Machine Learning in PDAC: Applications Analysis

ET Pinto[1], GC Santos[1], T Brito[1]
[1] Post-Graduate Program in Biomedical Engineering at the Federal University of São Paulo.

Abstract

The development of Artificial Intelligence (AI) has been taking place in many applications nowadays. Although there might be hesitation when it comes to healthcare, especially to image disease diagnosis, since human eye diagnosis cannot be replaced, there are many efforts to Deep Learning (DL) increasingly becoming accepted among health agencies around the world. With regards to pancreatic ductal adenocarcinoma (PDAC), recent studies have shown that image processing and analysis through DL algorithms are considerably briefer procedures than visual analysis of experienced doctors. The main imaging techniques of diagnosis are Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). There are already many studies about applying AI to images from both of these imaging modalities, aiming to transpose the limitations for some image diagnoses like for PDAC. In this paper, we aimed to analyze articles and proposals of Deep Learning applications for the early detection of PDAC in CT images.

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is associated with a dire prognosis and a 5-year survival rate of only 11.5% (2012–2018). However, for those patients with localized cancer where the tumor is confined to the primary site, the 5-year relative survival rate is 39.4%. It is estimated that in 2022, there will be 62,210 new cases of PDAC and an estimated 49,830 will die of this disease (Cancer Stat Facts). The premise of improved survival through early detection is that more individuals will benefit from potentially curative treatment. Because symptoms typically occur late in the disease course, pancreatic cancer early detection will possibly require screening asymptomatic subjects. Although it remains expensive and challenging with current technology to screen the general population for PDAC, the ability to define high-risk groups with an increased likelihood of harboring such lesions may lead to an earlier interception and improved survival (Kenner et al., 2020).
Medical imaging is a technique used to look at the human body, and diagnose, monitor, or treat medical statuses, including a visual representation of the functions done by some organs or tissues. Imaging techniques help screen for hidden features before visible symptoms, diagnose the conditions that would have developed into the now visible symptoms, and manage the disease stages or reaction to possible treatment. Blood and other laboratory tests usually determine the existence of PDAC. In pancreas imaging, computed tomography (CT) and magnetic resonance imaging (MRI) are usually used to help determine if the condition exists. Tumor classification through these methods can also help to track, predict and endorse customized therapy as part of effective treatment, without cancer invasions. If detected, then determine the stages and reaction to treatment, making it possible to predict the patient's overall survival rate (Bakasa & Viriri, 2021).
Predictive technique models such as classic Machine Learning (ML) and Deep Learning (DL) can be used for PDAC survival rate prognosis. As referred to Deep Learning algorithms,

convolutional neural network (CNN) is a class of neural network models that can extract features from images by exploring the local spatial correlations presented in images. CNN models are effective and powerful for addressing a variety of image classification problems, including medical imaging (Ma et al., 202).

Therefore, in this paper, we aimed to analyze articles and proposals of Deep Learning applications, specifically CNNs, for the early detection of PDAC in CT images.

2. Methodology

Focusing on CT images and CNN processing, the project seeks techniques applied nowadays surrounding early diagnosis, detection, and classification of PDAC. Therefore, the chosen literature is summarized below with its methodologies and common parameters utilized.

Xuan & You (2020) worked with a Hierarchical Convolutional Neural Network (HCNN) aiming to predict PDAC in the early stages with an end-to-end ML model, which has superior results and benefits in terms of performance and computational cost. The researchers applied CNN-RNN ML techniques, a P-Net model adapted for 3D images, to segmentation of the tumor area, and then HCNN to predict tumors. They tested the model with datasets beginning from 10 to 50 CT datasets. Its validation was done by dice index, sensitivity, and specificity rates.

Sekaran et al. (2019) invested in the CNN approach too, but the data was analyzed by the Gaussian Mixture Model (GMM) balanced by the EM algorithm; the authors called this part Lump Feature Extraction (LFE). This preprocessing helped to adjust the parameters of the CNN model and the features obtained served as input to it; the authors nominated that part as the Lump Recognition algorithm (LR) developed to identify the lump. Another input to the developed CNN model was a set of threshold parameters of lump growth, making a CNN-GMM method, predicting cancer on CT images by the percentage spread of the lesion on the pancreas. The dataset used consisted of 19,000 images from the Cancer Imaging Archive (TCIA). The study evaluated the model by recognition ratio versus the number of images on the training dataset.

Liu et al. (2020) did a retrospective study on pancreatic cancer detection by pattern recognition with ML, applying an edited form of the CNN model Visual Geometry Group (VGG) network. It was composed of three convolutional blocks in which each block consisted of two convolutional layers, a rectified linear unit, and a max-pooling layer at the end. Then, a flatten node was added to the last convolution block as three fully connected layers were attached at the end of the model. Aiming to balance the classification, a weighted binary cross-entropy was used as the loss function. For the process of classification to be optimized, two callbacks were added to the model, based on the loss function. They trained and tested the CNN model on a dataset consisting of 14,780 CT images, where 7,557 are controls. The success of the project was measured by sensitivity, specificity, and accuracy rates and area under the Receiver Operating Characteristic (ROC) curve (AUC).

