Advances and Challenges in Artificial Intelligence and Statistical Methods for the Early Diagnosis of Pancreatic Cancer

Yaxuan Yu[1,*]

[1]School of Suzhou Industrial Park Overseas Chinese Academy of Chiway Suzhou, Suzhou, Jiangsu, 215000, China
*Corresponding author: 26yuyaxuan@ocac-suzhou.net

Abstract:
This study reviews the research progress, current applications, and future challenges of artificial intelligence and statistical methods in the early diagnosis of pancreatic cancer. Dubbed the ‘king of cancers’ due to its insidious early symptoms and rapid metastasis, pancreatic cancer exhibits low early detection rates, with traditional CT, MRI, and biomarker diagnostics demonstrating limited sensitivity. Artificial intelligence demonstrates significant potential in deep learning, machine learning, multimodal data analysis, medical image recognition, and clinical decision support systems, substantially enhancing the detection rate of small lesions and diagnostic accuracy. By integrating multi-source medical imaging, genomic, and clinical data, AI not only improves screening efficiency but also assists clinicians in formulating personalised treatment strategies. However, current limitations include inadequate data standardisation, insufficient explainability, and insufficient external validation. Future efforts should rely on large-scale, multicentre clinical studies to advance the clinical translation and practical application of AI models in pancreatic cancer early screening, thereby providing new avenues for reducing mortality rates.

Keywords: Epidemiology; Multi-modal data analysis; Medical image analysis; CDSS.

1. Introduction

Nowadays, the artificial intelligence (AI) appeared in almost every aspect of humanity’s life including the transportation, agriculture, entertainment and in health care, which can provide immense convenience. And the most essential and fundamental trend is the research and application of AI in medical field, especially, being applied to the diagnosis of various early cancers that cannot be precisely detected through existing evaluation and judgements. AI can not only compensate the limitations of conventional methods of prognosis, but also significantly improve the early detection rate remarkably. Take the lung cancer as an example, gold nanocube superlattices (GNSs) can be used to detect it by utilising surface-enhanced Ramen

ISSN 2959-409X

scattering (SERS) signals from both internal and surface biomarkers. It can classify the normal cells and the damaged cells with a quite high accuracy about 98.95% [1]. Similarly, but much more difficult, is the diagnosis of pancreatic cancers which is often referred to as the 'king of cancers' due to its high lethality. Since the patients of it can be symptomless at the early stage which indicates that they cannot realize the existence of it as early as possible. And especially, during this period, it is easily metastatic. Hence, this will accelerate the deterioration. Unfortunately, the disease has caused a surprising death rate in America, which is 10.9 per 100000 [2]. Apart from that, according to the latest SEER 17 data registry, the overall survival rate of 5 years is only 12% and there is no improvement in the last few decades [3]. As for the existing method such as computed tomography (CT) and biomarkers, they have low sensitivity and specificity. At the same time, it also requires an abundant of experience of the operators and readers. About 40% patients with cancer may be ignored [4]. However, if the tumor can be diagnosed and treated immediately when it is less than 2 centimeters, the survival rate is expected to increase significantly to 80% [4]. So nowadays, increasing research efforts have been devoted to the rapid advancement of AI, since it can help to analyze and evaluate in a deeper way to give a more accurate result. The motivation for this review is to popularize related professional knowledge to the individuals and keep doctors fully abreast of cutting-edge advances. The main objective of this review is to summarize the application status and effect of AI in the diagnosis of the early stage of pancreatic cancers and the merits and demerits respectively by concluding more than 100 studies (2000-2025 years) from PubMed, Nature, Web of science and other professional sources. To be honest, there is still a lot of room and potential for improvement in this area. As a result, it is worthy to understand thoroughly about the up-to-date technology and allow the scientists to modify current technology.

