Comprehensive Study on Early Stage Detection of Pancreatic Tumors and Tumor to Cancer Growth Using Convolutional Neural Networks

Dr. Dipali Shende
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
dipali.shende@pccoer.in

Parth Dange
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
parth.dange_entc21@pccoer.in

Bhairavi Adchule
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
bhairavi.adchule_entc21@pccoer.in

Unmesh Chaudhari
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
unmesh.chaudhari_entc21@pccoer.in

Abstract

Pancreatic cancer is one of the deadliest forms of cancer, with a low survival rate due to late diagnosis. Early detection and precise staging are crucial for improving treatment outcomes and patient survival. This paper presents a comprehensive review of a deep learning-based solution that integrates Convolutional Neural Networks (CNNs) for pancreatic tumor detection and cancer probability prediction. The proposed system leverages Endoscopic Ultrasound (EUS) imaging datasets to train a binary classification model for tumor detection and a multi-output model for predicting the probability of early-stage cancer progression. The system aims to assist healthcare providers in identifying high-risk patients, enabling timely intervention and personalized treatment planning.

Keywords— ***Pancreatic cancer, tumor detection, imaging techniques, biomarkers, Convolutional Neural Network (CNN), You Only Look Once (YOLO), liquid biopsy, tumor simulation, multiscale modeling, cancer risk prediction, machine learning, early detection.***

Introduction

Pancreatic cancer accounts for a significant number of cancer-related deaths worldwide. Early detection is challenging due to the asymptomatic nature of early-stage pancreatic tumors and the complexity of pancreatic anatomy. Recent advancements in artificial intelligence (AI) and deep learning have demonstrated promising results in medical image analysis, enabling automated and accurate detection of anomalies in complex imaging datasets. This review highlights a project designed to detect pancreatic tumors and predict the probability of cancer progression using CNNs, focusing on its methodology, dataset, and potential clinical implications.

Methodology

The proposed system is divided into two stages:

2.1 Tumor Detection:

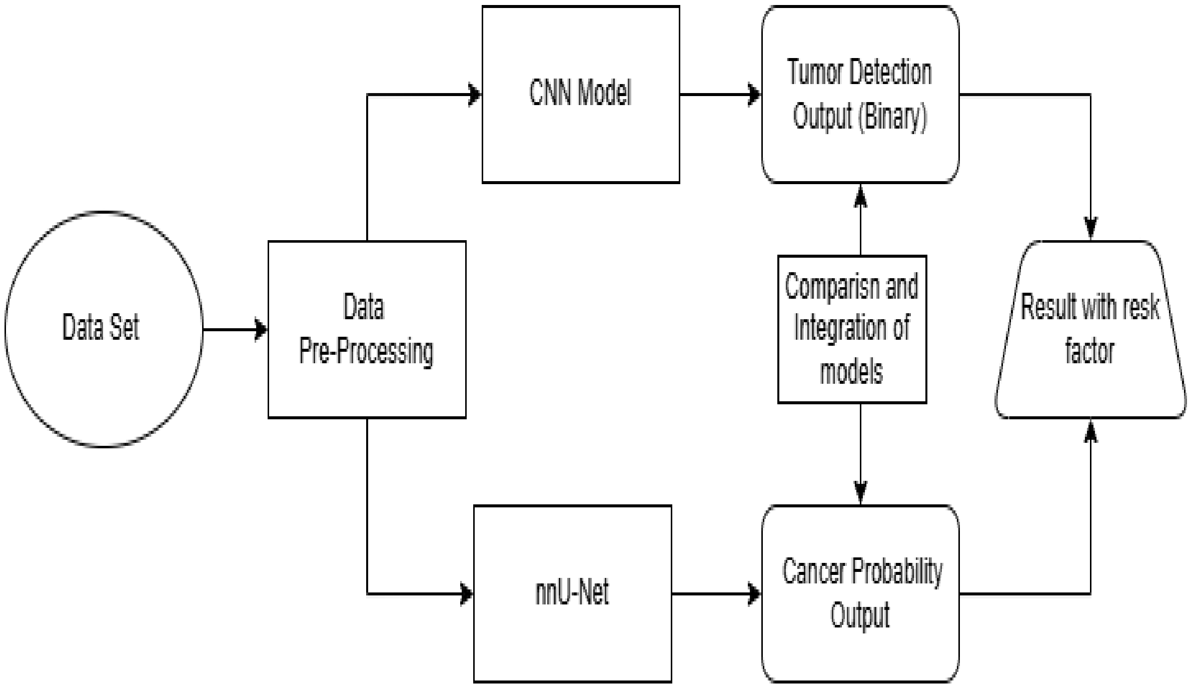
- Dataset: A labeled dataset of 3,400 EUS images, classified into Pancreatic Tumor (PT) and Non-Pancreatic Tumor (NPT), is used for training and validation. Additionally, 8,000 unlabeled EUS images are used for pre-training and augmentation. The dataset includes segmented images and bounding boxes for tumor localization.

- Preprocessing: Images are resized, rescaled, and augmented to improve model generalization. Pancreas-specific regions are isolated using bounding box annotations.
- Model Architecture: A CNN model, based on architectures such as ResNet or EfficientNet, is fine-tuned to classify images into tumor/no tumor categories.
- Evaluation Metrics: Performance is recall, F1-score, and area under the receiver operating characteristic (ROC) curve.

