SAR·DFP
Society of Abdominal Radiology
Disease-Focused Panels

Role of Artificial Intelligence (AI) in Pancreatitis

Pancreatitis Disease Focused Panel
Society of Abdominal Radiology
Educational material

Authors:
Kade Murphy BS, Yoganand (Yoga) Balagurunathan PhD and Harshna V. Vadvala MD

Layout

- **Introduction to Artificial Intelligence**
  - Overview (AI vs Machine Learning)
  - Types
    - Decision tree
    - Random forest
    - Support Vector Machines
    - Artificial Neural Networks
    - Convolutional Neural Networks
- **Acute Pancreatitis**
  - Diagnostic criteria
  - Clinical scoring systems
  - Utility of Artificial Intelligence
- **Summary of current research on utility of AI in AP**
  - Diagnosis of pancreatitis
    - Abilities of SVM and CNN modeling
  - Prediction of severity of pancreatitis
  - Radiomics
  - CAD (computer assisted diagnosis)
- **Usefulness of A.I in Clinic**
- **Limitations & Future Directions**

Perspective: *Why AI Now?*

- It is projected that AI has a potential to improve healthcare outcome by about 30% to 40%, simultaneously cutting treatment costs to half (*Artificial Intelligence in Healthcare, 2019*).

- Data revolution of 21st century has created a need to develop new technologies and there is a renewed focus to use AI/ML. An estimated compounded annual growth rate of about 37.3% expected from 2023 to 2030 (*McKinsey Global Institute & Forbes 2018).*

- AI publications represented 3.8% of all peer-reviewed scientific publications worldwide in 2019, up from 1.3% in 2011 (Elsevier/Scopus, 2020).

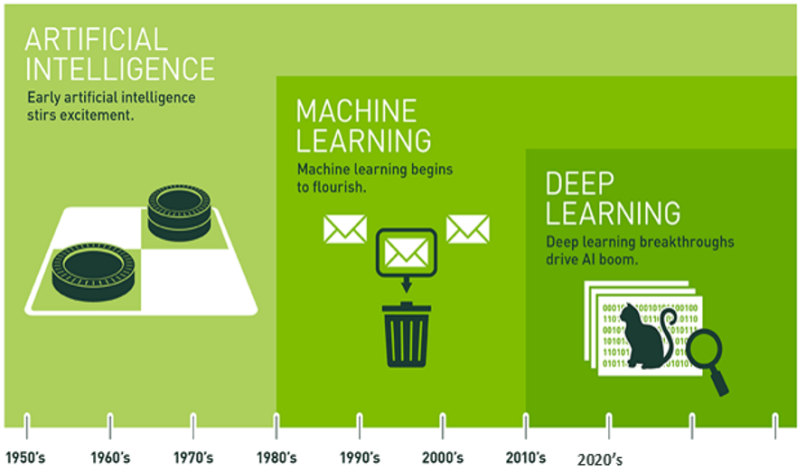
What's the difference: *Artificial Intelligence vs Machine Learning*

Artificial Intelligence: Machines that replicate human thought, emotion, and reason (and remain, for now, in the realm of science fiction); and narrow AI, technologies that can perform specific tasks as well as, or better than, humans. (*Jones et al., BJR, 7 (30), 2018*)

Lay terms: ability of a computer or an algorithm to demonstrate any form of "human" intelligence (i.e. problem solving, learning).

Machine Learning: Machine learning (ML) is the study of computer algorithms that can learn complex relationships or patterns from empiric data and make accurate decisions. (*Jones et al., BJR, 7 (30), 2018*)

Lay terms: The ability of a computer to be programmed to analyze data, learn patterns, make predictions, and subsequently improve performance via experience. Examples include; Decision Trees, Random Forests, Support Vector Machines, Artificial Neural Networks, Convolutional Neural Networks and many more.



Pictorial Courtesy: www.nvidia.com

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Overview: Machine Learning Methods

Most popular ML methods can be broadly categorized into Supervised and Un-Supervised learning. Examples of supervised learning are Tree based decision (Decision trees, Random forest, XGBoost), Discriminant function, Support vector machines and unsupervised include clustering methods.

- Decision Trees
  - A method of classifying information where data is progressively split along decision ("root") node into two or more mutually exclusive subsets. This continues until data encounters a "leaf" node, where the final result of all prior decision rules is established.
  - After a decision tree has been created with training data, it can be used to attempt to classify new samples.
  - Advantages: Simplicity and relative ease of understanding decision algorithms
  - Disadvantages: Sensitive to the training data, and large variance in accuracy when classifying new samples.

Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. 2015;27(2):130-5.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Machine Learning Methods: Tree Based

- Random Forest
  - An algorithm that uses many individually trained decision trees built **independently** from data to make decisions.

- XGBoost
  - An algorithm that uses many individually trained decision trees built **sequentially** from data, taking previously incorporated trees into account, to make decisions

- Advantages: Less sensitive to the training data and better able to classify new information than individual decision trees.
- Disadvantages: With increasing complexity, the underlying decision process can be less comprehensible than decision trees.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Shehadeh A, Alshboul O, Al Mamlook RE, et al. Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, lightgbm, and XGBoost regression. *Automation in Construction*. 2021;129:103827.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Machine Learning Methods: SVM

- Support Vector Machines (SVM)
  - Data is plotted in a multi-dimensional space, with the coordinates representing data inputs (“features”). The SVM then draws a “decision boundary” through the data, with the goal of maximizing the space between the data of the two categories (the “margin”).

  - Advantages: Great for distinguishing between two classes.
  - Disadvantages: To distinguish between three or more classes of data, more complex methods such as one-vs-one, one-vs-rest, or one-vs-all SVMs must be utilized.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Machine Learning Methods: Neural Networks

- Artificial Neural Networks (ANN)
  - Based loosely around the idea of interconnected neurons within the human brain, the input data is processed through one or more additional hidden layers where data is progressively modified.
  - In the output layer, the most “excited” neuron is responsible for the output classification.
  - “Learning” occurs when, through progressive modification of the weights and biases applied at each node with training data, the computer is able to provide an accurate output prediction based on novel inputs.