2. Epidemiology and Clinical Characteristics of Pancreatic Cancer

First, International Classification of Disease, 10th Revision (ICD-10) is introduced. It is managed by the World Health Organization (WHO), and it is a standard of classification and a code for disease [5]. It can record all kinds of health conditions. In the early 1990s, ICD-10 has already replaced the ICD-9, as it is much more specified and detailed which has over 54000 more codes compared to ICD-9 [5]. In ICD-10, C25 is the code for malignant neoplasm of pancreas [6]. And according to the different locations of the pancreas, the code varied as table 1 below shown. This coding system not only facilitates standardized documentation and international communication in clinical practice, but also enables the integration of structured medical data into AI algorithms for disease detection and prediction. By clearly defining site-specific codes for pancreatic cancer, researchers and clinicians can ensure that AI models are trained on accurate, location-specific data, thereby improving diagnostic precision (Table 1).

Table 1. Three Scheme comparing

ICD-10 Code	Location	ICD-10 Code	Location
C25.0	Head of Pancreas	C25.0	Head of Pancreas
C25.1	Body of Pancreas	C25.1	Body of Pancreas
C25.2	Tail of Pancreas	C25.2	Tail of Pancreas

At the same time, if specific metastasis and symptoms occur, secondary ICD-10 like the series of C78 is needed to be used together with C25. It can describe accurately about the cancer's complexity [7]. In addition, it is quite vital for the treatment and statistics required. Above all, the related codes are the most fundamental information needed to know about the application of AI. Gaining a deeper understanding of ICD-10 codes related to pancreatic cancer not only ensures standardization in case records and statistical analysis, but also lays a unified data foundation for the application of artificial intelligence in disease identification, disease progression tracking, and multi-center data integration. Clarifying this basic information is a prerequisite for promoting the precision of AI diagnostic technology and its clinical application.

3. Key Technologies for Artificial Intelligence in Cancer Diagnosis

In recent years, artificial intelligence technology has significantly transformed the paradigm of cancer diagnosis, with its core advancements primarily manifested in the following areas.

3.1 Machine Learning (ML) and Deep Learning (DL)

ML, a subfield of artificial intelligence, empowers computer systems to learn autonomously from training data and experiential input, thereby continuously self-optimizing and improving predictive accuracy. It makes autonomous decisions using mathematical and statistical models without direct human instructions [8]. High-quality data is essential for effectively training ML models [8]. Meanwhile, DL is a branch of ML which is a model combing multi-layer neural networks [9]. Algorithms such as Naive–Bayes, support vector machine, linear regression analysis, ensemble methods, decision tree, K-mode, hidden Markov model, hierarchical, Gaussian mixture, and neural networks have all been explored with different imaging data sets for distinguishing cancerous tissue from non-cancerous tissues [10]. Compared to machine learning (ML) models, deep learning networks demonstrate superior diagnostic capabilities by extracting all features from medical images, unlike ML which selectively extracts specific features [11]. Consequently, DL models are better suited for detecting digestive cancers and image segmentation. For instance, convolutional neural networks (CNNs), one of the most widely used supervised DL techniques, consist of an input layer containing distinct node clusters with specific features. These nodes interact with hidden layers sharing identical weights and biases through convolution operations on the inputs. The processed data is then aggregated and transformed to produce final outputs [12]. A typical CNN architecture includes input, convolutional, activation, pooling, fully connected, and output layers [13]. While CNNs offer computational efficiency, they require substantial processing power and can be slow. They provide probabilistic descriptions of complete images, making them more suitable for classification than segmentation tasks [14]. Among various CNNs, the U-Net algorithm with fewer convolutional layers is commonly employed in diagnosing digestive tract cancers like pancreatic cancer, achieving both feature classification and segmentation through specialized image analysis [15]. Therefore, CNN is specialized in disposing images which is the core technology to realize image recognition and pathological diagnosis automation. Meanwhile, artificial neural networks (ANN) is also very essential in the diagnosis of pancreatic cancer. the process of how it functions is shown below in figure 1. The advantage of it is that it can process large amounts of data and predict various interactions and relationships between the dependent and independent variables.

