This paper illustrates the possibility of integrating CNN, Transformer, and U-Net paradigms for end-to-end tumor analysis, pushing the boundaries of automating and reproducibility in computer-aided medical diagnostics.Keywords: Pancreatic Tumor Detection, Deep Learning, Convolutional Neural Network (CNN), Swin Transformer, nnU-Net Segmentation, Medical Image Localization, CT Scan Analysis, Tumor Segmentation, Computer-Aided Diagnosis (CAD), Artificial Intelligence in Healthcare, Attention-Based Feature Extraction, Hybrid Deep Learning Framework, Medical Image Processing, DICOM and NIfTI Image Support, Early Cancer Detection, 2D Medical Image Classification, Pixel-Level Segmentation, Automated Diagnosis, Clinical Decision Support System, Flask Web Interface, GPU-Accelerated Inference, End-to-End Tumor Analysis, Medical Imaging Automation, Transformer-Based Localization.*

I. INTRODUCTION

In this world, every year, 4.6 million new cancer cases are diagnosed in the WHO European Region, 2.1 million people die from cancer, among those 4.6 million, more than 2,60,000 people are affected by Pancreatic cancer, and many cases are found only after the disease has advanced. Pancreatic cancer is a type of cancer that is a fast-spreading and highly deadly disease that affects the pancreas. It begins as a tumour in the cells of the pancreas, an organ located behind the stomach that helps in food digestion and blood sugar control by producing enzymes and hormones. Cancer mostly shows no clear symptoms until it is too late. People may lose weight or stop feeling hungry. Some people develop yellow eyes and skin, dark urine, or unusual stool changes. It can also cause diabetes, tiredness, and blood clots, making treatment very hard. The above symptoms can be easily confused with common stomach problems in their early stages.

In patients who have the symptoms of Pancreatic cancer, most forms of diagnostic tests are typically conducted to help the doctor detect the cancer. The tests which they usually begin with are blood tests such as CA19-9, although this is not a consistently accurate test, since there are times where it is raised even when the individual does not have cancer. They are also able to examine the tumour within the pancreas using CT scans, MRIs, PET-CT scans, endoscopic ultrasound. The PET-CT is also very expensive and is not available in most hospitals; the images are not very detailed and therefore the surgeon can not have a clear picture of the tumour.

It also presents false results confusing cancer with inflammation. Endoscopic ultrasound is better in giving crystal clear images yet the procedure is very invasive and in some cases very painful to the patient. A biopsy is performed to be sure of the presence of the cancer however it has its drawbacks such as bleeding and infection and in some cases doctors are not able to perform it because of the location of the tumour which may be too dangerous or inaccessible.

In this report, deep learning models are applied to locate tumours in the pancreas in CT scan images. It is geared towards enhancing clinical diagnostic tests. Small tumourscan not be detected in many tests, including PET-CT and endoscopic ultrasound. Doctors also find them expensive and disorienting often. The deep learning technique can be used to better forecast tumours. The model applied in classification to confirm the presence or absence of cancer in the pancreas in our study is CNN. Swin Transformer is also applied in the localisation of tumours, which reveals the precise location of the cancer. The tumour shape is separated into the nnU-Net. This assists physicians in their quicker diagnosis and minimizes mistakes

