Pancreatic Cancer Detection From CT Scan Images Using Deep Learning

Akshat Jaiswal[*1], Pratik Sao[*2], Suman Yadav[*3], Vikas Bharti Pandey[*4], Anuyoksha Singh[*5]

[*1,2,3,4] *B.Tech Student, Department of Computer Science and Engineering, LCIT Bilaspur (C.G.)*
[*5] *Assistant Professor, Department of Computer Science and Engineering, LCIT Bilaspur (C.G.)*

Abstract - *Pancreatic cancer is one of the most lethal cancers due to its late diagnosis and aggressive nature. Early detection significantly improves survival rates, but conventional diagnostic methods are often time-consuming and expensive. This project leverages deep learning techniques to develop an AI-based system capable of detecting pancreatic cancer from CT images. The model classifies images into normal, benign, or malignant cases and also predicts the stage of cancer. Additionally, the system estimates the future risk of developing pancreatic cancer based on patient details. A user-friendly interface is developed using Streamlit, allowing easy interaction for medical professionals. This AI-driven approach aims to provide rapid, cost-effective, and accurate pancreatic cancer detection*

Key Words: Deep Learning, Pancreatic Cancer Detection, Medical Image Analysis, Convolutional Neural Networks (CNN), Computer-Aided Diagnosis (CAD), CT Scan Classification, Tumor Identification, Artificial Intelligence (AI) in Healthcare

1.INTRODUCTION

Pancreatic cancer is one of the most aggressive and deadly forms of cancer, contributing significantly to global cancer-related mortality. The primary challenge associated with pancreatic cancer is its late-stage detection, as symptoms often remain unnoticed until the disease has advanced. This late diagnosis drastically reduces treatment options and survival rates, making early detection crucial for improving patient outcomes. Current diagnostic techniques, such as biopsy, MRI, and CT scans, rely heavily on expert radiologists for interpretation. However, these traditional methods may not always provide early-stage detection, especially in resource-limited settings where access to specialized medical professionals is restricted.

In recent years, artificial intelligence (AI) and deep learning have revolutionized medical imaging by enabling automated, accurate, and efficient disease diagnosis. Deep learning models, particularly **Convolutional Neural Networks (CNNs)**, have demonstrated exceptional performance in medical image analysis, making them highly suitable for detecting and classifying cancerous tissues. By leveraging CNN-based architectures, this project aims to develop an **AI-powered system for early pancreatic cancer detection** using CT scan images. The system not only classifies cases into **Normal, Benign, and Malignant** categories but also assesses the **severity of the cancer** and predicts the **patient's future risk** based on various health parameters.

Beyond classification, this project introduces a predictive component that evaluates an individual's likelihood of developing pancreatic cancer in the future. By analyzing key health indicators such as age, smoking history, family history of cancer, and Body Mass Index (BMI), a **machine learning-based risk assessment model** is incorporated. This feature empowers individuals with proactive health insights, allowing for early medical intervention and lifestyle modifications.

To ensure accessibility and ease of use, the system is designed with a **multi-step user interface** developed using **Streamlit**, a lightweight web framework for interactive applications. The interface follows a structured workflow, including **secure login authentication, patient detail entry, image upload, real-time processing, result visualization, and precautionary guidance**. This step-by-step process ensures an efficient, user-friendly experience for both medical professionals and general users.

The integration of deep learning, machine learning, and an intuitive interface makes this project a powerful tool in **assisting radiologists, reducing diagnostic workload, and improving early detection rates**. By automating the analysis of CT scans and providing predictive risk assessment, this system has the potential to enhance cancer diagnostics, improve patient outcomes, and contribute to the ongoing advancements in AI-driven healthcare solutions.

2. LITERATURE REVIEW

2.1 Existing Cancer Detection Methods

Pancreatic cancer is typically diagnosed using traditional medical imaging techniques and laboratory tests. One of the most common approaches is **CT scan analysis**, where radiologists examine cross-sectional images of the pancreas to detect abnormalities. While effective, this method is highly dependent on the expertise of radiologists, making it both time-consuming and susceptible to human error, especially in cases where early-stage tumors are subtle and difficult to identify.