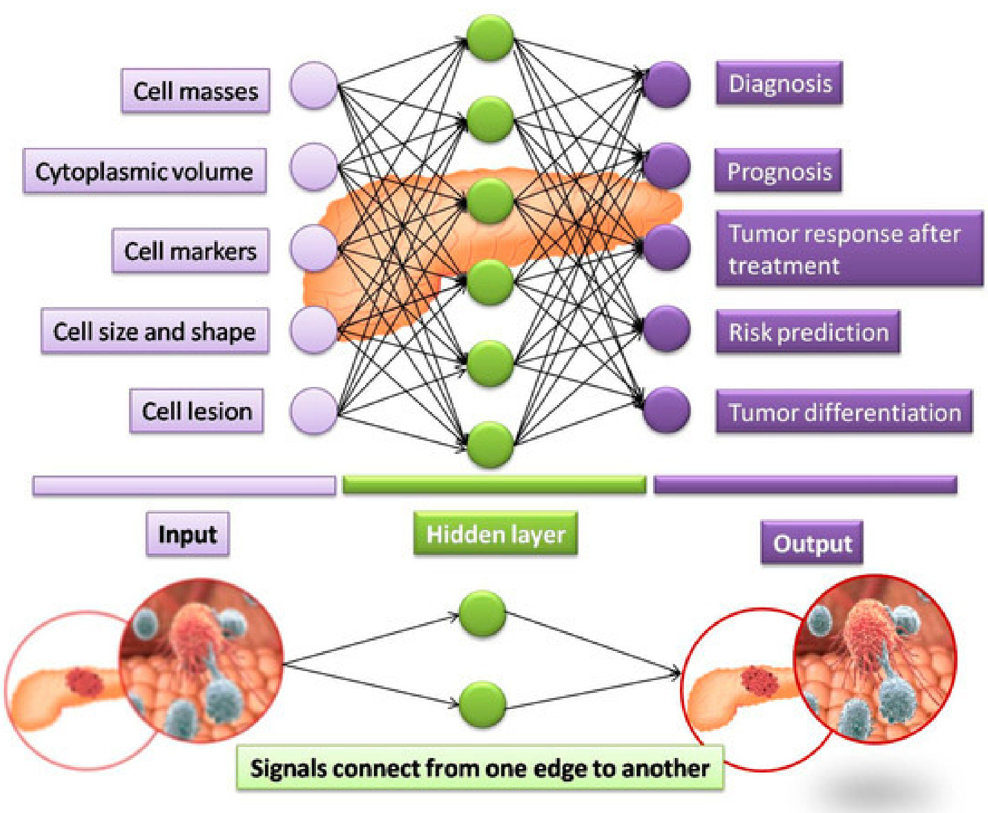

Fig. 1 Schematic diagram of the process flow in the sample ANN model used for diagnosing pancreatic cancer [16]

ISSN 2959-409X

3.2 Multi-Modal Data Analysis

Traditional single-modal data approaches often struggle to capture critical information in multimodal relationships. Multimodal analysis integrates diverse data sources like text, images, and audio to achieve comprehensive subject understanding [17]. This methodology not only reveals intricate connections between different modalities but also provides a holistic perspective for current challenges, thereby significantly enhancing prediction accuracy. According to an experiment, multi-modal data analysis can predict not only the long-term but also the short-term survival of patients with pancreatic cancers, showing potential as a clinical decision-making tool [18]. Since pancreatic cancer is very complex, by combining imaging (to look at shape), pathology (to look at cells), genetic testing (to look at genetic information), and clinical data (to look at overall health), AI models can make more comprehensive and accurate disease diagnoses, predict disease progression, and assess treatment outcomes. This is the key to achieve precision medicine.

3.3 Explainable AI and Assisted Decision Making

Explainable Artificial Intelligence (XAI) is a set of methods and techniques. Its goal is to make sure that machine learning algorithms produce outputs that people can understand and trust. XAI is a key part of the Fairness, Accountability, and Transparency (FAT) framework in machine learning. Because of this, it is often talked about when people discuss deep learning. For organizations seeking to build trust during AI deployment, XAI holds significant value. It enables users to gain deeper insights into model behavior while promptly identifying potential flaws or biases within the system. Take Grad-CAM as an example, it can provide a visualization of the Ramen features and by considering the reported biomarkers, it can further analyze the extracted features [19]. As the outcomes of an experiment demonstrated that the classification accuracy of TDN network can reach 98.344% for pancreatic part, and 99.471% for specific pancreatic part regardless of cancer occurrence [20]. Meanwhile, aid decision making must be carried out simultaneously as AI is solely in an auxiliary function. This is because misdiagnosis is extremely easy and the cost of it is a patient's life. Therefore, diagnostic decisions must be extremely careful and reliable. With the clues provided by XAI, surgeons can use their clinical knowledge and experience to verify it.