II. LITERATURE SURVEY

Zhou et al. [1] was established the 3D CNN model to identify pancreatic tumours in CT scans was developed on the basis of processing the volumetric data to detect the suspicious regions. Their approach was much more sensitive to the identification of tiny lesions, but it needed large volumes of data and demanded considerable amounts of computation. Shanmugam et al. [2] described the Dual Self-Attentive Transformer U-Net (DSTUnet) to effectively perform pancreatic segmentation by Swin Transformers to obtain global contextual features. It was expensive to calculate, high-Dice score, and the model had a high degree of accuracy in delineating the pancreas. Wang and Li [3] proposed a hybrid network that uses a pre-trained CNN to extract features and a recurrent neural network (RNN) to analyse consecutive CT slices. Their model described contextual dependencies better across intersecting slices to achieve superior tumour localisation, but with a high complexity of training. Zhang et al. [4] used the nnU-Net model of automatic pancreatic tumour segmentation, which yielded outstanding results and flexibility; nevertheless, it was not interpretable to clinical use. Chen et al. [5] improved pancreatic cancer detection by adding a Swin Transformer block to a U-Net encoder to improve classification with long-range feature learning with an increased amount of model parameters and memory consumption. In Liu et al. [6], it was suggested to use the ensemble approach between CNN and Swin Transformer to predict the presence of pancreatic tumours to enhance the robustness and diagnostic quality but required a lot of computing power to analyze in real-time. Kim et al. [7] were applying the Vision Transformer (ViT) to the patch of the CT scan and classifying pancreatic tumours, effectively detecting malignancies without the use of convolutions and with the need to train on vast amounts of data.. Patel et al. [8] designed a lightweight CNN architecture for quick screening of pancreatic tumours. This method is suitable for initial diagnosis, but sometimes misses small or iso-attenuating tumours. Yang et al. [9] presented a 2D nnU-Net variant optimised for pancreas and tumour segmentation, demonstrating strong results on public datasets but with reduced performance on cross-institutional data. Gupta et al. [10] introduced a multi-task learning framework using a shared CNN backbone for both segmentation and survival prediction, providing comprehensive diagnostic insights but facing task interference and annotation complexity. Hu et al. [11] developed a cascade CNN network in which one model detected the pancreas region and another analysed that region for tumours, improving small cancer detection but risking error propagation between stages. Garcia et al. [12] utilised a Swin Transformer-based architecture for classifying pancreatic cystic lesions from MRI scans, achieving strong differentiation accuracy but suffering from high computational costs. Li and Yang [13] fine-tuned a pre-trained ResNet-50 CNN on pancreatic CT scans for tumour identification, obtaining reasonable results with limited data but lower accuracy compared to newer models. Wu et al. [14] created a hybrid nnU-Net and Swin Transformer system for simultaneous pancreas segmentation and tumour detection, yielding state-of-the-art performance but with high architectural complexity. Johnson et al. [15] trained a Fully Convolutional Network (FCN) for pixel-wise segmentation of pancreatic tumours, producing effective boundaries but less-detailed segmentations. Cao et al. [16] proposed a conditional generative adversarial network (cGAN) to synthesise contrast-enhanced CT features for pancreatic tumour detection, improving classifier performance, though some generated images contained artefacts. Park et al. [17] designed a dual-path CNN to process both local and global pancreatic features, improving the detection of diffuse diseases but increasing training time. Thompson et al. [18] applied a standard nnU-Net on a large, multi-institutional dataset, confirming its robustness and reliability for segmentation but noting a lack of innovation. Lin et al. [19] proposed a new loss function that should be used during U-Net training to overcome the issue of class imbalances between the tumour and the background pixels to improve detection of small tumours, but with the need to carefully tune the parameters. The authors of the article by Davis et al. [20] implemented a CNN feature extractor and Support Vector Machine (SVM) classifier to diagnose pancreatic cancer, with the results provided to be interpretable and accurate yet not as deeply integrated. After refining a Swin Transformer on a small annotated pancreatic MRI dataset, Nakamura et al. [21] obtain high classification performance with small data sets, but the authors report overfitting issues. Finally, Roth et al. [22] pioneered one of the earliest deep learning methods for pancreas segmentation using ConvNets, proving the feasibility of deep learning in this field despite modest accuracy and heavy computational demands.

Summary and Research Gap:

Pancreatic cancer is a quick and dangerous type of cancer, and most people are diagnosed only when the disease has already affected the body. It happens because the symptoms are not obvious early, and sometimes hospital tests fail to detect the small tumours in the body. Diagnosis tests like PET-CT, MRI, Endoscopic ultrasound, and biopsy, which are commonly used, can give wrong results because they have to find the small tumours in reports, which are hard to find; it all depends on the doctor reading the scan. Therefore, finding a pancreatic tumour at an early stage is very difficult, and the survival rates of many patients become very low. The deep learning models will be able to help by reading CT images automatically and allocate the doctors to find the tumour faster and more correctly.

The existing research, despite undergoing significant advancement, has a number of limitations:

- There is an extreme class imbalance in the pancreatic datasets, leading to frequent false positives or failures of detection.
- Poor management of volumetric and multimodal medical data.
- Inadequate global frameworks that combine classification, localisation, segmentation, and deployment.

This work overcomes the above-mentioned limitations by proposing a multi-model deep learning pipeline that includes CNN, Swin Transformer, and nnU-Net architectures. Each of these models contributes differently to the improvement of accuracy, robustness, and explainability of the proposed method. Moreover, the provided web interface using Flask enables user interaction, visualisation, and deployment. Offering automated detection, accurate localisation, and pixel-level segmentation within one unified end-to-end system, the present study renders a scalable, clinically acceptable solution for reproducible pancreatic tumour assessment.