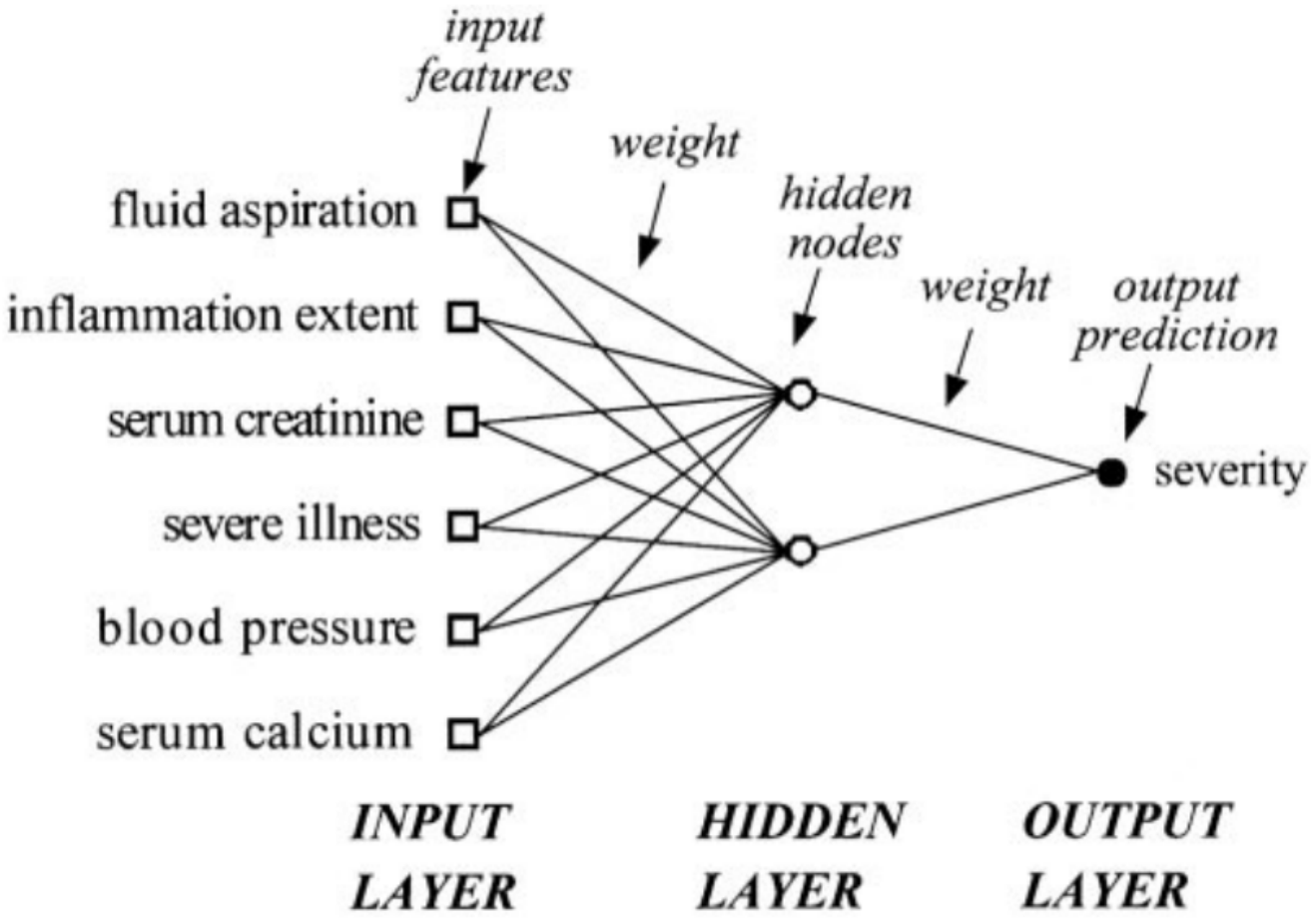
  - Advantage: Able to process large amounts of complex data
  - Disadvantage: Interpretability of the rationale for predictions is limited due to the interior complexity of the algorithm (“black box”).

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Mofidi R, Duff MD, Madhavan KK, et al. Identification of severe acute pancreatitis using an artificial neural network. *Surgery*. 2007;141(1):59–66.

Artificial Neural Network (ANN) Example



Keogan MT, Lo JY, Freed KS, et al. Outcome analysis of patients with acute pancreatitis by using an artificial neural network. *Academic Radiology*. 2002;9(4):410–9.

Convolutional Neural Networks (CNN)

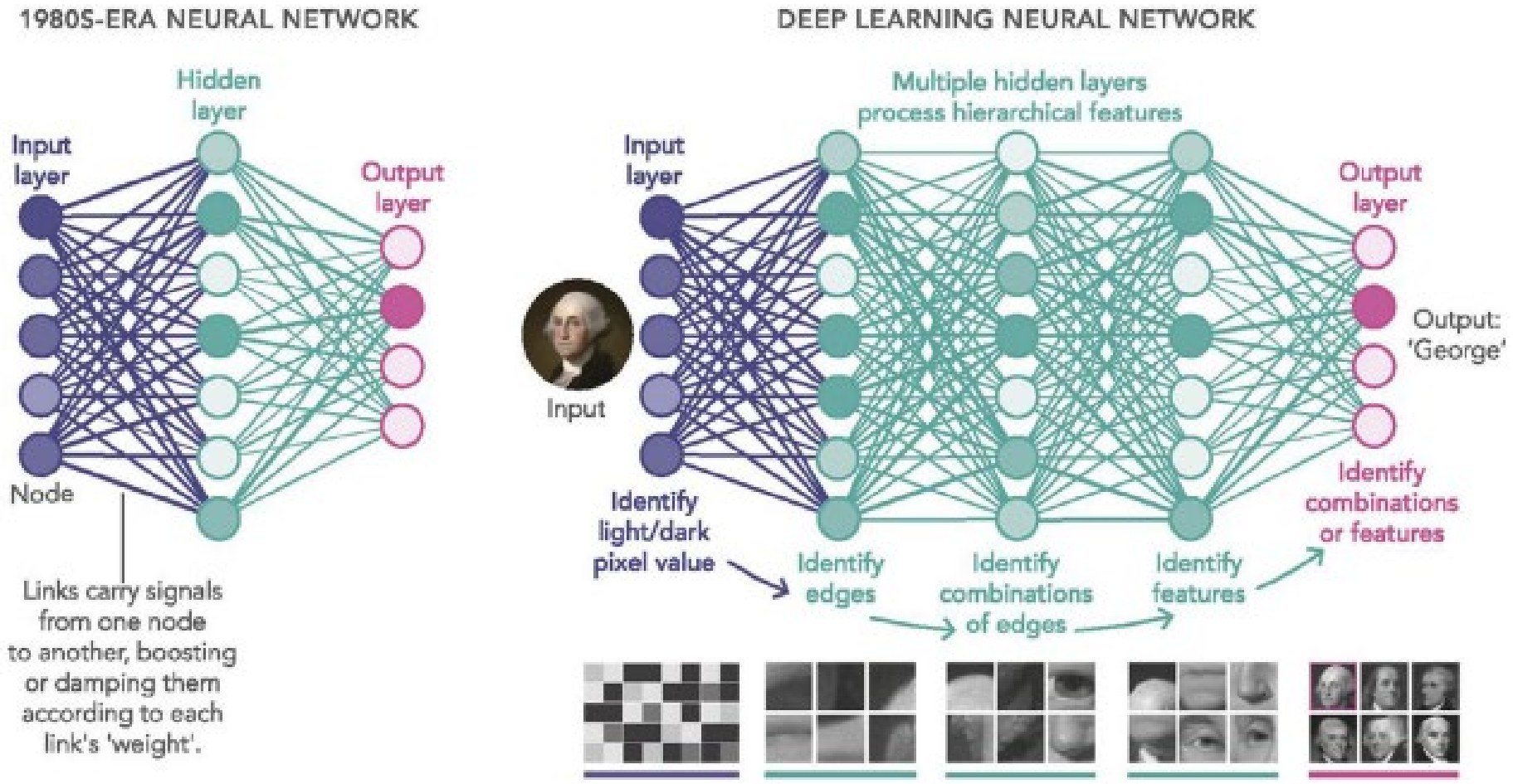
- ANN (Artificial Neural Networks) are useful in solving complex (non-linear) problems. While CNN (Convolutional Neural Networks) that are primarily used in computer vision problems. Use of Convolutional layer makes these network unique to solve vision-related applications. CNNs are fully connected feed forward neural networks that are effective in reducing the number of parameters and maintaining the quality of the models.
- To detect patterns in images, CNNs use a series of convolution operations at each layer to transform data before outputting to the next layer. Each of these layers can have filters which detect progressively more complex features until the entire structure is able to be identified.