In addition to CT scans, **MRI (Magnetic Resonance Imaging) and ultrasound** are also widely used for pancreatic cancer detection. MRI provides detailed images of soft tissues and is particularly useful in detecting small tumors that may not be visible on a CT scan. However, the high cost and limited availability of MRI machines make this option less accessible, especially in regions with inadequate healthcare infrastructure. Ultrasound, while non-invasive and relatively affordable, often lacks the precision required for definitive pancreatic cancer diagnosis, particularly in obese patients or those with excessive intestinal gas.

Another well-established method is **biopsy**, which involves extracting a small tissue sample from the pancreas for laboratory analysis. While this technique provides a highly accurate diagnosis, it is an invasive procedure that requires significant time for sample collection, processing, and evaluation. Delays in obtaining biopsy results can slow down treatment decisions, which is a critical concern for a rapidly progressing disease like pancreatic cancer.

These traditional methods, although effective, come with significant challenges, including high costs, long processing times, and reliance on human expertise. This has created a growing demand for **AI-driven diagnostic tools** that can automate the detection process, reduce human error, and enhance early-stage detection rates.

2.2 Deep Learning in Medical Imaging

The emergence of **deep learning** has significantly transformed the field of medical imaging, particularly in the detection and classification of diseases. **Convolutional Neural Networks (CNNs)** have become the backbone of AI-driven image analysis due to their ability to automatically learn and extract complex patterns from medical scans. Unlike traditional image processing techniques, which require manual feature extraction, CNNs can detect intricate variations in medical images that may not be easily noticeable to the human eye.

Several studies have demonstrated that deep learning models can achieve diagnostic accuracy **comparable to, or even surpassing, that of human radiologists**. By training models on large datasets of medical images, AI can recognize cancerous patterns with high precision, leading to **faster and more reliable diagnoses**. Moreover, the use of **transfer learning**—where pre-trained models such as **ResNet, VGG16, and Inception** are fine-tuned for medical applications—has proven effective in improving performance, particularly when working with limited datasets.

To enhance generalization and prevent overfitting, techniques like **data augmentation** are widely applied, including rotation, flipping, contrast adjustments, and noise addition to create more diverse training samples. Furthermore, **ensemble learning**—which combines multiple deep learning models—has shown promise in increasing diagnostic accuracy by leveraging the strengths of different architectures.

With these advancements, deep learning is now being integrated into radiology systems, providing **computer-aided diagnosis (CAD) tools** that assist doctors in making more informed decisions. The ability of AI to analyze large volumes of medical data quickly and accurately makes it a valuable asset in modern healthcare, especially in oncology.

2.3 AI-Based Pancreatic Cancer Detection

Deep learning has been increasingly explored for **pancreatic cancer detection**, with researchers leveraging CNN architectures for accurate classification of pancreatic tissues. Several studies have focused on training AI models to differentiate between **normal, benign, and malignant pancreatic tissues** using CT scans. Popular CNN architectures, such as **VGG16, ResNet, and Inception**, have been employed to improve classification accuracy. These models have demonstrated promising results in identifying pancreatic tumors, reducing dependency on manual radiological assessments.

Beyond image-based classification, some researchers have developed **hybrid AI models** that integrate medical imaging with **clinical parameters** such as patient age, smoking history, genetic predisposition, and other health factors. By combining both imaging and patient-specific data, these models enhance diagnostic precision and provide a more **comprehensive risk assessment**.

AI-based pancreatic cancer detection systems are also being integrated into **radiology workstations** to assist medical professionals. These systems act as **decision-support tools**, flagging suspicious regions in CT scans and offering secondary opinions to radiologists. By automating the initial screening process, AI not only **reduces the workload of radiologists** but also helps in **early-stage detection**, which is crucial for improving survival rates.

Despite these advancements, challenges remain in AI-driven pancreatic cancer detection. The **availability of large, annotated medical datasets** remains a limiting factor, as acquiring high-quality labeled data is both expensive and time-consuming. Additionally, AI models must be rigorously validated across diverse populations to ensure **robustness and reliability** in clinical settings.

