# Project Report: Analysis and Prediction of Pancreatic Cancer Using CNN

## Introduction

Pancreatic cancer is a life-threatening disease that requires early detection for effective treatment. This project leverages a Convolutional Neural Network (CNN) model to analyze and predict pancreatic cancer using an image dataset. The model aims to provide an automated method to assist in the diagnosis process.

## Dataset

- Source: The dataset is organized into three directories:

- Training Directory: `P:/Anusha/Images/validation`

- Validation Directory: `P:/Anusha/Images/Test`

- Test Directory: `P:/Anusha/Images/Test\_set`

- Data Preprocessing:

- Images are rescaled to have pixel values in the range [0, 1].

- Data augmentation is not applied in this instance.

- Images are resized to `128 x 128` pixels.

## Model Architecture

The CNN model comprises the following layers:

- Convolutional Layers: Extract features from the input images.

- First layer: 32 filters, kernel size (3x3), activation function: ReLU.

- Second layer: 64 filters, kernel size (3x3), activation function: ReLU.

- Pooling Layers: Reduce spatial dimensions using max-pooling with a (2x2) window.

- Flatten Layer: Converts the 2D feature maps into a 1D vector.

- Fully Connected Layers:

- Dense layer with 128 units and ReLU activation.

- Dropout layer with 50% rate to prevent overfitting.

- Output layer with 1 unit and sigmoid activation for binary classification.

## Implementation Details

- Frameworks Used:

- TensorFlow/Keras for CNN implementation.

- Matplotlib for data visualization.

- Flask for frontend integration.

- Compilation Settings:

- Optimizer: Adam

- Loss Function: Binary Crossentropy

- Metrics: Accuracy

## Training and Validation

- Training Configuration:

- Batch size: 32

- Epochs: 10

- Results:

- Training Accuracy: Visualized over 10 epochs.

- Validation Accuracy: Visualized over 10 epochs.

Training and Validation Accuracy Graph:

The graph demonstrates how the accuracy improves over time, with minimal overfitting observed.

## Evaluation

- Confusion Matrix:

- The confusion matrix is generated using the test dataset.

- Labels:

- Negative: Non-cancerous samples

- Positive: Cancerous samples

- Metrics Derived:

- True Positives (TP): Correctly predicted positive cases.

- True Negatives (TN): Correctly predicted negative cases.

- False Positives (FP): Incorrectly predicted positive cases.

- False Negatives (FN): Incorrectly predicted negative cases.

Confusion Matrix Visualization:

Accuracy: The model achieved an accuracy of \*\*(insert accuracy value)\*\* on the test set.

## Conclusion

This project successfully developed a CNN-based approach for analyzing and predicting pancreatic cancer. The results demonstrate the model’s potential as a diagnostic tool, though further optimization and testing on larger datasets are recommended to enhance reliability and accuracy.

## Future Work

- Integrate data augmentation techniques to improve model generalization.

- Incorporate additional metrics such as precision, recall, and F1-score.

- Explore the use of transfer learning with pre-trained models.

- Deploy the model as a web application using Flask for real-time predictions.