- Advantage: Detection of complex visual structures and images
- Disadvantage: Interpretability of the rationale for predictions is limited due to the interior complexity of the algorithm (“black box”), similar to ANNs.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Neural Networks Visualized



Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Acute Pancreatitis: Clinical Diagnosis

- Diagnosis of AP
  - Atlanta criteria

    AP is diagnosed when two of the three are met:
    - abdominal pain suggestive of pancreatitis,
    - serum amylase and/or lipase at least three times normal levels,
    - findings on imaging suggestive of pancreatitis

Colvin SD, Smith EN, Morgan DE, et al. Acute pancreatitis: An update on the revised Atlanta Classification. *Abdominal Radiology*. 2019;45(5):1222–31.

Acute Pancreatitis: Severity Scoring

- Major Clinical Scoring Systems
  - Ranson score:
    - High specificity resulting in larger utility in detecting extremes of disease, with uncertain significance of intermediate results.
    - Score >4 at 48 hrs after admission has good performance in detecting in-hospital mortality
  - Acute Physiology and Chronic Health Evaluation II (APACHE II):
    - Despite high diagnostic accuracy, it is a complex and time-consuming scoring system, resulting in limited utility for routine clinical use.
  - Modified Glasgow Score (MGS)
    - Has the greatest overall sensitivity with development of severe acute pancreatitis.

Venkatesh NR, Balasubramaniyan G, et al. Comparison of different scoring systems in predicting the severity of acute pancreatitis: A prospective observational study. *Cureus*. 2020;12(2):e6943.

Leung T-K. Balthazar Computed Tomography severity index is superior to Ranson criteria and Apache II scoring system in predicting acute pancreatitis outcome. *World Journal of Gastroenterology*. 2005;11(38):6049.

Colvin SD, Morgan DE, et al. Acute pancreatitis: An update on the revised Atlanta Classification. *Abdominal Radiology*. 2019;45(5):1222–31.

Acute Pancreatitis: Severity Scoring

- Balthazar Score
  - A grade of pancreatitis based on radiologic features
    - Grade A (0 points): normal pancreas
    - Grade B (1 point): enlarged pancreas
    - Grade C (2 points): inflammation in pancreas or peripancreatic fat
    - Grade D (3 points): one ill defined peripancreatic fluid collection
    - Grade E (4 points): two or more ill defined peripancreatic fluid collections
  - Occasionally used independently and separately from the CT Severity Index (CTSI) as a measure of pancreatitis severity

Leung T-K. Balthazar Computed Tomography severity index is superior to Ranson criteria and Apache II scoring system in predicting acute pancreatitis outcome. *World Journal of Gastroenterology*. 2005;11(38):6049.

Balthazar EJ. Acute pancreatitis: Assessment of severity with clinical and CT evaluation. *Radiology*. 2002;223(3):603–13.

Acute Pancreatitis: Severity Scoring

- Pancreatic necrosis score:
  - 0 points: 0% necrosis
  - 2 points: <30% necrosis
  - 4 points: 30-50% necrosis
  - 6 points: >50% necrosis

- CT Severity Index (CTSI)
  - The sum of both the **Balthazar Score** and **pancreatic necrosis score**.
  - Mild acute pancreatitis: 0-3 points
  - Moderate acute pancreatitis: 4-6 points
  - Severe acute pancreatitis: 7-10 points

Bollen TL, Maurer R, et al. A comparative evaluation of radiologic and clinical scoring systems in the early prediction of severity in acute pancreatitis. *American Journal of Gastroenterology*. 2011;107(4):612-9.

Acute Pancreatitis: Utility of AI

- Acute pancreatitis is associated with significant morbidity and mortality risks, and identification of high risk patients earlier improves outcomes.

- If AI can more accurately predict AP severity and risk of complications, earlier decisions can be made regarding provision of additional therapies or earlier transfer to specialized units, therefore improving outcomes.

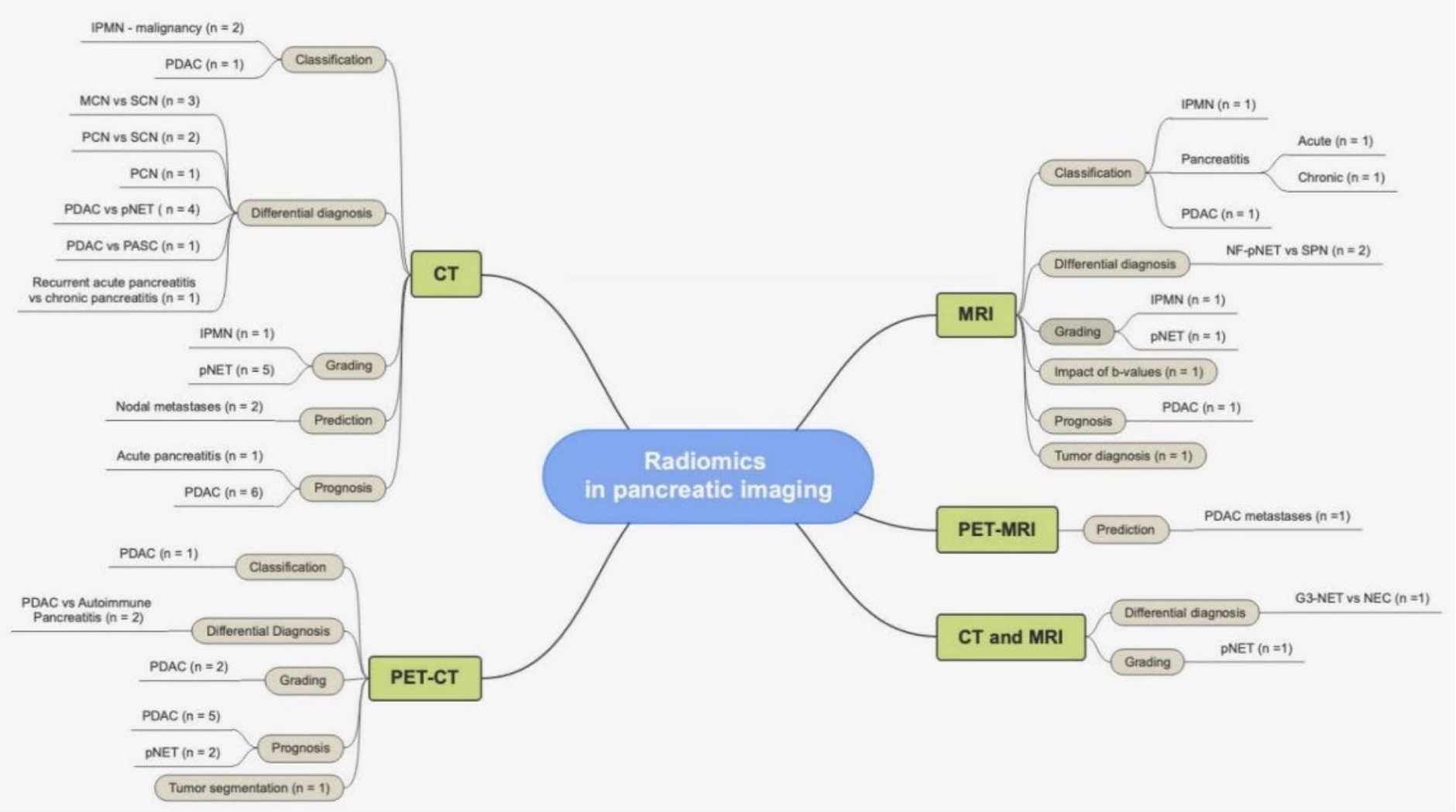
Venkatesh NR, Vijayakumar C, Balasubramaniyan G, et al. Comparison of different scoring systems in predicting the severity of acute pancreatitis: A prospective observational study. *Cureus*. 2020;12(2):e6943.