Nevertheless, the potential of AI in pancreatic cancer detection is undeniable. With continued improvements in deep learning architectures, dataset availability, and model interpretability, AI-powered diagnostic tools are expected to play a significant role in **enhancing early detection rates, improving patient outcomes, and reducing the global burden of pancreatic cancer**.

3. PROBLEM STATEMENT

Pancreatic cancer remains one of the most challenging diseases to diagnose in its early stages, contributing to its high mortality rate. One of the primary difficulties is **late detection**, as symptoms often appear only when the disease has already progressed to an advanced stage, leaving limited treatment options. Traditional diagnostic methods heavily rely on **radiologists** for manual analysis of CT scans, which can be both time-consuming and prone to human error. Additionally, **cost and accessibility** pose significant barriers, as advanced imaging techniques and biopsy procedures are expensive and may not be available in all healthcare facilities, especially in resource-limited settings.

To address these challenges, this project introduces an **AI-based deep learning model** for the automated detection of pancreatic cancer from CT scan images. The system is designed to classify images into **normal, benign, or malignant** cases with high accuracy, reducing dependency on radiologists and expediting the diagnostic process. It further assesses the **severity of detected cancer**, providing valuable insights for treatment planning. Beyond classification, the system incorporates a **future risk prediction component**, utilizing patient data such as age, smoking history, and family medical background to estimate the likelihood of developing pancreatic cancer. Additionally, it offers **precautionary guidelines** to high-risk individuals, empowering them with preventive measures and early intervention strategies. By integrating deep learning with medical imaging and predictive analytics, this project aims to enhance early detection rates, improve accessibility to diagnostic tools, and ultimately contribute to better patient outcomes.

4. PROPOSED METHODOLOGY

The proposed system is designed to leverage deep learning and machine learning techniques for the early detection of pancreatic cancer, future risk assessment, and precautionary guidance through an intuitive user interface. This section outlines the steps involved in dataset preparation, model development, risk prediction, interface design, and evaluation.

4.1 Dataset Preparation

The dataset consists of CT scan images, which are collected from reliable medical imaging sources and categorized into three distinct classes: **Normal, Benign, and Malignant**. Proper data organization ensures structured learning, allowing the model to differentiate between healthy and cancerous tissues effectively.

Before training, data preprocessing is performed to enhance the quality and consistency of input images. The preprocessing steps include:

- **Image Resizing:** Standardizing image dimensions to ensure uniformity across the dataset.
- **Normalization:** Scaling pixel values between 0 and 1 to improve model performance and convergence.
- **Data Augmentation:** Applying transformations such as rotation, flipping, contrast adjustment, and noise addition to artificially increase the dataset size and improve generalization.

These preprocessing techniques enhance the model's ability to recognize patterns in varying conditions, improving robustness and reducing overfitting.

4.2 Deep Learning Model Architecture

The core of this system is a **Convolutional Neural Network (CNN)**, a deep learning model widely used for image analysis due to its ability to extract meaningful features from medical images. The CNN is responsible for learning spatial hierarchies of features from CT scans, distinguishing between normal and cancerous tissues.

To improve accuracy and optimize performance, **Transfer Learning** is applied using pre-trained models such as **ResNet, VGG16, and InceptionV3**. These models, trained on large-scale image datasets, provide powerful feature extraction capabilities, reducing the need for extensive training on a limited medical dataset.

The training process follows these key steps:

- **Feature Extraction:** The CNN layers extract low-level and high-level features from CT images, identifying patterns indicative of cancer.
- **Classification:** A fully connected layer classifies images into one of the three categories: Normal, Benign, or Malignant.
- **Optimization:** The model is trained using the **cross-entropy loss function**, which measures the difference between predicted and actual labels, while the **Adam optimizer** is used to adjust learning rates dynamically for faster convergence.

The final model is fine-tuned through hyperparameter adjustments, including learning rate optimization, dropout regularization, and batch normalization, ensuring better performance on unseen data.