4. Typical Application Scenarios of AI in Early Screening and Diagnosis of Pancreatic Cancer

4.1 Medical Image Analysis

Currently, there is no standard procedure of screening images. But as for populations with inherited genetic syndromes or family history of pancreatic cancer, computed tomography (CT), magnetic resonance imaging (MRI) as well as endoscopic ultrasound (EUS) are used [20].

CT, especially spiral CT with a multidetector-row computed tomography (MDCT), remains the first-line imaging technology for initial detection and staging of pancreatic cancer. CT is excellent for visualizing the relationship between tumors and major pancreatic perivascular vessels (e.g., celiac trunk, superior mesenteric artery and vein), which is critical for determining surgical resectability. It can also be used to evaluate distant metastases. Overall sensitivity of MDCT for detecting tumors was reported to be between 76 and 92 per cent but decreased to between 63 and 77 per cent when the analysis included small tumors measuring less than 2cm in size [21]. Therefore, it has limited sensitivity for detecting early metastatic deposits and small tumors.

MRI provides excellent soft tissue contrast and functional imaging capabilities, making it particularly valuable in solving difficult problems and lesion characterization. Diffusion-weighted imaging (DWI) and contrast agents are employed to enhance MRI scanning details. By analyzing water molecules, DWI, MRI reveals tissue microstructures, enabling differentiation between healthy and pathological areas while improving detection sensitivity [22]. MRI with DWI has significantly improved diagnostic accuracy and early cancer detection. Dynamic Contrast-Enhanced MRI (DCE-MRI), achieved through contrast agent injection, allows qualitative and quantitative assessment of tumor lesions. Although quantitative DCE-MRI is currently limited to clinical trials, ongoing efforts are developing standardized protocols for its clinical application [23]. Since it does not need ionizing radiation, it is much safer. However, MRI is more susceptible to motion artifacts, takes longer to acquire, is more expensive, and has limited availability.

EUS provides an assessment of the pancreatic parenchyma and ducts by imaging from within the gastrointestinal lumen. EUS and related procedures form the cornerstone of early pancreatic cancer diagnosis, owing to their exceptional diagnostic accuracy, safety, and versatility in obtaining tissue samples for cytological and histological evaluation [23]. It can detect deep local tumors that are difficult to be identified by abdominal ultrasound. The possibility of obtaining tissue samples through FNA biopsy, then the samples are examined pathologically to confirm the diagnosis, which achieves a sensitivity of approximately 92% [24]. Compared with CT or MRI, it has higher sensitivity in detecting intramedullary papillary mucinous tumors less than 1 cm [25]. AI-powered image classification significantly enhances endoscopic physicians' capabilities by providing additional safety checks, while also helping to reduce diagnostic discrepancies in early-stage PC cases. However, it has a high expenditure [26].

In summary, CT, MRI, and EUS form complementary pillars for diagnosing and staging pancreatic cancer. CT remains the primary choice for initial staging, MRI excels in characterization and resolving complex cases, while EUS demonstrates unparalleled capabilities in tissue sampling and local staging.

4.2 Clinical Decision Support System (CDSS)

Due to massive data and machine learning technologies, CDSS represents advanced applications of artificial intelligence in clinical diagnosis and treatment. These systems integrate multi-source medical records, literature, and clinical research data to conduct comprehensive analyses of drug efficacy, product accessibility, adverse reactions, patients' financial status, and insurance types [27]. This helps create personalized treatment recommendations. These recommendations help clinicians make better care plans. Because of this, the technology has moved from everyday uses to important medical fields. It now plays key roles in areas like imaging diagnostics, pathological analysis, clinical decision support, prognosis assessment, and drug screening. However, the data used must be very accurate and complete to offer helpful information.

