

Early Detection of Pancreatic Tumor

¹A. Kumar,²Akash Kumar C, ³C Raj Kumar Reddy

⁴Pesala Madhusudhan Reddy, ⁵Vantari Uday Kiran

¹Assistant Professor, Department of Electronics and Communication Engineering

^{2,3,4,5}Students, Department of Electronics and Communication Engineering

Dhanalakshmi College of Engineering, Chennai, India

Abstract: Pancreatic cancer is one of the leading causes of cancer worldwide. However, pancreatic cancer can be cured if caught early. This article attempts to detect tumors on CT images. It uses image processing and CNN model architecture to detect tumors. After the image is created, the CNN Model Architecture is used to detect the tumor in the image. It has been determined that the training accuracy for cancer diagnosis is around 98.7%.

Keywords: Pancreatic Tumor, Normal, Pre-processing, Detect

I. INTRODUCTION

Since the implementation of Convolutional Neural Networks (CNNs) for visual data analysis, great success has been achieved in areas such as deep learning, computer vision, and image processing. It is possible to use deep learning for analysis of medical images, such as tumor detection for pancreatic cancer. Pancreatic cancer is one of the most common cancers and the 5-year survival rate is about 7%. The pancreas is a small organ located in the middle of the human body, and the complexity of research has increased. Also, remember that the best time for radical surgery is cancer. CT imaging is a medical surveillance technique that collects information about the location, size, and morphology of the tumor and is useful in the diagnosis and prognosis of pancreatic cancer compared to ultrasound imaging and magnetic resonance imaging (MRI). However, manual diagnosis requires physicians with a great deal of medical experience because the quality of CT images varies with different CT scanners or different operators, and it is difficult to distinguish features of different diseases. Therefore, there is a growing need for research to propose a deep learning-based algorithm to confirm pancreatic cancer.

1.1 Objective

The main purpose of this system is the automatic detection of pancreatic cancer using contrast-enhanced computed tomography (CT), which is widely used in the diagnosis and treatment of pancreatic cancer. Traditional manual methods only extract low-level features. But ordinary neural networks cannot use all the information about good content, which leads to bad results. In this paper, a new and efficient framework for the diagnosis of pancreatic disease has been established, aiming to use all the details of the various documents.

II EXISTING SYSTEM

Many methods have been proposed to avoid noise in natural image classification. Ren and others. A method is proposed to assign weights to the training model using additional hygiene methods. Their idea is to use small weights for noisy samples and increase weights to clean up training samples to improve gradient updating. There are some limitations to the development and use of popular classification systems in the medical literature.

Dgani et al. Model tag noise as part of a deep learning network to obtain ground truth for noise models to classify breast microcalcifications in multi-image mammograms.

III PROPOSED SYSTEM

In this project, try to detect the tumor with the CT scan image. These images are first used in image processing, then the CNN model architecture is used to classify tumors in the images. Data are from the Cancer Imaging Application Repository. CT scan images of this pancreatic tumor work for the body. Since these images are not in the same format, we convert them to jpg and only a small part is used because the images are very large.

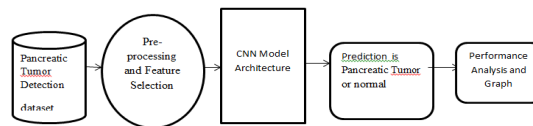
CNN model architecture is used for classification. Here, the system learns according to the classes defined in the image. Once the system learns to categorize by given attributes, it can classify test data into one of these categories.

We trained a CNN for patients to distinguish pancreatic cancer from healthy pancreatic cancer using contrast-enhanced CT images. CNN achieves very good accuracy and better understanding compared to radiologists' interpretation of independent test procedures and performance of test items from different patient types.

These results provide the first proof of concept that CNNs can capture CT abnormalities of cancer to aid and support radiologists in the diagnosis and diagnosis of tumors. eat pancreatic blood.

IV SYSTEM ARCHITECTURE

The description of the general features of the program is mixed with the meaning of the required products and the announcement of better orders. In architecture, many pages and their relationships are defined and created. Analyse and parse the main software in processing ideas and written procedures and analyse the relationship between modules.