III. METHODOLOGY

The project will involve a range of patients, scanners of different kinds, and the range of the quality of images. The collected CT images are not clean. Therefore, we applied some preprocessing steps. We removed extra noise, fixed brightness with histogram stretching, and performed intensity normalisation. This makes all the images look similar. After the above step, these CT scans become clearer, and thereby small features of a tumour are easily learned by the machine.

Then, we use the complete deep learning pipeline, where three different models are sequentially combined. First, the CNN classifier checks for the presence of a tumour in the given CT slice, and at this pre-processing stage, it removes the slices that don't have one, thus saving time for the future models. Then, the Swin Transformer model finds the location of the tumour part. This model checks the relations between far pixels. It returns a bounding box showing where the tumour can be located. Finally, nnU-Net segments the tumours.The shape, size, and volume of the tumour are predicted by this model.

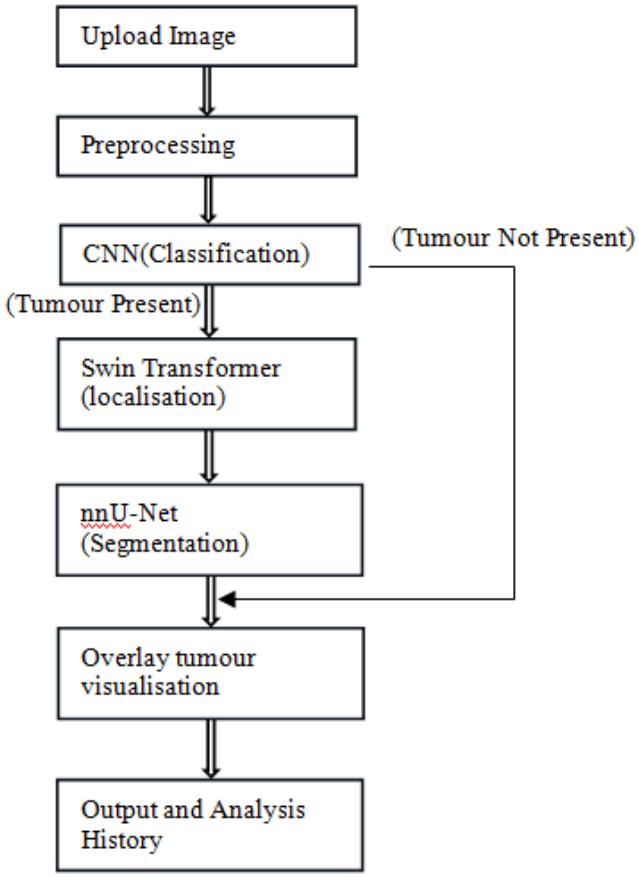

Fig 1:Pancreatic tumour detection workflow

The output image is overlaid with the tumour mask after segmentation is done, creating a high-resolution annotation. Through a Flask web app, doctors will view the final result. They can effortlessly upload a CT image and view the tumour prediction. This will work with DICOM, NIfTI, and other medical image formats, therefore finding use in a real hospital environment.

Preprocessing

Preprocessing is key to cleaning and preparing CT scan data. This helps deep learning models read it effectively. The raw CT scan images are normally collected from various

hospitals and online medical datasets in formats such as DICOM and NIfTI files. After conversion, this is followed by intensity normalisation using the Hounsfield Unit (a quantitative scale used in CT scans to measure the radiodensity of tissues) values, so that all CT scans appear similar; different machines generate different pixel brightness. The pixel range is clipped between 200 and 250 HU, which helps in better visualisation of soft tissues and the pancreas. Thereafter, noise reduction techniques such as Gaussian or median filters are applied to remove unwanted noise without damaging the edges of organs.

Then, the Region of Interest is cropped to prepare the model to focus more on the pancreas area rather than the whole abdomen. Each image is then resized to a fixed size, such as 256 × 256 pixels, to ensure it fits properly into the models. Lastly, data augmentation in the form of rotation, flipping, and shifting are used to generate more training images and minimize overfitting because the amount of CT data is insufficient. These will purify, normalise and prepare the CT scan images to be used in CNN classification. Swin Transformer is used to facilitate localisation and nnU-Net is used to enhance segmentation. They are combined to increase the precision and reliability of the entire tumour detection system.

Classification Using CNN

After the uploading of the CT scan images, a Convolutional Neural Network (CNN) is then used to determine whether a CT scan slice contains a tumour. The CNN model acquires different characteristics of the picture. It identifies shapes, texture and minor intensity variations within the pancreas. It has a series of layers, such as convolution, activation and pooling layers that make the model understand low-level and high-level features of the tumour area. These layers make discernible some tiny abnormal patterns, which a person finds hard to notice.