5. Summary

This research summarizes the artificial intelligence in the early detection of pancreatic cancers, including the methods of ML, DL, multi-modal data analysis, explainable AI and assisted decision making as well as the practical applications of medical image analysis and CDSS. In addition, their certain drawbacks are listed which will provide thorough information about how and when to choose them. With the help of such technologies, the rate of missed diagnosis can be diminished due to its higher specificity and sensitivity, especially for those small features that are hard to detect by the humans' eyes. Besides, they can provide more objective assessment outcomes and decrease the time for diagnosis. Meanwhile, there are certain challenges in each technology which results in no absolute solutions to detect the early stage of pancreatic cancers nowadays. Therefore, first, large-scale clinical trials are required in the future to provide more detailed data through various aspects. Secondly, multi-modal data analysis should be modified to be better and construct a comprehensive system. Thirdly, more effort can be put on explainable AI to improve the cooperation of AI and doctors and boost the accuracy of diagnosis. In sum, there is no doubt that AI is remodeling the methods of the diagnosis of pancreatic cancers. By motivating new research and experiments related to this, more and more patients can be found as early as possible and hence enhance their surviving rate.

References

[1] Chen H, Wang L, Fan D, Ma P, Zhang X, Lin K. SERS mapping combined with explainable deep learning for exosome analysis to enhance lung cancer detection. Analyst. https://doi.org/10.1039/D5AN00685F

[2] Zhang L, Sanagapalli S, Stoita A. Challenges in diagnosis of pancreatic cancer. World Journal of Gastroenterology, 2018, 24(19): 2047-2060. https://doi.org/10.3748/wjg.v24.i19.2047

[3] Rahib L, Coffin T, Kenner B. Factors driving pancreatic cancer survival rates. Pancreas, 2025, 54(6): e530-e536. https://doi.org/10.1097/MPA.0000000000002489

[4] Health Network. National Taiwan University Hospital has launched the AI system "Zhu Pan Jian" to break through the bottleneck in early diagnosis of pancreatic cancer. UHO Health Network, 2025-05-22. https://www.uho.com.tw/article-66684.html

[5] Fornero S. What is ICD-10? Quadrant Health, 2024-07-10. https://blog.quadrant.health/what-is-icd-10

[6] Pancreatic Cancer ICD-10-CM codes. Carepatron. https://www.carepatron.com/icd/pancreatic-cancer

[7] ICD List. ICD-10-CM diagnosis code C78.89 - Secondary malignant neoplasm of other digestive organs. ICD List. https://icdlist.com/icd-10/C78.89

[8] Weselá P, Eid M, Moravčík P, Vlažný J, Hlavsa J, Procházka V, Kala Z, Vaňhara P. Artificial intelligence in pancreatic cancer histopathology and diagnostics: Implications for clinical decisions and biomarker discovery? Cell Division, 2025, 20(1). https://doi.org/10.1186/s13008-025-00158-w

[9] Huang B, Huang H, Zhang S, Zhang D, Shi Q, Liu J, Guo J. Artificial intelligence in pancreatic cancer. Theranostics, 2022, 12(16): 6931-6954. https://doi.org/10.7150/thno.77949

[10] Hameed B S, Krishnan U M. Artificial intelligence-driven diagnosis of pancreatic cancer. Cancers, 2022, 14(21): 5382. https://doi.org/10.3390/cancers14215382