Mofidi R, Duff MD, Madhavan KK, et al. Identification of severe acute pancreatitis using an artificial neural network. *Surgery*. 2007;141(1):59–66.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Summary of Research

Radiomics applications in pancreatic imaging

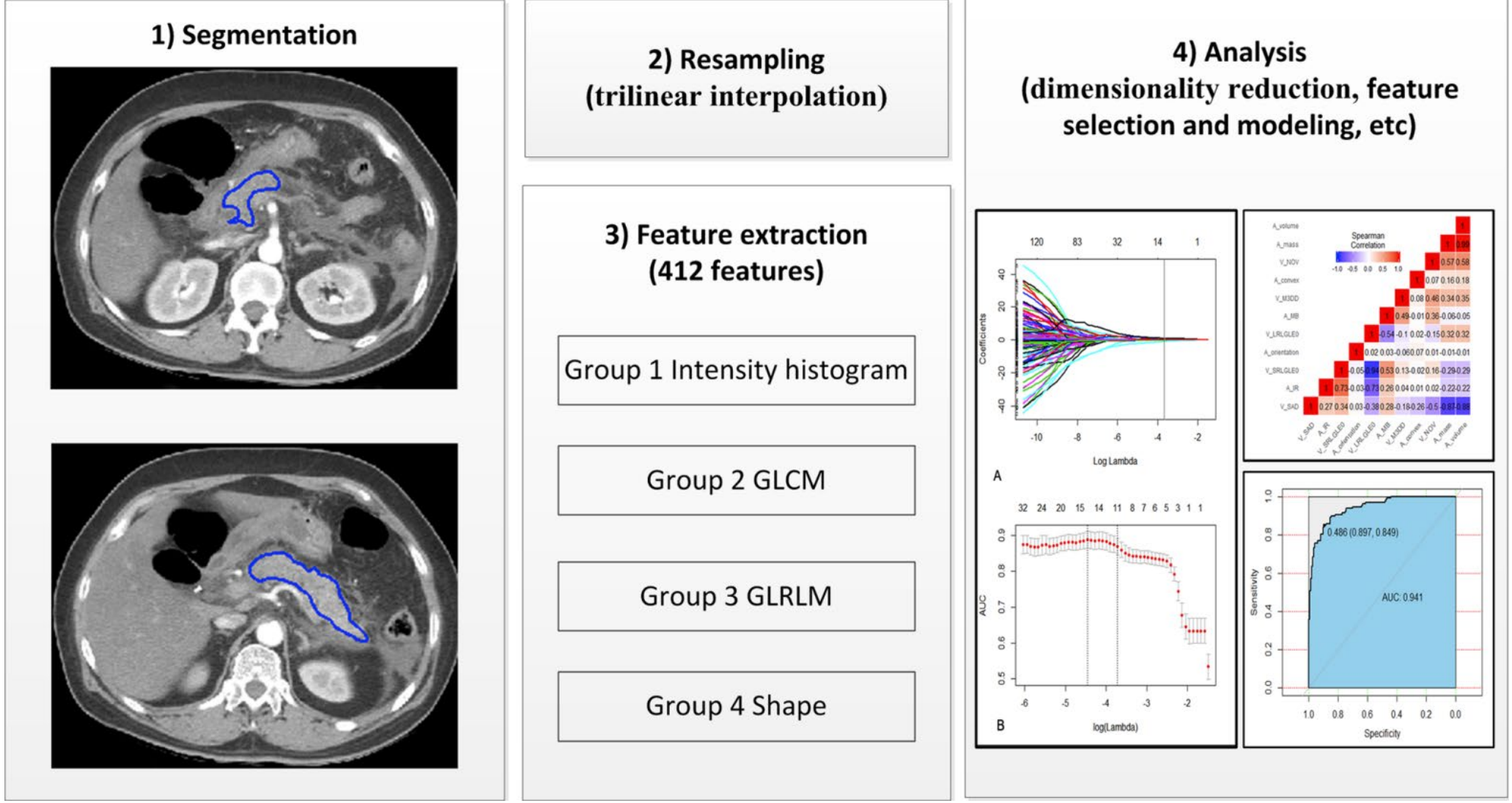


Laino ME, Ammirabile A, Lofino L, et al. Artificial Intelligence applied to Pancreatic Imaging: A narrative review. *Healthcare*. 2022;10(8):1511.

Radiomics in Pancreatic Imaging

- In 2018, a prospective cohort study was completed by Chen et al. using 389 first-attack acute pancreatitis patients.
- A radiomic model was developed using texture features of the pancreas on initial CT scans and used to predict recurrence of acute pancreatitis.
- The radiomics model containing 10 features strongly outperformed the clinical model in prediction of recurrent attacks of AP (AUC of 0.929 vs 0.671).

Chen Y, Chen TW, Wu CQ, et al. Radiomics model of contrast-enhanced computed tomography for predicting the recurrence of acute pancreatitis. *European Radiology*. 2018;29(8):4408–17.



