

A Deep Learning Approach for an Early Diagnosis of Pancreatic Cancer

Saraswathi H S

Dept of CSE
Jain Institute of Technology
Davangere

Kotramma T S

Dept of CSE
Jain Institute of Technology
Davangere

Uma K

Dept of CSE
Jain Institute of Technology
Davangere

Thejaswini K R

Dept of CSE
Jain Institute of Technology Davangere

Tejashree K

Dept of CSE
Jain Institute of Technology Davangere

Abstract:- Pancreatic cancer has a high incidence due to a lack of early detection procedures for this particularly aggressive type of cancer. The goal of this study was to find biomarkers in the serum proteome for early identification of resectable pancreatic cancer using high-throughput protein profiling technologies[10]. After fractionation on an anion exchange resin, protein profiles were obtained from the serum of 49 pancreatic cancer patients and 54 healthy individuals employing surface-enhanced laser desorption/ionization mass spectrometry. Using two multivariate analysis methodologies, classification and regression tree and logistic regression, these spectral data were used to create classification models that could identify pancreatic cancer serum samples from control serum samples. During the examination.

Keywords ; ANN, CNN, Corelation Matrix, Accuracy.

I. INTRODUCTION

The pancreas is a digestive organ located behind the stomach. It features a huge head, a tapering body, and a slender, pointed tail that resembles a fish. In adults, it's around 6 inches long and less than 2 inches wide. The head of the pancreas is located on the right side of the abdomen, behind where the stomach meets the duodenum (belly)[2][3]. The pancreas body is located behind the stomach. The pancreatic tail is located close to the spleen on the left side of the belly. With accurate views of endocrine cells secreting hormones into blood vessels and exocrine cells secreting hormones into the digestive tract, the pancreas and pancreatic duct are depicted in relation to the liver, gallbladder, common bile duct, duodenum, and ampulla of Vater.

In the United States, pancreatic cancer is the fourth highest cause of death with pancreatic ductal adenocarcinoma (PDAC)[4] accounting for over 90% of all pancreatic malignancies and neuroendocrine tumours accounting for the remaining 10%. PDAC[4] has a five-year survival rate of less than ten percent.

Symptoms

- Your skin is going yellow, as are the whites of your eyes (jaundice).
- Loss of appetite or rapid weight loss.
- Diabetic patients who have recently been diagnosed or whose diabetes is becoming more difficult to manage.
- Clots in the blood.

II. OBJECTIVE OF THE STUDY

- To compile a database of patients with pancreatic cancer who had serum biomarkers.
- To prepare the data for analysis.
- To create the classifier that will be used to train the model.
- To assess the model's correctness and analyse its performance for various classifiers based on the parameters.

A. ANN

The approached ANN-based Image classifier for early detection of pancreatic tumour Abdominal ultrasonography has a significant disadvantage in that information about staging is not apparent, and it fails to detect tiny tumours that are missed by MRI or CT scan[1]. They achieved a 98 percent accuracy rate in adopting an ANN-based image classifier to increase patient survival rates and detect pancreatic tumours at an early stage using PET scan pictures.

Using an Artificial Neural Network Model, researchers were able to diagnose pancreatic cancer based on symptoms. To train the network, they adopted the approach of the Levenberg-Marquardt back propagation algorithm. They used this technique to detect different stages of cancer in the pancreas. The accuracy of the results was 87 percent, indicating the benefits of utilising an ANN model.

The brain tumour detection that is adaptable Image processing is employed in medical instruments to detect tumours; however, MRI pictures are unable to identify the tumorous region and must be preprocessed with K-Means segmentation. They employed SVM (Support Vector Machine) to make this system adaptive. SVM is used in an unsupervised manner to develop and maintain patterns for future use. They must also discover the feature to train SVM for patterns. They discovered the texture and colour properties for this purpose. It is believed that the suggested system's experimental outcomes will be superior to those of other existing systems.