V. SYSTEM REQUIREMENTS

Hardware Requirements

System : Pentium i3 Processor
Hard Disk : 500 GB.
Monitor : 15'' LED
Input Devices : Keyboard, Mouse
Ram : 2 GB

Software Requirements

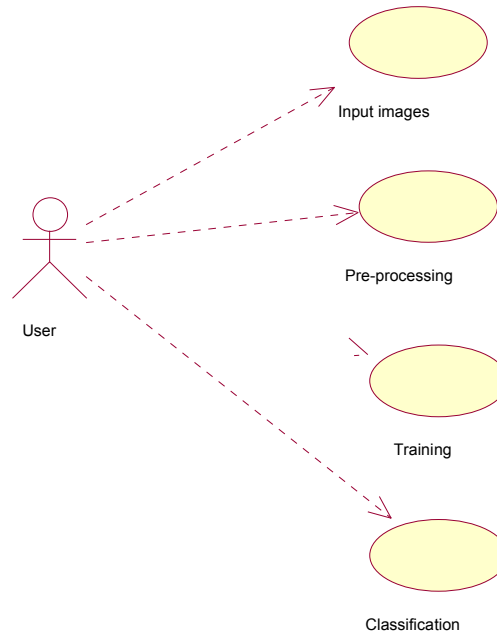
Operating system : Windows 10
Coding Language : Python

VI. METHODOLOGY

In this project, we use augmented reality to facilitate shopping. The techniques used here are site search and product discovery. In surface detection and object detection, tools use maps/objects and see the size of objects to help identify real-world objects.

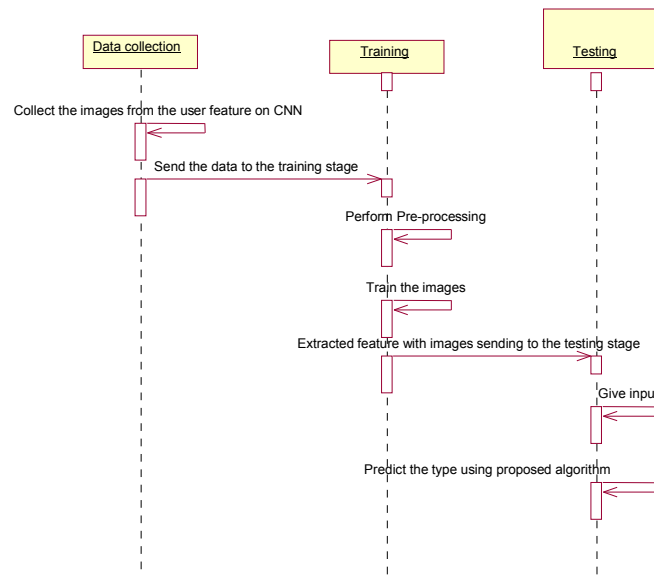
6.1 Use Case Diagram

A data diagram in Unified Modeling Language (UML) is a diagram that shows behavior defined and generated using data analysis. Its purpose is to provide an overview of the work provided by the system regarding the actors, their purpose (represented as use cases), and all the intervention effects of these uses. The main purpose of the graph using data is to show what work is done for which actor. The role of actors in the system can be explained.



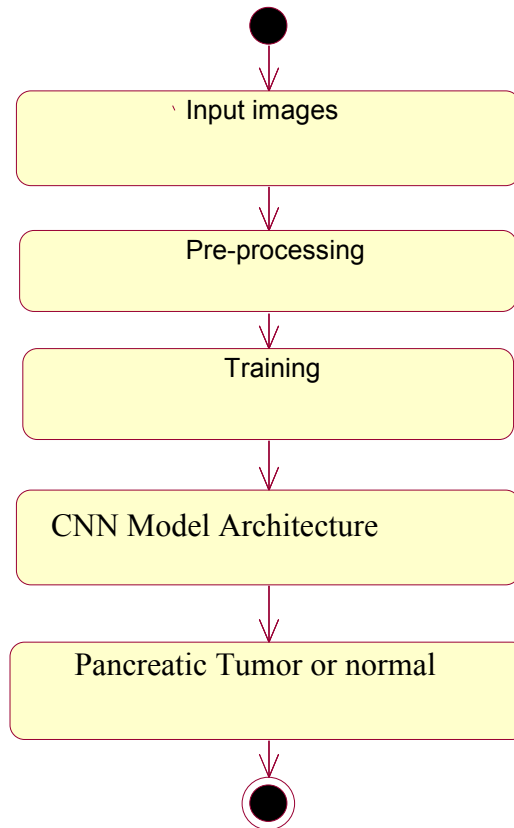
6.2 Sequence Diagram

Sequence diagrams in Unified Modelling Language (UML) are one of a kind, what is the order? It is the construction of the sentence system. Sequence diagrams are also sometimes called flow diagrams, sequence diagrams, and sequence diagrams.



6.3 Activity Diagram

A game map is a graphical representation of the operation of a set of operations and operations that supports selection, iteration, and concurrency. In Unified Modeling Language, game diagrams can be used to describe the work and periodic operations of components in a system. The diagram shows the overall control.