A sequential diagram of the process of segmentation, radiomic analysis to extract specific features, and analysis of aforementioned features to find those which are most predictive recurrence of acute pancreatitis.

Chen Y, Chen TW, Wu CQ, et al. Radiomics model of contrast-enhanced computed tomography for predicting the recurrence of acute pancreatitis. *European Radiology*. 2018;29(8):4408–17.

Radiomics to differentiate AIDP and PDAC

- In 2019**, Zhang et al.** used 2D and 3D PET/CT images from 111 patients to evaluate whether quantified radiomics could **differentiate between AIP and PDAC.** Multiple feature selection strategies were utilized and compared.
- Multidomain, 3D, CT features were shown to be superior to single domain, 2D, PET features respectively.
- Overall, the **SVM-RFE feature selection strategy** and **linear SVM classifier** had the **highest diagnostic performance** (AUC=0.93).
- For the differentiation between AIP and PDAC in F-FDG PET/CT images, the **quantified radiomics model** was concluded to be **significantly superior in accuracy and specificity to human doctors as well as** clinical prediction models.

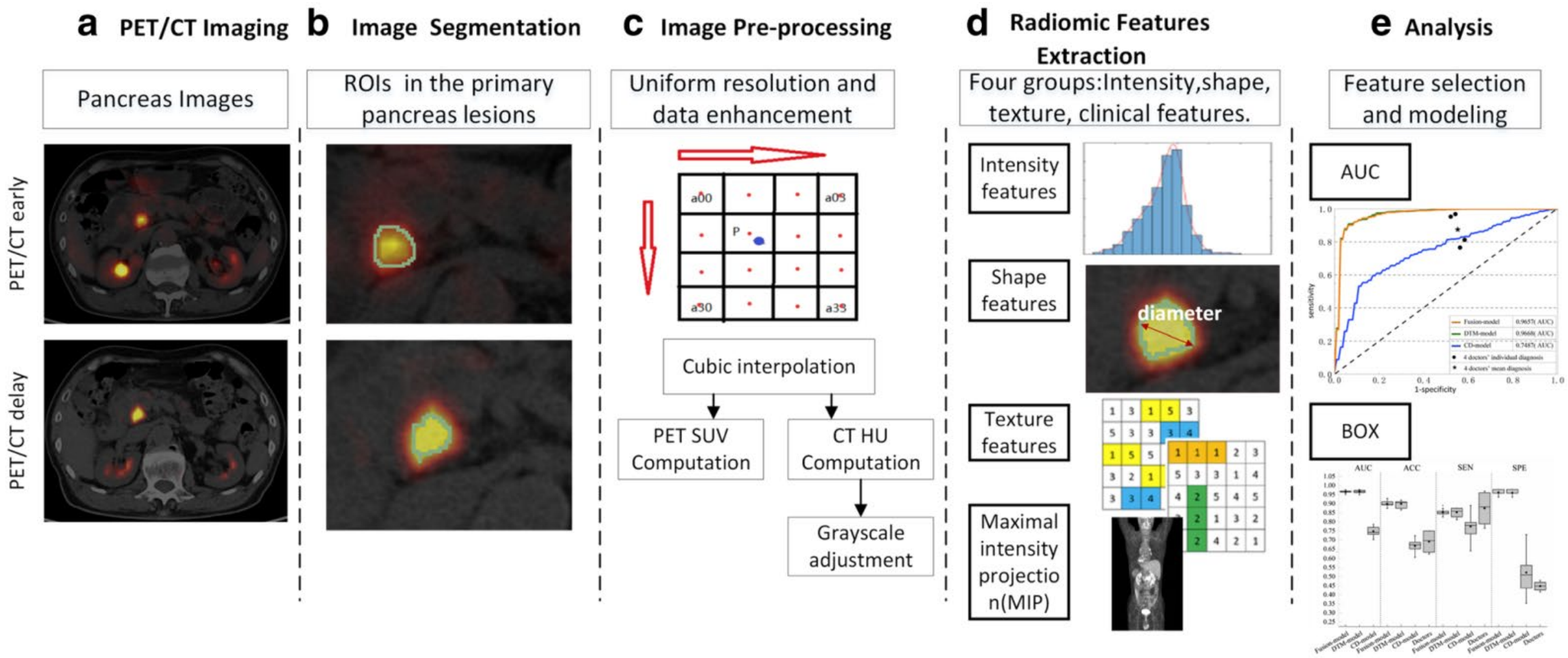
Zhang Y, Liu Z, et al. Radiomics analysis for the differentiation of autoimmune pancreatitis and pancreatic ductal adenocarcinoma in F-FDG PET/CT. *Medical Physics*. 2019;46(10):4520–30.

Radiomics to differentiate AIDP and PDAC

- In 2021, **Liu et al.** developed another radiomics-based prediction model to **distinguish PDAC from AIP** in 112 patients.
- Radiomic features were extracted and analyzed using SVM-RFE and linear SVM models.
- Overall, the model obtained an average AUC of 0.9668, an accuracy of 89.91%, a sensitivity of 85.31%, and a specificity of 96.04%.
- Therefore, **the radiomics model** based on F-FDG PET/CT dual time images demonstrated **good performance in discriminating between patients with AIP and PDAC lesions**, showing promise for clinical use.

Liu Z, Li M, Zuo C, et al. Radiomics model of dual-time 2-[18F]FDG PET/CT imaging to distinguish between pancreatic ductal adenocarcinoma and autoimmune pancreatitis. *European Radiology*. 2021;31(9):6983–91.

Process of Radiomic Feature Extraction and Analysis



Liu Z, Li M, Zuo C, et al. Radiomics model of dual-time 2-[18F]FDG PET/CT imaging to distinguish between pancreatic ductal adenocarcinoma and autoimmune pancreatitis. *European Radiology.* 2021;31(9):6983–91.

Radiomics in prediction of AP severity

- In 2020, **Lin et al.** developed a radiomics model using portal venous phase MRI with contrast to predict **the severity of acute pancreatitis** in the early phases of disease.
- Their radiomic model had an AUC of 0.848, **outperforming conventional scoring systems (MRSI, APACHE II, and BISAP** with AUC of 0.719, 0.725, and 0.708 respectively)

Lin Q, JI Y, Chen Y, et al. Radiomics model of contrast-enhanced MRI for early prediction of acute pancreatitis severity. *Journal of Magnetic Resonance Imaging*. 2019;51(2):397–406.

Radiomics in classification of chronic pancreatitis

- In 2020, **Frøkjær et al**. used radiomic texture analysis of MRIs from 99 patients to develop a model for **the classification of chronic pancreatitis vs healthy controls.**
- To this end, model achieved a 97% sensitivity, 100% specificity, and 98% accuracy, **demonstrating the potential for clinical utility in diagnosis**.

Frøkjær JB, Lisitskaya MV, Jørgensen AS, et al. Pancreatic magnetic resonance imaging texture analysis in chronic pancreatitis: A feasibility and validation study. *Abdominal Radiology*. 2020;45(5):1497–506.