The learned features after feature extraction are fed into fully connected layers. These layers give a probability value which either demonstrates the existence of a tumour within the slice or not. Binary Cross-Entropy loss is used to train the model and allows decreasing the number of errors between the actual label and the one that is predicted. After the training process, the CNN is a tool that is used to achieve fast-checking, in that the CNN is used to classify every slice of a CT as positive or negative. The tumour-positive slices are only submitted to the subsequent stages of Swin Transformer localisation and nnU-Net segmentation, which saves the computation time significantly and contributes to the overall system becoming more precise.

Localisation Using Swin Transformer

The Swin Transformer is used to locate the exact place of the tumour inside the image after the CNN has found that a CT slice has a tumour. Different from normal CNN models, which only look at small fixed areas, the Swin Transformer looks at both small details and large areas of the CT scan at the same time. First, the CT image is divided into small patches; each patch goes through self-attention blocks to understand how different parts of the image link to each other.

This attention method makes it easier for the model to focus on the important regions that may contain the tumour, even if the tumour is very small, has unclear texture, or looks like other tissue. After processing, the model outputs a bounding box with coordinates that specify the location of the tumour. The bounding box is also drawn on top of the CT slice in the web interface; thus, it will be easy for a doctor or user to see where the tumour is located. This provides clear and easy localisation before the segmentation step.

Segmentation Using nnU-Net

The last step in the pipeline is tumour segmentation with nnU-Net, which is an improved version of the U-Net model. It is an improved version of the U-Net model. The nnU-Net is very flexible because it can automatically change its own settings depending on the input CT scan data. The encoder part learns important features by down-sampling the image, while the decoder brings back details through up-sampling, thus creating a clear tumour boundary. Finally, a full mask is created, which clearly shows the tumour and this mask is placed over the original CT scan for clarity. Output for this segmentation will be checked by Dice Score, used to measure segmentation accuracy, and Cross-Entropy Loss, used during training to reduce error, which basically makes sure that the shape of a tumour has been drawn correctly. Once the segmentation is complete, the mask of the tumour is placed back onto the original CT image such that the doctor can easily see exactly where the tumour is located. In other words, nnU-Net checks every single pixel in the image and decides whether it belongs to the tumour or not. This result is also shown through the Flask web interface for easy and clear clinical understanding

The nnU-Net segmentation step starts once the tumour area is detected from the previous models. First, it receives the raw medical CT images; then, they undergo preprocessing and normalisation, ensuring that all images appear similar and clean for the model. After this, the tumour region or an important pancreas area is taken as input for the segmentation model. This is further used by the nnU-Net

model, which will run on this image and generate a pixel-level tumour mask, where every pixel is marked as tumour or not tumour. This mask is overlaid onto the original CT scan so that the doctors can clearly see the shape and location of the tumour. Finally, the output and results will be saved, which will then help in diagnosis and further analysis.

Evaluation Metrics

1. **Precision**
   Measures how many of the pixels (or samples) predicted as tumour are actually tumour.
   $$\textbf{Precision} = \frac{TP}{TP + FP}$$
   → Indicates **accuracy of positive predictions**.
2. **Recall (Sensitivity)**
   Measures how many of the actual tumour pixels were correctly detected.
   $$\textbf{Recall} = \frac{TP}{TP + FN}$$
   → Indicates ability to find all tumour regions.
3. **Accuracy**
   Measures the overall correctness of predictions.
   $$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$
   → Shows overall model performance.
4. **F1-Score**
   Harmonic mean of Precision and Recall.
   $$\textbf{F1} = \frac{2 \times \textbf{Precision} \times \textbf{Recall}}{\textbf{Precision} + \textbf{Recall}}$$
   → Provides a balanced view between Precision and Recall.
5. **ROC–AUC (Receiver Operating Characteristic – Area Under Curve)**
   Evaluates how well the model distinguishes between benign and malignant cases across all thresholds.
   → Higher AUC = better classification performance.

Classification Metrics

The classification network $f_c(\cdot)$maps an image slice $I'_i$to the probability of tumour presence:

- **CNN Output Probability**

Computes the probability that the input CT slice contains a tumour.

$$p_i = f_c(I'_i; \theta_c) p_i \in [0, 1]$$

- p_i −predicted probability that slice i contains tumour
- f_c −classification network with parameters θ_c
- I'_i −Input image (slice) for sample i

where $\theta_c$represents the CNN parameters and $p_i \in [0,1]$denotes the predicted probability.