The modified Self-Organizing Map (SOM) network is a neural network-based technique for medical image analysis that is compared to the standard SOM network. 30 Breast Ultrasound Images and 10 Magnetic Resonance Images are included in this dataset (MRI)[9]. For the study, one

Computerized Tomography (CT)[1] head image and a set of brain images were employed. evaluation. To create the model, a two-dimensional Discrete Wavelet Transform (DWT) is used. The network's input feature space. The network's filtering property minimises the image size. To a large extent, noise. In compared to SOM-based methods, the proposed method is more accurate. The Jaccard, Rogers, and Tanimoto indexes all support the network technique.

B. CNN

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning system that can take an input image and prioritize (learnable weights and biases) various aspects/objects in the image while also recognizing them. Other classification techniques necessitate far more pre-processing than ConvNet. With adequate training, ConvNets can learn these filters/characteristics, but core techniques involve filter engineering by hand.

The convolutional neural network resembles the architecture of the Visual Cortex and is analogous to connecting a network of neurons in the human brain. Every neuron responds to stimuli of the receptive field. Receptive is a small part of the Visual field. The group of series is connected to the field and merged with the visual fields.

Correlation matrix

The Correlation matrix is in the form of tables that presents the coefficients of correlation between variables. Each table presents the relationship between two variables A correlation matrix can be used to summarise data, as an input to a bigger study, or as a diagnostic tool for future research. In the field of statistical classification, a confusion matrix, also known as an error matrix, is a table structure that allows visualization of the performance of a machine learning algorithm, most commonly a supervised learning algorithm (unsupervised learning is usually referred to as a matching matrix).

The rows of the matrix indicate the examples in a predicted class, whereas the columns reflect the occurrences in an actual class (or vice versa).

If the system is integrating two types of data, the name is self-evident (i.e. frequently mistaking one for another). It's a unique contingency table with two dimensions ("actual" and "predicted"), both with the same set of "classes". In probability theory and statistics, a covariance matrix (also known as an auto-covariance matrix, dispersion matrix, variance matrix, or variance-covariance matrix) is a square matrix that shows the covariance between each pair of components of a random vector. The variances or covariances of each component with itself are represented by the diagonal of the matrix.

III. METHODOLOGY

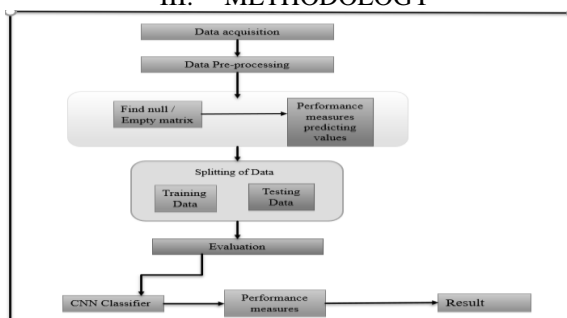


Fig 1: Methodology of CNN

A. Dataset Acquisition

The Dataset Acquisition refers to collecting raw data from the real world. benign sample diagnosis, plasma CA19 9, CEA, sample id, patient cohort, age, sex, diagnosis, stage, benign sample diagnosis.

B. Data preprocessing

Preprocessing data is a frequent preparing raw input in a way that the network can accept is the first stage in the deep learning workflow. You can, for example, change the image input layer size to match the image input layer size You can also preprocess data to improve desired features or remove artifactsthat may bias the network.

Splitting datasets into training and testing datasets

Divide the modelling dataset into two parts: training and testing. Assign two-thirds of the data points to training and one-third to testing. As a result, before applying the model to the test set, we train it on the training set. This is how we may judge the performance of our model.

C. Evaluation

Accuracy, The Confusion Matrix, Precision, and Recall are all used to evaluate deep learning models.

- Binary Classification Confusion Matrix
- Multi-Class Classification Confusion Matrix
- Scikit-learn is used to calculate the Confusion Matrix.
- Accuracy, Precision, and Recall are three words that come to mind while
- thinking about accuracy, precision, and recall.
- Which is more important: precision or recall?
- Conclusion.

D. Performance measure

Every machine learning pipeline has performance measurements. They tell you if you're progressing and give you a number. A metric is required for all machine learning models, whether linear regression or a SOTA approach like BERT.