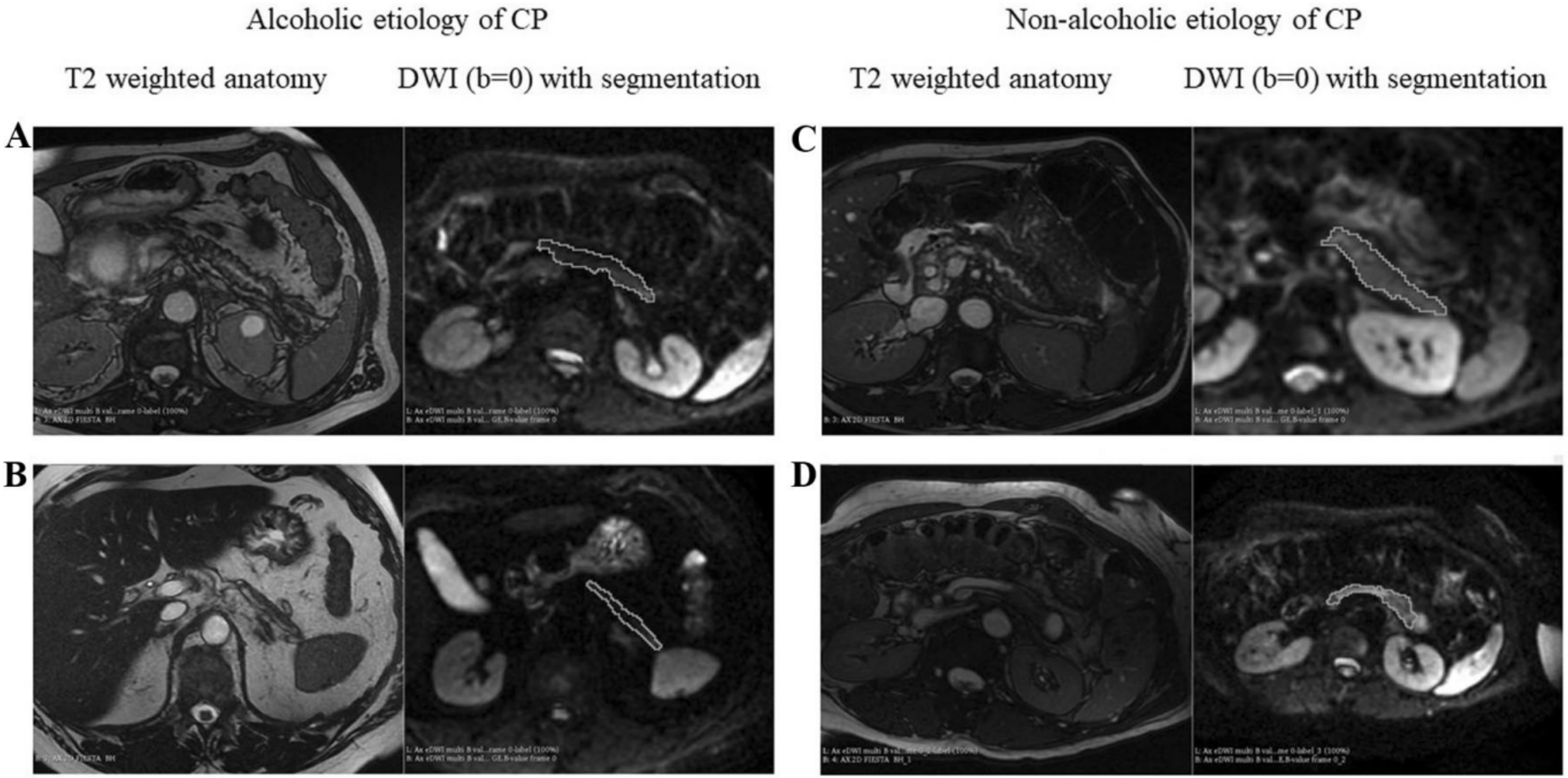
Radiomics in classification of chronic pancreatitis

- Subgroup classification was also performed, allowing for additional evaluation of alcoholic vs non-alcoholic etiology of CP, tobacco use, diabetes, and exocrine pancreatic dysfunction.
- Even with limited sample sizes, **the model demonstrated utility in classifying patients into etiology sub-groups** as well.

Classifier	Number of selected features	Accuracy	Sensitivity	Specificity	PPV
Disease (CP vs. healthy controls)	5	0.9798	0.9740	1.0000	1.0000
Alcohol (alcoholic vs. nonalcoholic etiology of CP)	9	0.8831	0.9111	0.8438	0.8913
Tobacco (use of tobacco vs. no use of tobacco)	10	0.8571	0.8627	0.8462	0.9167
Diabetes (diabetes vs. no diabetes)	4	0.8310	0.7143	0.8800	0.7143
Exocrine pancreatic function (PEI vs. normal exocrine function)	3	0.8219	0.8148	0.8421	0.9362

CP chronic pancreatitis, *PEI* pancreatic exocrine insufficiency, *PPV* positive predictive value

Frøkjær JB, Lisitskaya MV, Jørgensen AS, et al. Pancreatic magnetic resonance imaging texture analysis in chronic pancreatitis: A feasibility and validation study. *Abdominal Radiology*. 2020;45(5):1497–506.



T2 and DWI MRs demonstrate the texture feature values of both alcoholic and non-alcoholic etiology of chronic pancreatitis.
A and B represent the values furthest from the non-alcoholic etiology group.
C and D represent the average feature values of the non-alcoholic etiology group.

Frøkjær JB, Lisitskaya MV, Jørgensen AS, et al. Pancreatic magnetic resonance imaging texture analysis in chronic pancreatitis: A feasibility and validation study. *Abdominal Radiology*. 2020;45(5):1497–506.

Computer Assisted Diagnosis in Pancreatic Imaging

- **Computer assisted diagnosis (CAD)** has also shown great promise in recent years in the support of **disease diagnosis, classification, and pancreatic segmentation.**