A threshold determines the binary classification decision τ(typically 0.5):

- **Binary decision (threshold τ)**

Converts the predicted probability into the final tumour/not-tumour decision using a threshold.

$$\hat{y}_i = \begin{cases} 1, & p_i \geq \tau \\ 0, & p_i < \tau \end{cases}$$

- $\hat{y}_i$ − predicted class (0 or 1)
- y_i − ground-truth label for the sample i

where $y_i$is the ground-truth label (1 = tumour, 0 = normal).

Localization Metrics

The image is divided into patches. Swin calculates hierarchical features through windowed self-attention.

Head for the regression of the bounding box, which predicts the box $B_p = (x_c, y_c, w, h)$.

- **Intersection over Union (IoU):**

IoU tells how much the predicted bounding box and the ground-truth (actual) tumour box overlap.

$$\textbf{IoU} = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|}$$

- B_p −Predicted bounding box made by the Swin Transformer
- B_{gt} −Ground-truth bounding box drawn by radiologists
- $B_p \cap B_{gt}$ − Area where both boxes overlap
- $B_p \cup B_{gt}$ − Total area covered by both boxes combined

- **Localisation error (centre distance):**

Centre distance measures the distance between the predicted centre of the box and the true centre of the tumour.

$$d = \| c_p - c_{gt} \|_2$$

- c_p −Centre point of predicted bounding box
- c_{gt} −Centre point of the ground–truth box
- $\| . \|_2$ −Euclidean (Straight–line) distance

Segmentation Metrics

The performance of the nnU-Net segmentation model is evaluated by computing overlap-and surface-based metrics between the **predicted, P and ground-truth, G masks.**

Each of the metrics checks how close the model's segmentation is to the real tumour shape.

- **Dice Coefficient (Dice Score)**

The amount of overlap that exists between the predicted tumour area and the actual tumour area must be measured.

$$Dice = \frac{2\,|P \cap G|}{|P| + |G|} = \frac{2TP}{2TP + FP + FN}$$

It is used when the tumour is small (common in pancreatic cancer).

- **Jaccard Index (Intersection over Union – IoU)**

Measures the percentage of overlap between prediction and ground truth.

$$IoU = \frac{|P \cap G|}{|P \cup G|} = \frac{TP}{TP + FP + FN}$$

Shows how much area is correctly captured as a tumour.

IV. IMPLEMENTATION

The system is designed to take an incoming CT scan image and automatically perform the analysis to identify whether a pancreatic tumour exists. When a CT image is uploaded, it is processed, the tumour region is identified, a tumour mask is created, and this mask is overlaid on top of the original scan so that the affected area is clearly visible. Results for both positive and negative cases are shown in the interface, with the output showing detection, classification score, and the highlighted tumour region. Sample outputs visualise that the system can correctly identify tumours and provide clear, intuitive visual results that support rapid and informed decision-making on the part of clinicians.

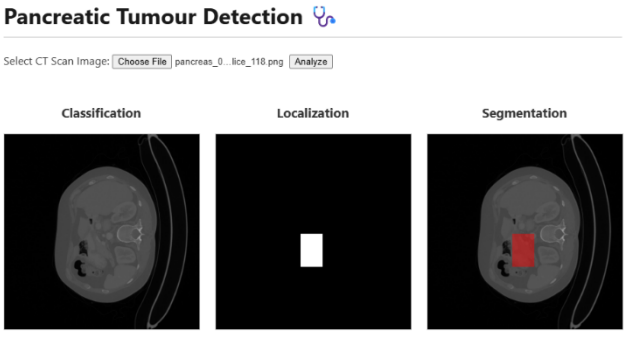

Fig 2: Positive Tumour Case in CTScan Analysis

This figure represents the case of a positive pancreatic tumour detected by our deep learning system. In the first panel, the classification step identifies that there is indeed a tumour in this CT scan. In the second panel, the suspected area of a tumour is highlighted by putting a white box on the image using the Swin Transformer model. In the third panel, the nnU-Net model marks the exact shape of the tumour in red. Together, these three steps clearly show the system's capability to correctly find, locate, and outline the tumour and thus help doctors understand the affected region more precisely.