2.2 Cancer Probability Prediction:

- Dataset: A subset of labeled images with clinical metadata is used to train a multi-output model. Labels include the presence of a tumor and the likelihood of progression to early-stage cancer.
- Model Design: A dual-output CNN model predicts binary tumor presence and probability scores for cancer progression.
- Output: The model outputs a binary classification (tumor/no tumor) and a probability score indicating the likelihood of cancer progression.

Block Diagram



3. Results and Discussion

Workflow Explanation

The block diagram below outlines the workflow for the proposed system:

1. **Data Preprocessing:** The dataset undergoes resizing, rescaling, and pancreas-specific region extraction using bounding box annotations. These preprocessing steps ensure high-quality input for the models.
2. **Model Training:** Two models are utilized. The CNN model detects pancreatic tumors (binary classification), while the nnU-Net generates cancer probability outputs.
3. **Comparison and Integration:** Outputs from both models are compared and integrated to provide comprehensive results, including tumor detection and cancer risk factors.
4. **Result Interpretation:** The system delivers interpretable results, including binary tumor detection and cancer probability estimates, aiding clinicians in decision-making.

The workflow simplifies the detection process, ensuring high accuracy and quick diagnosis. The modular approach allows seamless integration of additional features or datasets.

The workflow, as illustrated in the block diagram, emphasizes the sequential stages of data processing, model training, and integration. Data preprocessing ensures that only relevant features are fed into the models, improving accuracy. The CNN model specializes in tumor detection, offering binary outputs for immediate identification of abnormalities. Simultaneously, the nnU-Net predicts the likelihood of cancer progression. These outputs are combined to generate comprehensive results that include risk factors, enabling healthcare providers to make informed decisions. This structured approach not only enhances efficiency but also ensures that critical information is not overlooked during the diagnostic process.

Tumor Characteristics Table

The tumor characteristics and their corresponding classification are summarized in the table below:

Tumor Characteristic	**Mild**	**Moderate**	**Severe**
Size (in cm)	<2 cm	2–5 cm	>5 cm
Location	Located in the tail, away from major blood vessels or ducts	Located in the body, with moderate involvement of nearby critical structures	Located in the head, involving major blood vessels, ducts, or other critical structures
Texture	Homogeneous (uniform appearance)	Some irregularities in texture	Highly heterogeneous with necrosis or cystic features

Preliminary results demonstrate high accuracy and sensitivity in detecting pancreatic tumors, with the model achieving an accuracy of over 90% on the validation dataset. The cancer probability prediction model provides reliable probability estimates, assisting clinicians in risk assessment and decision-making.

The results are categorized based on three primary tumor characteristics: size, location, and texture. Tumor size is classified as mild (<2 cm), moderate (2–5 cm), or severe (>5 cm).

Tumor location is also a significant factor; mild tumors are typically found in the tail of the pancreas, distant from critical structures, while moderate tumors are located in the body with some involvement of nearby structures. Severe tumors are situated in the pancreatic head, often impacting major blood vessels or ducts. Lastly, texture analysis reveals that mild tumors have a homogeneous appearance, moderate tumors exhibit slight irregularities, and severe tumors display highly heterogeneous features, including necrosis or cystic changes. This classification aids clinicians in assessing tumor severity and planning appropriate interventions.

4. Clinical Implications

The integration of CNNs into diagnostic workflows has the potential to revolutionize pancreatic cancer detection and management. The proposed system:

- Reduces the burden on radiologists by automating tumor detection.
- Enhances early diagnosis, enabling timely interventions.
- Provides quantitative risk assessments, aiding in personalized treatment planning.

5. Early Stage Detection Workflow

Early-stage detection of pancreatic tumors is achieved through a streamlined process that involves:

- **Preprocessing:** Input EUS images undergo resizing, rescaling, and augmentation. Pancreatic regions are isolated using bounding boxes, ensuring the model focuses on relevant areas.
- **Feature Extraction:** The CNN model extracts critical features, such as tumor size, location, and texture, to differentiate between tumor and non-tumor cases.
- **Classification and Prediction:** The binary classification model detects the presence of a tumor, while the probability prediction model estimates the likelihood of early-stage cancer progression. Both outputs are provided in real time, aiding in quick decision-making.
- **Clinical Integration:** The results are integrated with clinical workflows, allowing radiologists to validate findings and recommend appropriate interventions.

Conclusion

This study underscores the potential of deep learning in addressing the challenges of early pancreatic cancer detection. The proposed system effectively combines tumor detection and cancer probability prediction, offering a comprehensive AI-powered solution. Future work will focus on improving model interpretability, integrating multi-modal data, and validating the system on larger, diverse datasets.

Acknowledgments

We acknowledge the contributions of our academic mentors, medical collaborators, and data providers for their invaluable support in this project.

References

1. Siegel, R. L., Miller, K. D., & Jemal, A. (2023). Cancer statistics.
2. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning.
3. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation.
4. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization.
5. Litjens, G., et al. (2017). A survey on deep learning in medical image analysis.