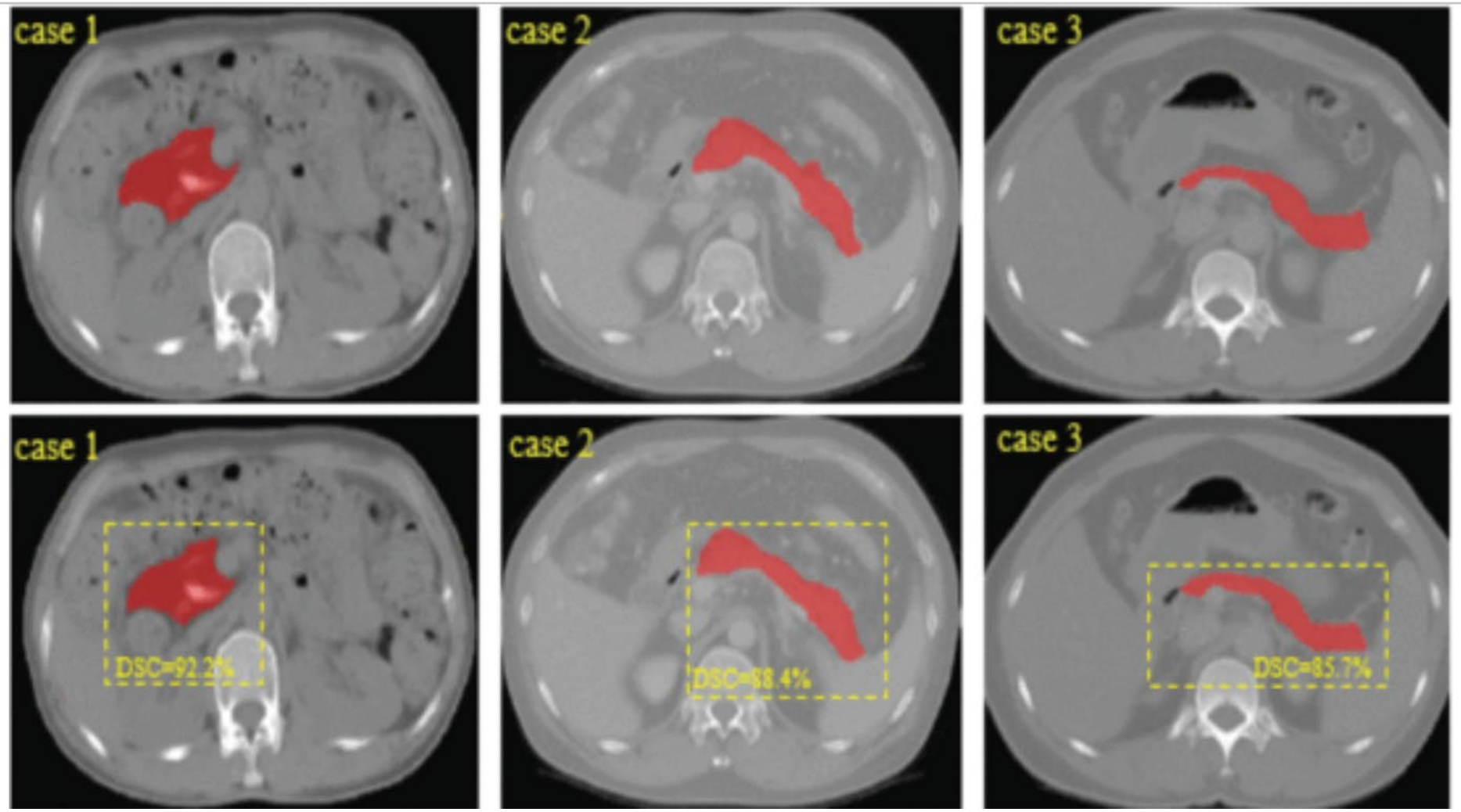
- **Pancreatic segmentation** in particular is both a time and labor-intensive process. Automation using artificial intelligence has proven challenging due to the high degree of variability in the size, shape, and location of the pancreas as well as similarity in voxel intensity to adjacent structures within the abdomen.

Laino ME, Ammirabile A, Lofino L, et al. Artificial Intelligence applied to Pancreatic Imaging: A narrative review. *Healthcare*. 2022;10(8):1511.

Computer Assisted Diagnosis in Pancreatic Imaging

- In 2021, **Xue et al**. used a cascaded multitask 3D fully convolutional network to **attempt to segment the pancreas**.
- This technique used a two-stage approach: an initial fast localization of the pancreas followed by a localized fine-grain segmentation
- This procedure proved to be **more accurate than current state-of-the-art methods of segmentation** with a Dice score of 86.9%

Xue J, He K, Nie D, et al. Cascaded multitask 3-D fully convolutional networks for pancreas segmentation. *IEEE Transactions on Cybernetics*. 2021;51(4):2153–65.



This image shows the pancreatic segmentation results of three different patients within the dataset. The top row contains the expert segmentations, and the bottom row contains the corresponding automatic segmentations.

Automatic Pancreatic Segmentation

- **Zheng et al.** (2020) also utilized a **two-stage process for pancreatic segmentation** involving localization followed by refinement with their 2.5D convolutional neural network.
- This process also achieved what is considered **state-of-the-art performance in pancreatic segmentation** for the used datasets, achieving a Dice score of 86.21%. Sensitivity and specificity were 87.49% and 85.11% respectively.

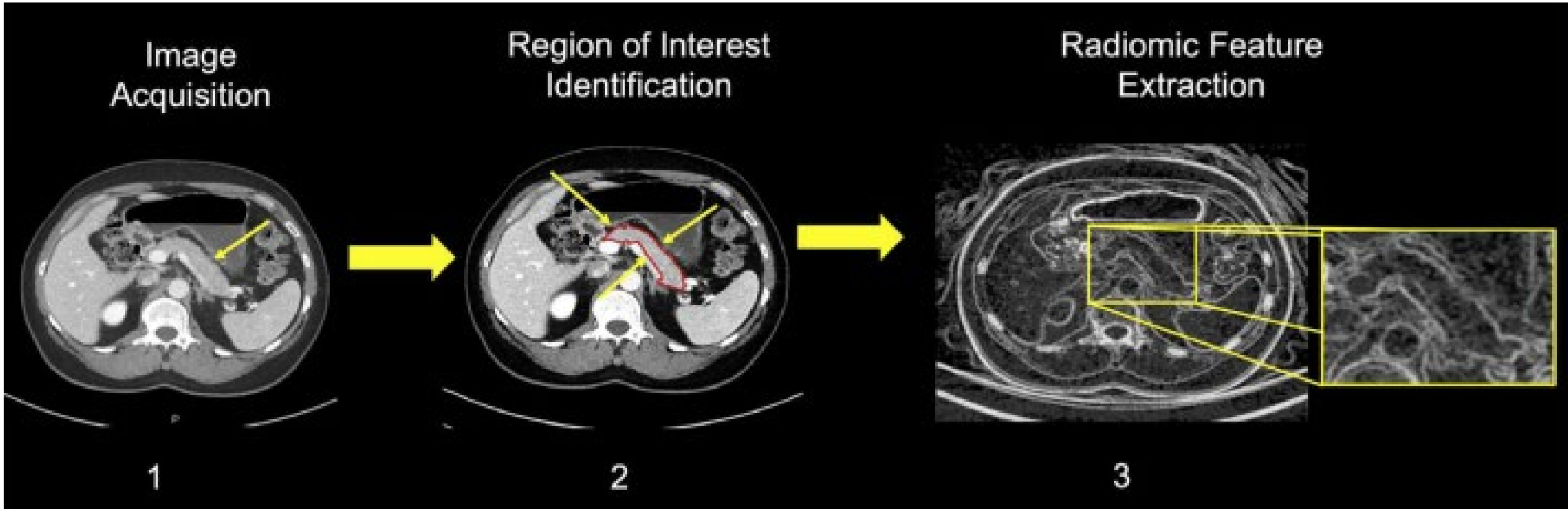
Zheng H, Qian L, Qin Y, et al. Improving the slice interaction of 2.5D CNN for Automatic pancreas segmentation. *Medical Physics*. 2020;47(11):5543–54.