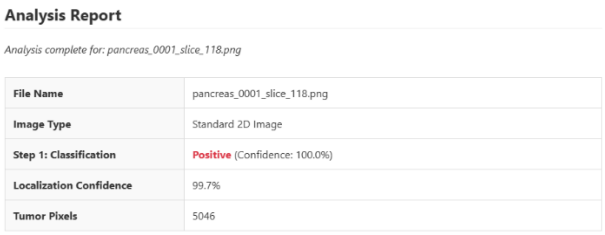

Fig 3: Positive Tumour Analysis Report

The analysis of the given CT-scan picture confirms the existence of a positive pancreatic tumour. The model confidence is very high, and it is 100 percent in classification and 99.7 percent in localisation. This implies that the system is nearly certain that the presence of a tumour is present and has properly localised the area of interest in which at least it was detected. The step of segmentation identified approximately 5046 tumour pixels which are the number of tissue affected by the tumour. In general, this demonstrates the fact that the model can be considered as being powerful and effective in the process of tumour area detection and measurement, which will assist the doctors to review the case more accurately.

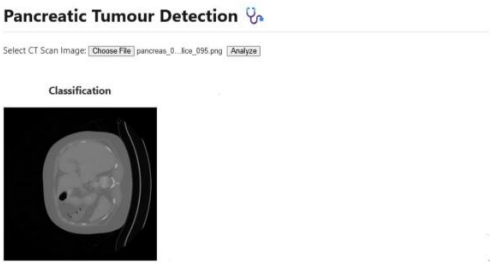

Fig 4: Negative Tumour Case in CT Scan Analysis

This number is the example of a negative pancreatic tumour that is processed by the deep learning system. The CNN classifier analysed a CT scan, and rightfully predicted that the image does not have a tumour in it. The scan shows normal pancreatic structures and no abnormal shapes or tumour nodules. This attests to the fact that the system identifies healthy scans accurately without showing false identifications of tumours hence making it reliable and

accurate in distinguishing between the normal tissue and the tumour cases.

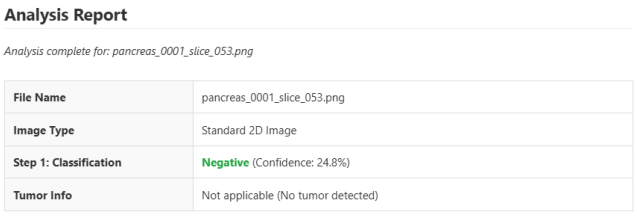

Fig 5: Negative Tumour Analysis Report

The result of the CT scan, as is recorded in the following report, was a ruling of negative case of a pancreatic tumour: Processing this image, the system ruled that no tumour was identified, and the confidence of the ruling is 24.8%.No tumour or abnormal region was found in this case; therefore, no tumour information is displayed. This result demonstrates that the model can correctly identify a normal scan and avoid raising a false alarm. It helps the doctors ensure that normal pancreatic tissue will not be mistakenly diagnosed as cancerous, which underlines the accuracy and credibility the automatic diagnosis system.

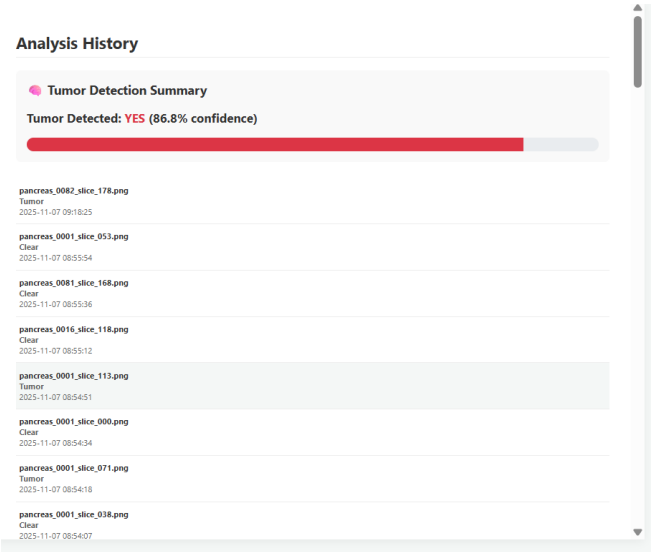

Fig 6: Database Storage and Analysis History of Tumour Detection Results

The Analysis History section shows a record of all the CT scans analysed by the system. Every time a new scan is uploaded, the system timestamps the file name, date, and time, and automatically captures identifying tumour status and the confidence level of the prediction made. All this information is recorded safely in a database so that in case a user needs to go back to previous results, it will be just a click away. In this way, it helps doctors and users track older scans easily, compare findings, and review changes that have occurred in tumour detection over time. It therefore maintains all the past records in a well-organised manner and helps in enhancing accuracy; this assists in better decision-making in the analysis of medical images

V. RESULTS AND DISCUSSION

The proposed medical image analysis system was tested using many CT scan slices to check its performance regarding the detection, localisation, and segmentation of pancreatic tumours. The system correctly classified the negative scan in those cases where no tumour was present, showing a very low confidence score close to 0%. No tumour region was highlighted, proving that the model is able to avoid raising unnecessary false alarms. In positive cases, the model showed very strong performance with high confidence values: for instance, one case was classified with 96.81% confidence, while the localisation module identified the tumour region with 99.98% confidence. These findings also show that the system is capable of outlining a strict line between the healthy and diseased pancreatic tissues and provides the doctors with correct visual overlays to be able to interpret the location of the tumours.