VII. MODULES

- Dataset
- Importing the necessary libraries
- Retrieving the images
- Splitting the dataset
- Building the model
- Apply the model and plot the graphs for accuracy and loss
- Accuracy on test set
- Saving the Trained Model

7.1 Module Description

Dataset:

In this model, we develop a method to obtain input data for training and testing purposes. We extracted information from pancreatic cancer diagnosis using image processing.

Link: <https://www.cancerimagingarchive.net/collections/>

The dataset consists of 1411 brain Tumor images

Importing the necessary libraries:

We will use the Python language for this. First, we will import necessary libraries such as keras to build key models, sklearn for classification of training and test data, PIL to convert images to numerical arrays, and other building libraries like pandas, NumPy, matplotlib and TensorFlow.

Retrieving the images:

We will keep the photos and their tags. Then set the image to (224,224) because each image needs to be the same size. Then convert the image to a NumPy array.

Splitting the dataset:

Divide the data into training and testing. 80% study material and 20% test material.

Copyright to IJARSCT

DOI: 10.48175/IJARSCT-10223

www.ijarsct.co.in



The purpose behind the first module of Convolutional Neural Networks is:

1. Understanding the convolution function
2. Understanding the joint function
3. Note the terms used in convolutional neural networks (fill, step, filter, etc.)
4. Create a convolutional neural net for multiple image classes

Building the model:

We will use the array from the keras library to create it. Then we will add layers to create a neural network. In the first 2 layers of Conv2D, we use 32 filters with kernel size (5.5). In layer MaxPool2D, we keep the pool size (2.2), which means it will choose the maximum value of each 2 x 2 area of the image. By doing this, the size of the image will be reduced by 2 times.

In the output method we keep the output value = 0.25, which means that 25% of the neurons are removed.

We apply these 3 layers again with some change in parameters. Then we apply flatten layer to convert 2-D data to 1-D vector. This layer is followed by dense layer, dropout layer and dense layer again. The last dense layer outputs 2 nodes as the brain tumour or not. This layer uses the SoftMax activation function which gives probability value and predicts which of the 2 options has the highest probability.

Apply the model and plot graphs for accuracy and loss:

We will collect the sample and use it through the fitting function. The batch size will be 2. Then we will explain the facts and loss of image. We achieved an average of 100% accuracy and 98.7% training accuracy.

Accuracy on test set:

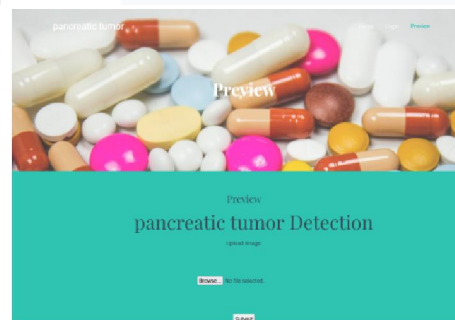
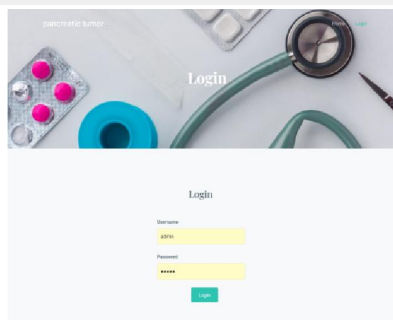
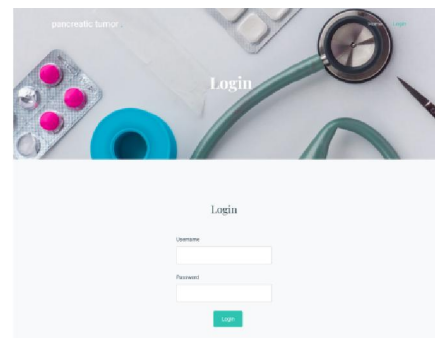
We achieved 100% accuracy in the testing process.

Saving the trained model:

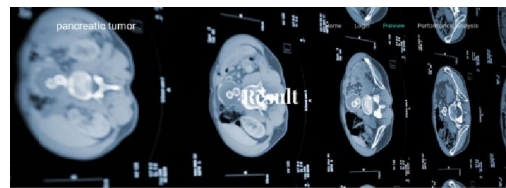
When you're confident enough to move your training and test models to a production-ready environment, the first step is to save them as .h5 or .pkl files using a library like pickle. Make sure you have marinade installed in your environment. Next, let's import the module and save the model as a .pkl file.