Liu et al. (2019) studied the prediction of PDAC using a faster region-based CNN model (Faster R-CNN). First, they utilized a VGG16, a pre-trained in ImageNet containing 13 convolutional layers and three fully connected layers, for feature extraction; its output fed the

CNN model chosen. The study analyzed more than 5,000 CT images, either from pancreatic cancer patients or not, and the model's performance was measured by AUC.
Ma et al. (2020) also developed a CNN model for binary classification applied to PDAC diagnosis. They had more than 6,000 CT images of two distinct groups: patients already diagnosticated and other volunteers with normal pancreas, both randomly selected in a bigger dataset. The strategy developed was three convolutional layers and one fully connected, and then a batch normalization (BN), a rectified linear unit (ReLU), and a max-pooling layer. The model was evaluated using accuracy, specificity, and sensitivity rates.
Between the chosen articles, there is some resemblance: they all used CT images as the data to develop the CNN models and the acquisition of these images was made with contrast application. Even so, all the studies had their parameters and forms of success evaluation, such as accuracy, the ROC curve, AUC, the Dice-performance ratio, and others. Therefore, in the Results section, each article has an individual approach.

3. Results

This section aims to display the results achieved by the methods applied in the chosen papers. Therefore, the results are organized in Table 1.

Table 1. Summarized results of literature.

Article	Techniques	Results
Xuan & You, 2020.	HCNN for detection and CNN-RNN for segmentation.	- 95.10% Stabilized precision ratio; - 97.33% Dice index; - 97.66% Sensitivity Ratio; - 96.12% Specificity Ratio.
Sekaran et al., 2019.	Gaussian Mixture Model (GMM) with Expectation Maximization (EM) for feature extraction and a CNN model for classification.	The recognition ratio using 1,000 CT images to feature extraction on model training was 99.9%.
Liu et al., 2020.	CNN model derived from the Visual Geometry Group (VGG) network for classification.	<u>Local Test set 1</u> - 97.3% Sensitivity Ratio; - 100% Specificity Ratio; - 98.7% Balanced Accuracy. <u>Local Test set 2</u> - 99.0% Sensitivity Ratio; - 98.9% Specificity Ratio; - 98.9% Balanced Accuracy. <u>External Test set</u> - 79.0% Sensitivity Ratio; - 97.6% Specificity Ratio; - 88.3% Balanced Accuracy.
Liu et al., 2019.	Faster region CNN: feature	- AUC: 0.9632.

	extraction, region proposal network (RPN), proposal classification and regression network.	
Ma et al., 2020.	Three convolutional layers and one fully connected, and then a batch normalization (BN), a rectified linear unit (ReLU), and a max-pooling layer.	Binary classifiers Accuracy: 95.4%; Sensitivity: 98.3%; Specificity: 91.6%. Ternary classifiers Accuracy: 82.1%; Sensitivity: 98.6%; Specificity (tail/body): 52.0%; Specificity (head/neck): 46.2%.

Source: the authors.

The study of Sekaran et al. (2019) compared its method of lump recognition with other techniques used for the same goal like Gray Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), K-Means, and K-Medoids. Between these methods, the one which got closer to LFE + LR with GMM + CNN was GLCM, with a 91.9% of recognition rate with 1,000 CT images.

Liu et al. (2020) did the classification analysis with the CNN model in two ways a (1) patch-based and a (2) patient-based and did the current process of diagnosis with specialized radiologists for the two Local Test sets. The radiologists' performances were lower than the ML technique developed, achieving 94.4% and 91.7% of the Sensibility Ratio, respectively for the sets.

Liu et al. (2019) did a three-step algorithm that consisted of features extraction; RPN, in order to identify the regions of interest (ROI); and then proposal classification and regression network, in which coordinates and probability scores are assigned to regions of the image. It was evaluated by the AUC, which was calculated to be 0.9632.

Ma et al. (2020) evaluated their model in both binary and ternary classifiers. The former was compared to the analysis of imaging specialists and had a performance, indicating that it can be utilized for the classification of cancerous or non-cancerous mass. Ternary classifiers, in turn, have the goal of localizing the mass in the pancreas. That could be used by doctors to choose the surgical technique to be done.