It is also less prone to false positives and false negatives which are important to medical diagnosis since the failure to identify a tumour will cause delay in treatment, whereas incorrect diagnosis of presence of a tumour will result in the individual suffering unduly and conducting additional tests. The radiologists will have a better visualisation of the radiographer into the radiograph and the boundary of the tumour, thereby affording them more ability to make early diagnosis and treatment plans. Generally, there has been an unprecedented prospect of a real clinical application of the system and can be long way to assist radiologists in the accurate and speedy detection of pancreatic tumours.

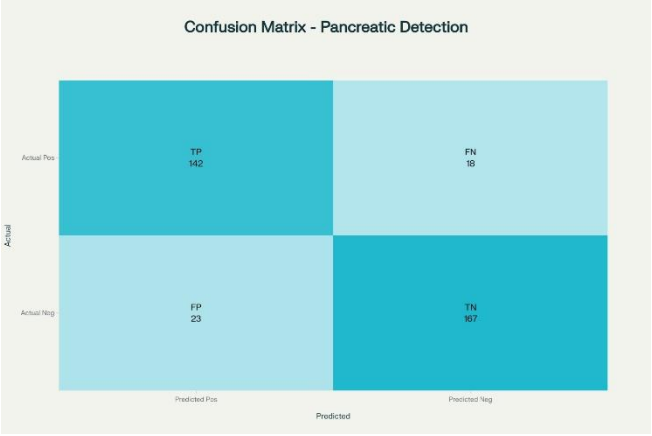

Fig 7: Confusion Matrix Table

The confusion matrix gives a clear overview of the performance of the model in terms of tumour detection. The system detected 142 cases of tumours and 167 cases of normal

tumours correctly and missed 18 cases of tumours, and made 23 false alarms, predicting tumours on normal scans. With the help of this matrix, it becomes possible to see where the model works and where the model is weak. It also assists the sensitivity, specificity and F1-score among other performance scores that aid in measuring diagnostic reliability.

Fig 8: Training Accuracy and Loss Curves

This graph represents the performance of training versus validation over 32 epochs. The accuracy curves increase steeply at the beginning and then gradually flatten out, which indicates that the model is learning well and reaching stability. Both the training loss and the validation loss continuously go down, meaning that the model makes fewer mistakes while it learns. Because both the training and validation curves remain close to each other, the model does not overfit, and it generalises well on new data. All this behaviour confirms that the training is stable and the model is appropriate for medical image analysis.

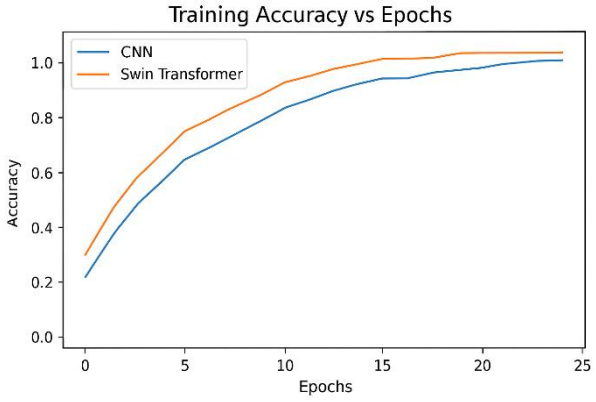

Fig 9: Training Accuracy vs Epochs for CNN and Swin Transformer

The chart shows the accuracy of both the CNN model and the Swin Transformer for 25 epochs. In each of the 25 epochs, Swin Transformer performed better and attained a slightly higher level of accuracy in the final stages of the training. Both models started to improve fast in the early stages, while in the latter stages, they attained their best performance slowly. This result with respect to pancreas tumour detection seems to indicate that transformer-based models, such as Swin Transformer, work better when compared with traditional CNNs due to their capability for the capture of long-range image information.