VI. RESULTS

Screenshots



Output:



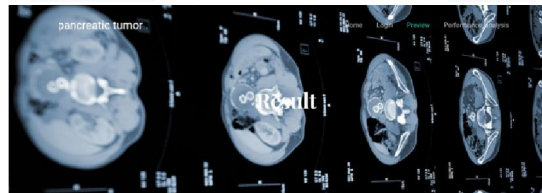
Pancreatic Tumor Detection

Here's the result of testing using image processing

Pancreatic Tumor Detection



Prediction: Normal



Pancreatic Tumor Detection

Here's the result of testing using image processing

Pancreatic Tumor Detection



Prediction: Pancreatic Tumor



PERFORMANCE ANALYSIS

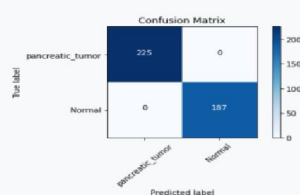
Accuracy: 1.000

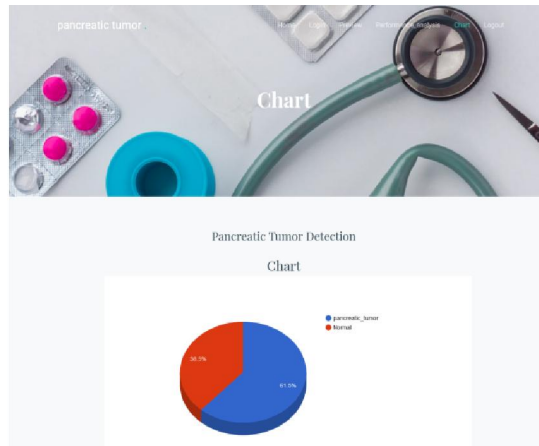
Precision: 1.000

Recall: 1.000

F-Measure: 1.000

Confusion Matrix





Jupyter screenshots:

```

In [1]: import tensorflow as tf
import os
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras import layers, models
import numpy as np
import random as rd
from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping
from tensorflow.keras import metrics

In [2]: train_data_dir = 'data/train'
val_data_dir = 'data/val'
img_height = 224
img_width = 224

In [3]: train_image_generator = ImageDataGenerator(rescale=1/255)
val_image_gen = ImageDataGenerator(rescale=1/255)
train_data_gen = train_image_generator.flow_from_directory(train_data_dir, target_size=(img_height, img_width),
class_mode='binary')
val_data_gen = val_image_gen.flow_from_directory(val_data_dir, target_size=(img_height, img_width),
class_mode='binary')
Found 800 images belonging to 2 classes.
Found 400 images belonging to 2 classes.

In [4]: model = Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(img_height, img_width, 3)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(128, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(256, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Flatten())
model.add(layers.Dense(1000, activation='relu'))
model.add(layers.Dense(1, activation='sigmoid'))

In [5]: model.compile(optimizer='adam',
loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),
metrics=['accuracy'])

```

```

In [6]: model.compile(optimizer='adam',
loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),
metrics=['accuracy'])

In [7]: img_height, img_width = 224, 224
img_data_gen = train_data_gen.flow_from_directory(train_data_dir, target_size=(img_height, img_width),
class_mode='binary')
val_data_gen = val_data_gen.flow_from_directory(val_data_dir, target_size=(img_height, img_width),
class_mode='binary')

In [8]: history = model.fit(img_data_gen, val_data_gen,
epochs=100,
validation_data=val_data_gen,
callbacks=[ModelCheckpoint('model.h5', save_best_only=True),
EarlyStopping(monitor='val_loss', patience=10)])

In [9]: model.evaluate(img_data_gen, val_data_gen)
100/100 [====] 100s 1000 steps - loss: 0.0000 - accuracy: 0.9999 - val_loss: 0.0000 - val_accuracy: 0.9999
100/100 [====] 100s 1000 steps - loss: 0.0000 - accuracy: 0.9999 - val_loss: 0.0000 - val_accuracy: 0.9999

