Pancreatic Cancer Detection Using Machine Learning

Karan S
Dept. Of Master of Computer Application
AMC Engineering College
Bengaluru, India
karans2742@@gmail.com

***Abstract*—** This paper delivers into the significant research being conducted on medical health systems, which is driving the development of computing systems with the latest innovations. These advancements are facilitating more efficient implementations of medical systems, including the automatic identification of health-related disorders. Among the most crucial health research areas is the prediction of cancer, which can manifest in various forms and affect different parts of the body. Pancreatic cancer, one of the most common and currently considered incurable cancers, is a primary focus. Previous studies have identified a panel of three protein biomarkers (LYVE1, REG1A, and TFF1) in urine that can help detect resectable PDAC (pancreatic ductal adenocarcinoma). This study aims to improve this panel by replacing REG1A with REG1B, using data extracted into CSV format. Creatinine is a protein commonly used as an indicator of kidney function, LYVE1 may assist in tumor spread, REG1B is associated with pancreatic regeneration, and TFF1 is linked to urinary tract regeneration and repair. Effective treatment of pancreatic cancer is challenging once it is diagnosed. However, machine learning and neural networks are showing promise for accurate real-time segmentation of pancreatic images for early diagnosis. This research explores how to analyze pancreatic tumors using ensemble approaches in machine learning. Preliminary data suggest that the proposed technique enhances classifier performance for the early diagnosis of pancreatic cancer.

Keywords—Deep Learning, Convolutional Neural Network, Machine Learning, Decision Tree.

I. INTRODUCTION

According to medical health news analysis, cancer is one of the most challenging diseases, often seeming invincible. It may be hereditary, caused by genetic abnormalities that control cellular functions in the human body. These genetic changes can be inherited or result from lifestyle factors. The pancreas, a vital organ with internal and external secretory functions, is susceptible to various diseases. Surgical access is often difficult, and pre-biopsy is frequently impossible. Pancreatic cancer is the fourth most common cause of cancer death and the second leading cause of death from digestive system neoplasms. However, regular segmentation of the pancreas remains challenging due to low soft tissue contrast on CT images and significant anatomical variations. The pancreas varies greatly in size and location within the abdominal cavity of patients, making it a deformable, yielding tissue. Consequently, its outline and appearance differ greatly among individuals. PDAC, a particularly lethal form of pancreatic cancer, poses significant diagnostic and treatment challenges.

II. LITERATURE SURVEY

1. Pancreatic Cancer Prediction through Convolutional Neural Networks

Due to the late onset of cancer-specific symptoms, diagnosing pancreatic cancer early is challenging, and there is no reliable screening method to identify high-risk patients. CT, MRI, and EUS are the most commonly used techniques for detecting abnormalities in the pancreas. Therefore, precise and automated classification methods are necessary to reduce the human death rate. To minimize the time radiologists spend on cancer identification while maintaining accuracy, various approaches are being explored. However, detecting pancreatic cancer with MRI scans and EUS is complex and time-consuming, and these techniques are not widely adopted. To address these challenges and develop a more accurate, affordable, and practical method for identifying pancreatic cancer, this study employs a range of deep learning techniques using convolutional neural networks alongside other machine learning algorithms..

2. Pancreatic Cancer Detection using Machine and Deep Learning Techniques

Despite extensive research, pancreatic cancer remains associated with a grim prognosis, boasting a five-year survival rate of only a few individuals. The rationale behind early detection and improved survival rates lies in the potential for more widespread access to effective treatments. Across the healthcare landscape, machine learning and deep learning algorithms have emerged as promising tools for classifying and identifying pancreatic cancer risk. Consequently, our study delved into multiple methodologies proposed by researchers for diagnosing pancreatic cancer using these advanced models. Additionally, the report underscored their accomplishments and the persistent challenges within this domain. By evaluating numerous strategies, we aimed to draw meaningful conclusions from the available data.

3. Review on computer aided diagnosis of pancreatic cancer using Artificial Intelligence System

Malignant growth refers to abnormal cell tissue growth. Pancreatic cancer ranks among the leading causes of death worldwide. It originates in the pancreas tissue, which secretes digestive enzymes and hormones regulating sugar metabolism. Typically diagnosed at advanced stages, pancreatic cancer spreads quickly and carries a grim prognosis. This paper explores several artificial intelligence approaches for detecting pancreatic cancer and proposes a novel AI method to identify subtle patterns, offering precise information to pathologists

.4. Earlier Detection of Pancreatic Cancer Using Neural Network Based Optimization Technique

Medical healthcare systems are under extensive study, providing ample room for innovation in computer technology. A critical focus of medical research involves predicting cancer, which manifests in various forms and affects multiple body organs. Pancreatic cancer stands out as one of the most common and deadly diseases, challenging to cure once diagnosed, often located unexpectedly in the abdomen near the stomach. Techniques like CT and MRI facilitate computer-aided diagnosis (CAD), offering quantitative assessments and automated segmentation of pancreatic cancer. These methods in cancer classification can aid in identification, prediction, and personalized treatment strategies, aiming to mitigate malignant progression. Optimization techniques such as Flying Squirrel Optimization for segmentation, feature extraction, and classification, alongside CNN integrated with Frog Leap Optimization, play crucial roles. The proposed approach employs Frog Leap Optimization to discern normal and abnormal images, effectively minimizing classification errors. Adaptive algorithms like Flying Squirrel Optimization further refine cancer size and severity assessment through image segmentation. Notably, the CNN-FLFS method demonstrates high predictive accuracy of 99% in diagnosing pancreatic cancer.