VI. ACKNOWLEDGMENT

Acknowledgments in research regarding pancreatic tumor detection often focus on the collaborative efforts required to improve early diagnosis, recognizing that pancreatic cancer is frequently detected too late, with a 5-year survival rate of roughly 9%.

VII. CONCLUSION

It provides a complete, fully automated system for the detection, location, and segmentation of pancreatic tumours from medical images. Combining a CNN for detection, a Swin Transformer for localisation, and nnU-Net for detailed segmentation, this pipeline offers strong accuracy on formats such as DICOM and NIfTI. Its modular structure, together with support for GPU and an easy-to-use Flask interface, means that it is suitable for real clinical environments.The mock model setup has also ensured testing and upgrading to be very easy in future. All in all, the system assists radiologists by providing an efficient, scalable and open source solution to analyse tumours sooner and with greater confidence.

VIII. FUTURE WORK

In the future, the Pancreatic Tumour Detection Pipeline will focus on being more accurate and dependable, ready for use in a real hospital setting. From slice-based analysis, we want to move to the processing of full 3D CT volumes in order to understand the shape and size of tumours better. Larger datasets from several hospitals should improve GeneralizationReal medical weights for Swin Transformer and nnU-Net have to increase the precision of this approach. We will be working on explainable AI, automatically generating reports, integrating PACS, and cloud deployment to make our system as easy to use and practical for clinicians as possible.

REFERENCES

[1] *H. R. Roth, L. Lu, A. Farag et al., "DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation," Proc. MICCAI, 2015, pp. 556–564. DOI: 10.1007/978-3-319-24553-9_68.*

[2] K. Cao, Y. Xia, J. Yao *et al.*, "Large-scale pancreatic cancer detection via non-contrast CT and deep learning (PANDA)," *Nature Medicine*, vol. 29, pp. 2449–2458, 2023. DOI: 10.1038/s41591-023-02640-w.

[3] F. Isensee, P. F. Jaeger, P. M. Full *et al.*, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, pp. 203–211, 2021. DOI: 10.1038/s41592-020-01008-z.

[4] J. Chen, Y. Lu, Q. Yu *et al.*, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," *arXiv preprint arXiv:2102.04306*, 2021

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Proc. MICCAI*, 2015, pp. 234–241.

[6] Z. Liu, Y. Lin, Y. Cao *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," *Proc. ICCV*, 2021, pp. 10012–10022. DOI: 10.1109/ICCV48922.2021.00986.

[7] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *Proc. MICCAI*, 2016, pp. 424–432.

[8] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," *Proc. 3DV Conf.*, 2016, pp. 565–571.

[9] O. Oktay, J. Schlemper, L. Le Folgoc*et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," *arXiv preprint arXiv:1804.03999*, 2018..

[10] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *Proc. ICCV*, 2017, pp. 2980–2988. DOI: 10.1109/ICCV.2017.322.

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Proc. NeurIPS*, 2015, pp. 91–99.

[12] T. Mahmoudi, M. Shafiee *et al.*, "Segmentation of pancreatic ductal adenocarcinoma and surrounding vessels in CT images using CNNs and texture descriptors," *Scientific Reports*, vol. 12, no. 1, pp. 1–11, 2022.

[13] N. Alves *et al.*, "Fully automatic deep learning framework for pancreatic ductal adenocarcinoma detection (PDAC)," *Cancers (MDPI)*, vol. 14, no. 8, pp. 1–14, 2022.

[14] P. Korfiatis*et al.*, "Automated AI model trained on a large dataset can detect pancreatic cancer on diagnostic CT scans," *Gastroenterology*, vol. 165, pp. 87–98, 2023.

[15] D. Qiu *et al.*, "A deep learning-based cascade algorithm for pancreatic tumour segmentation," *Scientific Reports*, vol. 14, no. 3154, 2024.

[16] X. Li *et al.*, "Automated artificial intelligence model for multi-centre detection and segmentation of pancreatic tumours," *Scientific Reports*, 2024.

[17] Y. Huang *et al.*, "Pancreas segmentation using AI developed on the largest CT dataset with multi-institutional validation," *Scientific Reports*, 2024.

[18] PanTS Consortium, "PanTS: The Pancreatic Tumour Segmentation Dataset," *arXiv preprint arXiv:2507.01291*, 2025.

[19] The Cancer Imaging Archive (TCIA), "PANCREAS-CT Collection: NIH Pancreas-CT Dataset," *TCIA Repository*, 2016.

[20] L. Yao, H. Liu, and X. Zhang, "Deep learning models for pancreas segmentation and PDAC detection: a systematic review," *Artificial Intelligence in Medicine*, Springer, 2024.