F.Result Analysis

```

Epoch 145/150
63/63 [=====] - 0s 2ms/step - loss: 0.0261 - accuracy: 0.9915
Epoch 146/150
63/63 [=====] - 0s 2ms/step - loss: 0.0393 - accuracy: 0.9865
Epoch 147/150
63/63 [=====] - 0s 2ms/step - loss: 0.0265 - accuracy: 0.9910
Epoch 148/150
63/63 [=====] - 0s 2ms/step - loss: 0.0438 - accuracy: 0.9860
Epoch 149/150
63/63 [=====] - 0s 2ms/step - loss: 0.0447 - accuracy: 0.9837
Epoch 150/150
63/63 [=====] - 0s 2ms/step - loss: 0.0278 - accuracy: 0.9910
32/32 [=====] - 0s 2ms/step - loss: 0.3291 - accuracy: 0.9250
test_acc: 0.925000011920929
32/32 [=====] - 0s 1ms/step
1/1 [=====] - 0s 33ms/step
[[1. 0.]]
0
The Patient is not having Cancer
  
```

Fig 2: Accuracy and Result

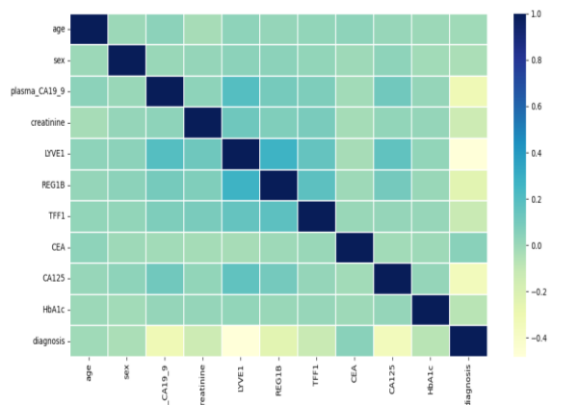


Fig 3: Correlation Matrix

CONCLUSION

Finally, an AI tool may integrate risk factor information into the context of a patient's disease history. The researchers believe that using data other than disease codes, such as prescriptions, laboratory values, and data, will improve AI prediction accuracy.

REFERENCES

- [1] Deep Supervision for Pancreatic Cyst Segmentation in Abdominal CT Scans Zhou, Yuyin | Xie, Lingxi | Fishman, Elliot K. | Yuille, Alan
- [2] The Use of Biomarkers in Early Diagnostics of Pancreatic Cancer Kroupa, Radek | Kala, Zdenek | Tesarikova, Pavla | Kysela
- [3] Jagannathan, P., Rajkumar, S., Frnda, J., Divakarachari, P.B. and Subramani, P., 2021. Moving vehicle detection and classification using gaussian mixture model and ensemble deep learning technique. *Wireless Communications and Mobile Computing*, 2021.
- [4] Deep Radiomics Analytics pipeline for prognosis of pancreatic Ductal Adenocarcinoma Zhang, Yucheng
- [5] Xue, C J Scarlett, L Chung, G Butturini, A Scarpa, R Gandy, S R Wilson, R C Baxtyer & R C Simth British Journal of Cancer volume 103, pages391–400 (2010) Cite this article 1168 Accesses 41 Citations 9 Altmetric Metrics
- [6] Randall E. Brand, Brian M. Nolen, Herbert J. Zeh, Peter J. Allen, Mohamad A. Eloubeidi, Michael Goldberg, Eric Elton, Juan P. Arnoletti, John D.
- [7] Christein, Selwyn M. Vickers, Christopher J. Langmead, Douglas P. Landsittel, David C. Whitcomb, William E. Grizzle and Anna E. Lokshin DOI: 10.1158/1078-0432.CCR-10-0248 Published February 2011
- [8] Rui Liu, Xi Chen, Yiqi Du, Weiyan Yao, Lin Shen, Cheng Wang, Zhibin Hu, Rui Zhuang, Guang Ning, Chunni Zhang, Yaozong Yuan, Zhaoshen Li, Ke Zen, Yi Ba, Chen-Yu Zhang *linical Chemistry*, Volume 58, Issue 3, 1 March.
- [9] Use of Radiographic Criteria Predict the Outcomes of following Surgrey for pancreatic cancer Reynolds, Rae Brana | Justin, Folloder
- [10] Pancreatic cancer: Classifying pancreatic cancer using gene expression profiling. Ayars, Michael | Goggins, Michae