5. Computational learning of small RNA regulation in pancreatic cancer progression

Based on rapidly accumulating genomics data, modeling gene regulation networks has become the primary computational approach to deduce interactions among genes and their regulators, such as transcription factors (TFs) and microRNAs, thereby studying (post-)transcriptional regulation in biological systems. However, extracting dynamic behaviors from data in complex diseases poses a significant challenge due to the intricate interplay of regulatory mechanisms, which vary over time. Therefore, developing scalable learning models capable of encompassing diverse regulatory types by leveraging heterogeneous data is crucial, ensuring each biological species is treated uniquely and meaningfully. This study investigates integrated methods for learning gene networks in human cancer, with a specific emphasis on microRNA-mediated regulation. We propose a framework that combines RNA expression and interactome data, integrating distinct graphical models to shift from predicting static interactions to identifying semi-conditional relationships. Analyzing data from human pancreatic cancer reveals unique gene regulatory networks across four progressive stages, highlighting 15 microRNA-gene interactions conditionally associated with these stages. Functional analysis uncovers significant dysregulation by microRNAs in key cancer hallmarks, particularly affecting adaptive immune response and lymphocyte proliferation, illuminating their regulatory roles in pancreatic cancer progression. We propose that this integrated model serves as a robust tool for discovering critical regulatory features in other complex biological systems.

III. EXISTING SYSTEM

The five-year survival rate drops to less than 10% once diagnosed. However, if the cancer is detected early, when tumors are still small and manageable, the five-year survival rate can climb to 70%. Unfortunately, many cases of pancreatic cancer remain undetected until the disease has advanced significantly. Therefore, a diagnostic test to identify pancreatic cancer patients could be very advantageous. Traditionally, blood has been the primary source of biomarkers, but urine presents a viable alternative. It allows for non-invasive sampling, high-volume collection, and easy repeated measurements. Currently, there are no reliable biomarkers for early detection of PDAC. The only biomarker used in clinical practice, serum CA19-9, lacks the specificity and sensitivity needed for screening and is mainly used as a prognostic marker and for monitoring treatment response.

IV. PROPOSED SYSTEM

Despite requiring invasive samples, he improves cancer diagnosis when combined with other urine indicators in a study. Previous research identified a panel of three protein biomarkers (LYVE1, REG1A, and TFF1) in urine that can help detect significant PDAC. In this study, we enhanced this panel by substituting REG1A with REG1B. Ultimately, we will analyze four significant biomarkers found in urine: creatinine, LYVE1, REG1B, and TFF1. Creatinine is a protein commonly used as a kidney function indicator. Lymphatic vessel endothelial hyaluronan receptor 1 (LYVE1) is a protein that may facilitate tumor spread.

V. CONCLUSION

This study examines the use of ensemble methods in machine learning for analyzing pancreatic tumors. Researchers are incorporating features such as active attention and in-line memory, enabling folding neural networks to evaluate novel elements significantly different from the training data. Preliminary results indicate that the proposed approach enhances the classifier's performance in the early detection of pancreatic cancer. This method closely mimics the mammalian visual system, suggesting a more intelligent artificial image recognition classification. Despite collecting invasive samples, cancer diagnosis improves when combined with other urine indicators. In this study, we improved the panel by replacing REG1A with REG1B. Consequently, we will analyze four significant biomarkers found in urine: creatinine, LYVE1, REG1B, and TFF1. Creatinine is commonly used as an indicator of kidney function. LYVE1 (Lymphatic Vessel Endothelial Hyaluronan Receptor 1) may assist in tumor spread. REG1B is associated with pancreatic regeneration, and TFF1 (Trefoil Factor 1) is linked to urinary tract regeneration and repair. This regularization of the form's continuity allows for smooth pancreatic segmentation. The preliminary results reflect the state of the art in pancreatic cancer prediction and achieve a high level of precision. However, further research is required for early pancreatic cancer detection, especially since pancreatic damage induced by COVID-19 infection has received minimal attention.

VI. REFERENCES

1. H. Matsubayashi, H. Ishiwatari, K. Sasaki, K. Uesaka, and H. Ono,``Detecting early pancreatic cancer: Current problems and future prospects," Gut Liver, vol. 14, no. 1, pp. 3036, Jan. 2020.
2. Z.-Y. Wang, X.-Q. Ding, H. Zhu, R.-X. Wang, X.-R. Pan, and J.-H. Tong,``KRAS mutant allele fraction in circulating cell-free DNA correlates with clinical stage in pancreatic cancer patients," Frontiers Oncol., vol. 9,p. 1295, Nov. 2019.
3. Shamsuddin, A. S., Rashid, T. A., Al-Rashid Agha, R. AAl-Salihi, N. K., &Mohammadi, M. (2019). Donkey and smuggler optimization algorithm: A collaborative working approach to path finding. Journal of Computational Design and Engineering, 6(4), 562-583.