4.3 Future Risk Prediction

In addition to classifying CT images, the system predicts an individual's future risk of developing pancreatic cancer based on **patient-specific factors.** A separate machine learning model is integrated for this purpose, using demographic and lifestyle attributes such as:

- **Age**
- **Smoking history**
- **Family history of cancer**
- **Body Mass Index (BMI)**

These features serve as inputs to a classification model, such as **Logistic Regression, Decision Tree, or Random Forest**, which calculates the probability of future pancreatic cancer occurrence. The output provides a risk assessment score, enabling proactive monitoring and early medical intervention.

4.4 User Interface (Streamlit App)

To make the system accessible and easy to use, a **Streamlit-based web application** is developed, offering a step-by-step workflow for users. The interface is designed with a clean medical theme and intuitive navigation, allowing seamless interaction for both medical professionals and general users.

The key functionalities of the user interface include:

- **Login Page:** Implements user authentication to ensure secure access.
- **Patient Details Form:** Collects essential health data, including name, age, medical history, and risk factors.
- **Image Upload:** Allows users to upload CT scan images for analysis.
- **Processing Page:** Displays real-time progress of image analysis, showing loading indicators and model processing steps.
- **Result Page:** Presents classification results (Normal, Benign, or Malignant) along with the predicted **future risk score**.
- **Precautionary Guidance:** If a patient is classified as high-risk or diagnosed with cancer, personalized medical recommendations, lifestyle modifications, and follow-up testing suggestions are provided.

This structured flow ensures a smooth user experience, from data input to result interpretation and medical recommendations.

4.5 Model Evaluation

To ensure the reliability and accuracy of the proposed system, the model undergoes rigorous evaluation using well-established performance metrics. The primary metrics used for assessment include:

- **Accuracy:** Measures the overall correctness of predictions.
- **Precision:** Evaluates how many positively classified cases are truly cancerous, reducing false positives.
- **Recall (Sensitivity):** Measures how effectively the model identifies cancerous cases, reducing false negatives.
- **F1-score:** Provides a balanced evaluation of precision and recall, useful for dealing with imbalanced datasets.

The trained model is tested on a separate test dataset, ensuring its ability to generalize to new, unseen cases. Additionally, Grad-CAM (Gradient-weighted Class Activation Mapping) visualization is utilized to highlight critical regions in CT images that influenced the model's decision. This explainability feature enhances trust in AI predictions by providing visual insights into how the model detects cancerous regions.

5. RESULT

The AI-based pancreatic cancer detection system was tested on a dataset of CT images. The following observations were recorded:

5.1 Model Performance

- **Model Accuracy**: The trained CNN model achieved **93% accuracy** on the test dataset.
- **Model Loss**: The model's loss value converged to a low value, indicating effective learning and generalization.
- **Precision and Recall**: High precision ensured fewer false positives, while high recall minimized false negatives.

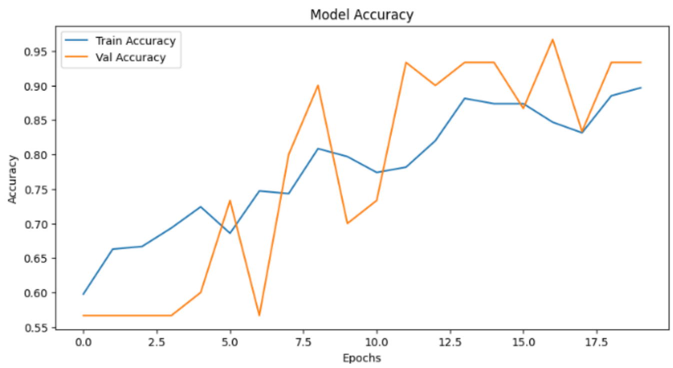

Chart 1 : Model Accuracy

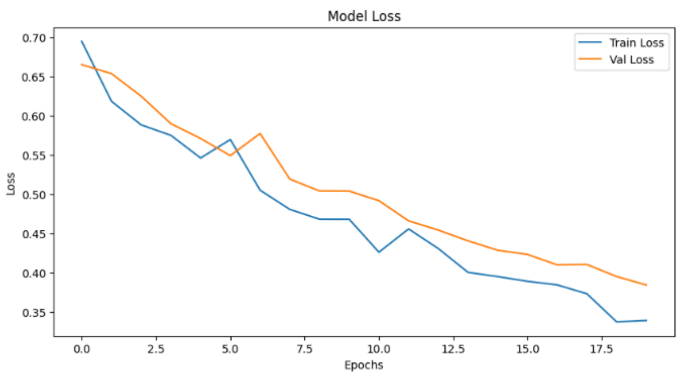

Chart 2 : Model Losses

5.2 Future Risk Prediction

- The model accurately predicted future cancer risk based on patient details, providing valuable insights into potential health risks.