ISSN 2959-409X

[11] Zhang M M, Yang H, Jin Z D, Yu J G, Cai Z Y, Li Z S. Differential diagnosis of pancreatic cancer from normal tissue with digital imaging processing and pattern recognition based on a support vector machine of EUS images. Gastrointestinal Endoscopy, 2010, 72(5): 978-985. https://doi.org/10.1016/j.gie.2010.06.042
[12] Kim P. MATLAB deep learning. Apress eBooks, 2017. https://doi.org/10.1007/978-1-4842-2845-6
[13] Liu Y H. Feature extraction and image recognition with convolutional neural networks. Journal of Physics Conference Series, 2018, 1087: 062032. https://doi.org/10.1088/1742-6596/1087/6/062032
[14] Kumar A, Kim J, Lyndon D, Fulham M, Feng D. An ensemble of fine-tuned convolutional neural networks for medical image classification. IEEE Journal of Biomedical and Health Informatics, 2016, 21(1): 31-40. https://doi.org/10.1109/jbhi.2016.2635663
[15] Siddique N, Paheding S, Elkin C P, Devabhaktuni V. U-Net and its variants for medical image segmentation: A review of theory and applications. IEEE Access, 2021, 9: 82031-82057. https://doi.org/10.1109/access.2021.3086020
[16] Coudray N, Ocampo P S, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira A L, Razavian N, Tsirigos A. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature Medicine, 2018, 24(10): 1559-1567. https://doi.org/10.1038/s41591-018-0177-5
[17] Vashisth V. What is multi-modal data analysis? Analytics Vidhya, 2025-07-25. https://www.analyticsvidhya.com/blog/2025/07/what-is-multi-modal-data-analysis
[18] Schuurmans M, Saha A, Alves N, Vendittelli P, Yakar D, Sabroso-Lasa S, Xue N, Malats N, Huisman H, Hermans J, Litjens G. End-to-end prognostication in pancreatic cancer by multimodal deep learning: A retrospective, multicenter study. European Radiology, 2025. https://doi.org/10.1007/s00330-025-11694-y
[19] Xia J, Li J, Wang X, Li Y, Li J. Enhanced cancer classification and critical feature visualization using Raman spectroscopy and convolutional neural networks. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2024, 326: 125242. https://doi.org/10.1016/j.saa.2024.125242
[20] Xin X, Danping H, Shanshan H, Yang S. Research on the diagnosis method of pancreatic lesions by endoscopic ultrasound based on twin network structure. IET Image Processing, 2025, 19(1). https://doi.org/10.1049/ipr2.70063
[21] Francis I R. Pancreatic adenocarcinoma: Diagnosis and staging using multidetector-row computed tomography (MDCT) and magnetic resonance imaging (MRI). Cancer Imaging, 2007, 7(Special issue A): S160-S165. https://doi.org/10.1102/1470-7330.2007.9010
[22] Ben-Amitay S, Jones D K, Assaf Y. Motion correction and registration of high b-value diffusion weighted images. Magnetic Resonance in Medicine, 2011, 67(6): 1694-1702. https://doi.org/10.1002/mrm.23186
[23] Luo W, Wang J, Chen H, Cao Z, Qiu J, Liu Y, Fu Y, Yang G, Tao J, Weng G, Liu T, Wang Y, Ye L, Ding C, Lu X, Zhang M, Huang H, Xiao J, You L, Zhang T. Advances in screening and early diagnosis of pancreatic cancer. Cancer Screening and Prevention, 2024, 3(3): 154-162. https://doi.org/10.14218/csp.2024.00006s
[24] Raut C P, Grau A M, Staerkel G A, Kaw M, Tamm E P, Wolff R A, Vauthey J N, Lee J E, Pisters P W T, Evans D B. Diagnostic accuracy of endoscopic ultrasound-guided fine-needle aspiration in patients with presumed pancreatic cancer. Journal of Gastrointestinal Surgery, 2003, 7(1): 118-128. https://doi.org/10.1016/S1091-255X(02)00150-6
[25] Gheonea D I, Săftoiu A. Beyond conventional endoscopic ultrasound: Elastography, contrast enhancement and hybrid techniques. Current Opinion in Gastroenterology, 2011, 27(5): 423-429. https://doi.org/10.1097/MOG.0b013e328349cfab
[26] Kaur S, Baine M J, Jain M, Sasson A R, Batra S K. Early diagnosis of pancreatic cancer: Challenges and new developments. Biomarkers in Medicine, 2012, 6(5): 597-612. https://doi.org/10.2217/bmm.12.69
[27] Shortliffe E H, Sepúlveda M J. Clinical decision support in the era of artificial intelligence. JAMA, 2018, 320(21): 2199-2200. https://doi.org/10.1001/jama.2018.17163