```

Model Summary

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 224, 224, 32)	320
max_pooling2d (MaxPooling2D)	(None, 112, 112, 32)	0
conv2d_1 (Conv2D)	(None, 112, 112, 64)	5120
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 64)	0
conv2d_2 (Conv2D)	(None, 56, 56, 128)	16384
max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 128)	0
conv2d_3 (Conv2D)	(None, 28, 28, 256)	65536
max_pooling2d_3 (MaxPooling2D)	(None, 14, 14, 256)	0
flatten (Flatten)	(None, 4096)	0
dense (Dense)	(None, 1000)	4097000
dense_1 (Dense)	(None, 1)	1001
Total params: 4,417,120		
Trainable params: 4,417,120		
Non-trainable params: 0		

```

In [10]: model.compile(optimizer='adam',
loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),
metrics=['accuracy'])

In [11]: history = model.fit(img_data_gen, val_data_gen,
epochs=100,
validation_data=val_data_gen,
callbacks=[ModelCheckpoint('model.h5', save_best_only=True),
EarlyStopping(monitor='val_loss', patience=10)])

In [12]: model.evaluate(img_data_gen, val_data_gen)
100/100 [====] 100s 1000 steps - loss: 0.0000 - accuracy: 0.9999 - val_loss: 0.0000 - val_accuracy: 0.9999
100/100 [====] 100s 1000 steps - loss: 0.0000 - accuracy: 0.9999 - val_loss: 0.0000 - val_accuracy: 0.9999

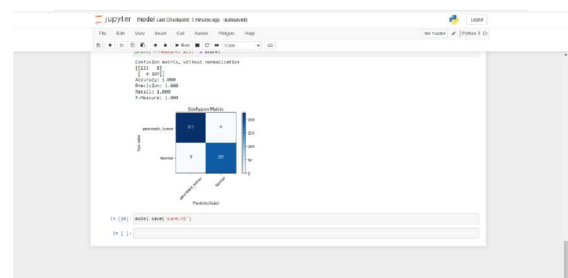
```

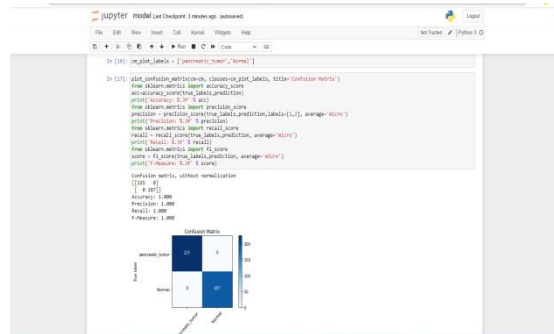
```

In [13]: def plot_confusion_matrix(cm, classes,
normalize=False,
title='Confusion matrix',
cmap=plt.cm.Blues):
"""This function prints and plots the confusion matrix.
Normalization can be applied by setting 'normalize' to True.
"""
plt.imshow(cm, interpolation='nearest', cmap=cmap)
plt.title(title)
plt.colorbar()
tick_marks = np.arange(len(classes))
plt.xticks(tick_marks, classes, rotation=45)
plt.yticks(tick_marks, classes, rotation=45)
if normalize:
cm = cm.astype('float') / cm.sum(axis=1)
print('Normalized confusion matrix')
else:
print('Confusion matrix, without normalization')
print(cm)
plt.tight_layout()
plt.show()

In [14]: from sklearn.metrics import confusion_matrix
import numpy as np
import matplotlib.pyplot as plt
cm = confusion_matrix(y_true=batch_labels, y_pred=model.predict(img_data_gen))
plot_confusion_matrix(cm, classes,
normalize=True,
title='Confusion matrix, with normalization')

```





VII. CONCLUSION

The article concludes that the CNN model architecture for cancer prediction is effective in our system. The system classified the images as normal images and pancreatic cancer images. The accuracy of the model is 100%. The system provides proof of concept that CNN can distinguish pancreatic cancer from CT images. The CNN model is expected to be a computer-aided diagnosis that will assist radiologists and doctors in diagnosing pancreatic cancer.

REFERENCES

- [1]. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014; 74: 2913–21.
- [2]. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019; 69: 7–34.
- [3]. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med* 2014; 371: 1039–49.
- [4]. Al-Hawary MM, Francis IR, Chari ST, et al. Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the Society of Abdominal Radiology and the American Pancreatic Association. *Radiology* 2014; 270: 248–60.
- [5]. Dewitt J, Devereaux BM, Lehman GA, Sherman S, Imperiale TF. Comparison of endoscopic ultrasound and computed tomography for the preoperative evaluation of pancreatic cancer: a systematic review. *Clin Gastroenterol Hepatol* 2006; 4: 717–25; quiz 664.
- [6]. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: current applications and future directions. *PLoS Med* 2018; 15: e1002707.
- [7]. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115–18.
- [8]. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; 316: 2402–10.
- [9]. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast enhanced CT: a preliminary study. *Radiology* 2018; 286: 887–96.
- [10]. Frampas E, David A, Regenet N, Touchefeu Y, Meyer J, Morla O. Pancreatic carcinoma: key-points from diagnosis to treatment. *Diagn Interv Imaging* 2016; 97: 1207–23.