4. Discussion

In the “Detection and Diagnosis of pancreatic tumor using HCNN” article (Xuan, W., & You, G., 2020), the combination of validation and training subsets together were used to further refine the network until the validation subset’s performance converges into accuracy. The validated subset also represents the design of the P-Net, to indicate training accuracy after the model selection. As a result, the proposed HCNN method achieves 95.1% of Stabilized Precision ratio, which is therefore assessed in the confidence intervals between impact estimates and the predictive potential of proposed indices. Also, computer-aided diagnostics (CAD) systems are designed to assist radiologists by offering a second option to reduce understanding of volume time, variance, errors, and booting specificity and sensitivity in this context. With this purpose, this method achieved a dice index of 97.33%, a Sensitivity Ratio of 97.66%, and a Specificity Ratio of 96.12%.

The method adopted by Sekaran et al (2019) analyzed the lump of the pancreas by its size, shape, depth, length, and weight parameters. The result was a trained system for the effective identification of the lump and they measured the model performance by the recognition rate of the lump versus the quantity of CT images utilized in the training process. Compared with other techniques of lump recognition, the LFE + LR with GMM + CNN method had the best results, achieving a 99.9% recognition ratio with 1,000 images on the training set.

In the proposal of a VGG network utilized by Liu et al. (2020) the classification occurred in the same way for both strategies (1) and (2): using the cutoff that achieved the highest Youden index in the ROC curve that was constructed using the sets. The validation form for all analyses was done by sensitivity, specificity, and accuracy of the model on the set of data separated to the test process and the rates of the CNN model are higher than the human performance of diagnosing the samples, achieving an accuracy of nearly 99%. And, between the patch-based and patient-based analysis, the second one has had greater results with 95% of confidence.

In Liu et al. (2019), the features used for the algorithm were age, sex, number of metastatic lymph nodes via CT, number of metastatic lymph nodes via pathology, N stage (tells whether cancer has spread to the nearby lymph nodes), degree of tumor differentiation, presence or absence of intravascular tumor thrombus, and presence or absence of nerve infiltration. Images were analyzed by senior experienced radiologists (5 years of working experience) that used the following labels for classification: pancreatic tumor, normal pancreatic tissue, chronic pancreatitis, and benign pancreatic tumor. A clear advantage of the AI method was the time required for analysis, which was about 3 seconds rather than 8 minutes for imaging specialists.

5. Conclusions

As briefly shown in this paper, Deep Learning techniques are very promising in PDAC detection, however, there are still many considerations to be taken if one desires to use it on a large scale. Even though the performance measurements such as AUC of the ROC, accuracy, specificity, and sensitivity are high enough to believe and be enthusiastic about applying DL algorithms for pancreatic cancer detection purposes, limitations such as concentrated population sampling and short health history of patients don’t allow modeling the entire humankind. It might even be better to develop models demographically, for example.

References

Bakasa, W. and Viriri, S., 2021. Pancreatic Cancer Survival Prediction: A Survey of the State-of-the-Art. *Computational and Mathematical Methods in Medicine*, *2021*.

Cancer Stat Facts: Pancreatic Cancer. Available at: https://seer.cancer.gov/statfacts/html/pancreas.html. Accessed December 8, 2022.

Kenner, B., Chari, S. T., Kelsen, D., Klimstra, D. S., Pandol, S. J., Rosenthal, M., Rustgi, A. K., Taylor, J. A., Yala, A., Abul-Husn, N. and Andersen, D. K., 2021. Artificial intelligence and early detection of pancreatic cancer: 2020 summative review. *Pancreas*, *50*(3), p.251.

Liu, K.-L., Wu, T., Chen, P.-T., Tsai, Y. M., Roth, H., Wu, M.-S., … Wang, W. (2020). Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation. The Lancet Digital Health, 2(6), e303–e313. doi:10.1016/s2589-7500(20)30078-9

Liu, S.-L., Li, S., Guo, Y.-T., Zhou, Y.-P., Zhang, Z.-D., Li, S., & Lu, Y. (2019). Establishment and application of an artificial intelligence diagnosis system for pancreatic cancer with a faster region-based convolutional neural network. Chinese Medical Journal, 1. doi:10.1097/cm9.0000000000000544

Ma, H., Liu, Z. X., Zhang, J. J., Wu, F. T., Xu, C. F., Shen, Z., Yu, C. H., & Li, Y. M. (2020). Construction of a convolutional neural network classifier developed by computed tomography images for pancreatic cancer diagnosis. World journal of gastroenterology, 26(34), 5156–5168. https://doi.org/10.3748/wjg.v26.i34.5156

Sekaran, K., Chandana, P., Krishna, N. M., & Kadry, S. (2019). Deep learning convolutional neural network (CNN) With Gaussian mixture model for predicting pancreatic cancer. Multimedia Tools and Applications, 79(15-16), 10233–10247. doi:10.1007/s11042-019-7419-5

Xuan, W., & You, G. (2020). Detection and diagnosis of pancreatic tumor using deep learning-based hierarchical convolutional neural network on the internet of medical things platform. Future Generation Computer Systems. doi:10.1016/j.future.2020.04.037