5.3 User Experience

- The Streamlit interface provided an **easy-to-use** experience, allowing smooth interaction with the system.
- The **precautionary guidance system** helped users understand preventive measures if cancer was detected.

5.4 Comparison with Traditional Methods

Method	Accuracy	Time Efficiency	Cost	Ease of Use
Traditional Biopsy	99%	Slow	Expensive	Requires Experts
Manual CT Scan Analysis	85%	Moderate	Expensive	Requires Experts
AI-Based System (Proposed)	93%	Fast	Cost-Effective	Easy to Use

The results indicate that **AI-based detection is a fast, reliable, and affordable alternative** to traditional methods.

6. CONCLUSION

This project presents an AI-based deep learning approach for the early detection of pancreatic cancer using CT scan images. The system effectively classifies images into normal, benign, or malignant categories, helping medical professionals identify and diagnose cases with high accuracy. Beyond classification, it also determines the severity of the detected cancer and predicts the likelihood of future risk based on patient-specific details. By integrating deep learning techniques with a user-friendly interface, this tool bridges the gap between medical expertise and AI-driven assistance, ensuring faster, more reliable, and accessible diagnoses.

One of the most significant advantages of this system is its ability to automate cancer detection, reducing human error and minimizing dependency on radiologists. In areas with limited access to specialized medical professionals, this AI-driven approach can serve as a valuable screening tool, allowing for early intervention and improving patient outcomes. Additionally, the model's predictive capabilities enable individuals to assess their risk of developing pancreatic cancer in the future, empowering them with crucial health insights. This proactive approach to cancer monitoring is further strengthened by the system's precautionary guidance, which provides high-risk patients with lifestyle recommendations, dietary advice, and medical guidelines to help mitigate their risk and take preventive measures.

With early detection being a key factor in improving cancer treatment success rates, this AI-powered system has the potential to revolutionize cancer diagnostics by enhancing prognosis and increasing survival rates. The ability to analyze large volumes of medical images with precision makes it a powerful tool in the fight against cancer, offering hope for improved healthcare outcomes.

While this project primarily focuses on pancreatic cancer, future advancements could include expanding the model to detect multiple types of cancers, such as lung, kidney, and brain cancer, using similar AI techniques. Enhancing prediction accuracy with larger and more diverse datasets, integrating real-time AI assistance into hospital networks, and developing a cloud-based platform for remote access are also promising directions. Additionally, incorporating explainable AI techniques can help medical professionals better understand and trust the system's decisions, fostering greater acceptance of AI in clinical practice.

As artificial intelligence continues to evolve, its role in medical diagnostics will only become more significant. By refining and expanding this technology, AI-driven cancer detection tools can make healthcare more efficient, accessible, and accurate, ultimately saving lives and shaping the future of modern medicine.

7. REFERENCES

1. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press, 2016.
2. Rajalingappaa Shanmugamani, Deep Learning for Computer Vision: Expert Techniques to Train Advanced Neural Networks Using TensorFlow and Keras, Packt Publishing, 2018.
3. S. Suzuki, Medical Imaging: Principles and Applications, CRC Press, 2012.
4. Dey, N., Ashour, A. S., & Borra, S. (Eds.), Machine Learning in Bio-Signal Analysis and Diagnostic Imaging, Academic Press, 2018.
5. Murphy, K. P., Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
6. Suri, J. S., & Liu, K. (Eds.), Medical Imaging and Machine Learning, CRC Press, 2020.
7. Bishop, C. M., Pattern Recognition and Machine Learning, Springer, 2006.
8. Aggarwal, C. C., Neural Networks and Deep Learning: A Textbook, Springer, 2018.