SMV Guided Diagnosis Determination

- In 2020**, Mashayekhi et al.** conducted a retrospective single center cohort study to determine **whether radiomic features of the pancreas on CT could differentiate patients with recurrent acute pancreatitis, functional abdominal pain, and chronic pancreatitis.**
- 56 patients were included in the study. Using 54 radiomic features, a one-vs-one Isomap and Support Vector Machine was trained to classify patients into diagnostic groups.
- On analysis, the **IsoSVM classifier had an overall accuracy of 82.1% in predicting diagnostic groups using 11 radiomic features.**
- The **PPV** of **functional abdominal pain was 100%.**
- **Recurrent acute pancreatitis had the lowest rate of misclassification.**

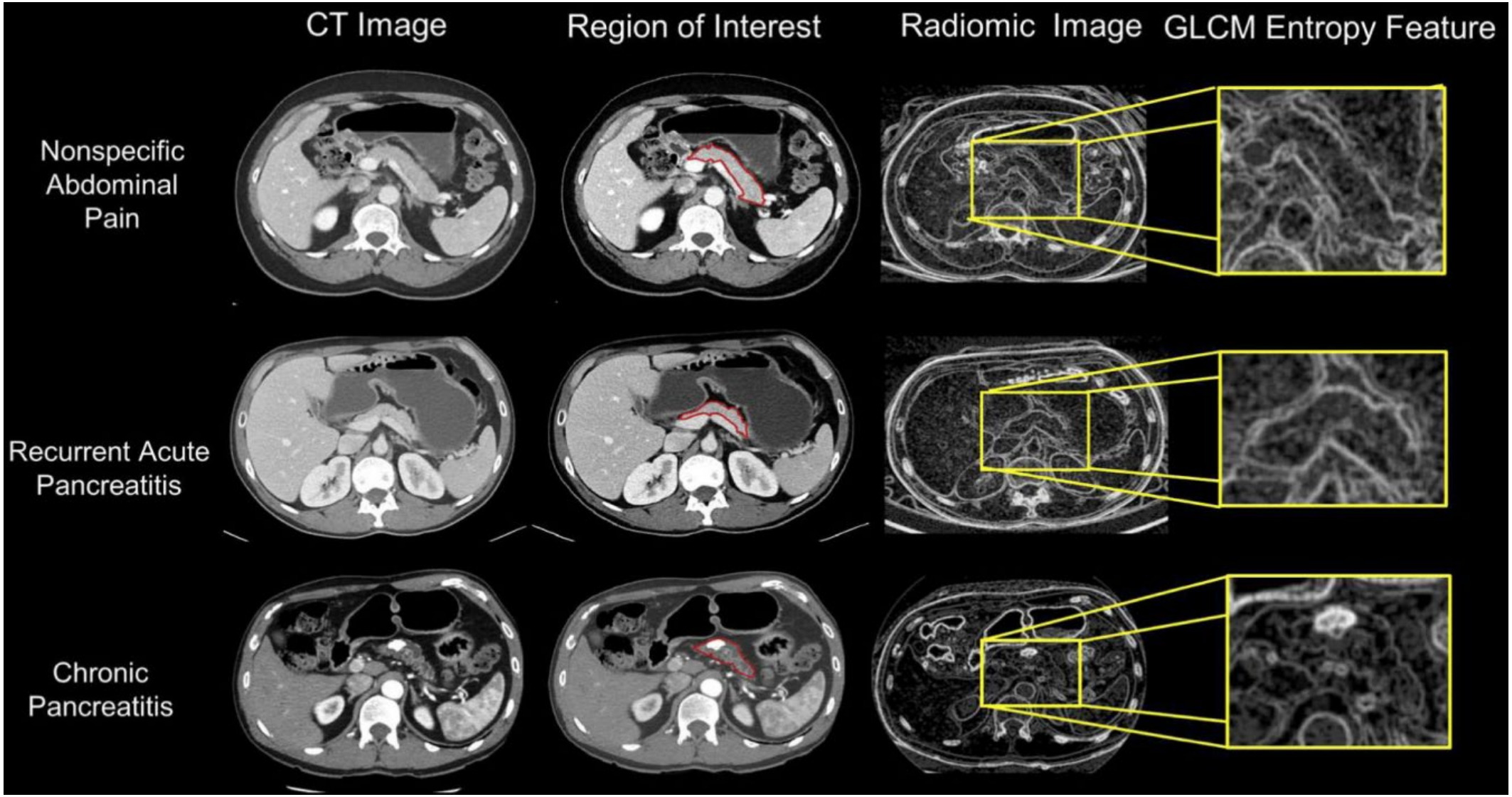
Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. *European Journal of Radiology*. 2020;123:108778.

Generating Radiomic Features



Visual of the process of generating radiomic features

Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. *European Journal of Radiology*. 2020;123:108778.



Most important radiomic features for textural analysis and prediction were GLCM.

- **GLCM (Grey Level Co-Occurrence Matrix)** describes the texture of a region in terms of how heterogenous the values of adjacent voxels are. In other words, a measure of **entropy**.

Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. *European Journal of Radiology*. 2020;123:108778.

IsoSVM Diagnoses using Radiomic Features

- Recurrent acute pancreatitis
  - Sensitivity: 95%
  - Specificity: 78%
  - AUC: 0.88
- Functional abdominal pain
  - Sensitivity: 79%
  - Specificity: 100%
  - AUC: 0.91
- Chronic pancreatitis
  - Sensitivity: 71%
  - Specificity: 95%
  - AUC: 0.90

Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. *European Journal of Radiology*. 2020;123:108778.

CNN Guided Diagnosis Determination

- In 2020**, Marya et al**. conducted a retrospective single center cohort study to see if CNNs could be used to **help distinguish autoimmune pancreatitis from pancreatic ductal adenocarcinoma (PDAC), chronic pancreatitis (CP), or normal pancreatitis (NP).**
- 583 patients with either AIP, PDAC, CP, or NP were included.
- A convolutional neural network model was trained with endoscopic ultrasound images of the pancreas and then tested against the diagnoses of expert human endosonographers.
- **Occlusion heatmap analysis** was also utilized to demonstrate the most discriminating sonographic features of AIP and PDAC specifically.

Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

Results

- The four-way CNN diagnosis classifier **had higher diagnostic accuracy than the human endosonographers** (75.6% vs 61.6%, p=0.026)
- Overall CNN vs. Human EUS Interpretation
  - *Sensitivity=88.2% vs. 53.8%*
  - Specificity=82.5% vs. 86.7%

Table 3 Performance characteristics of human endosonographers versus the CNN model for the identification of AIP

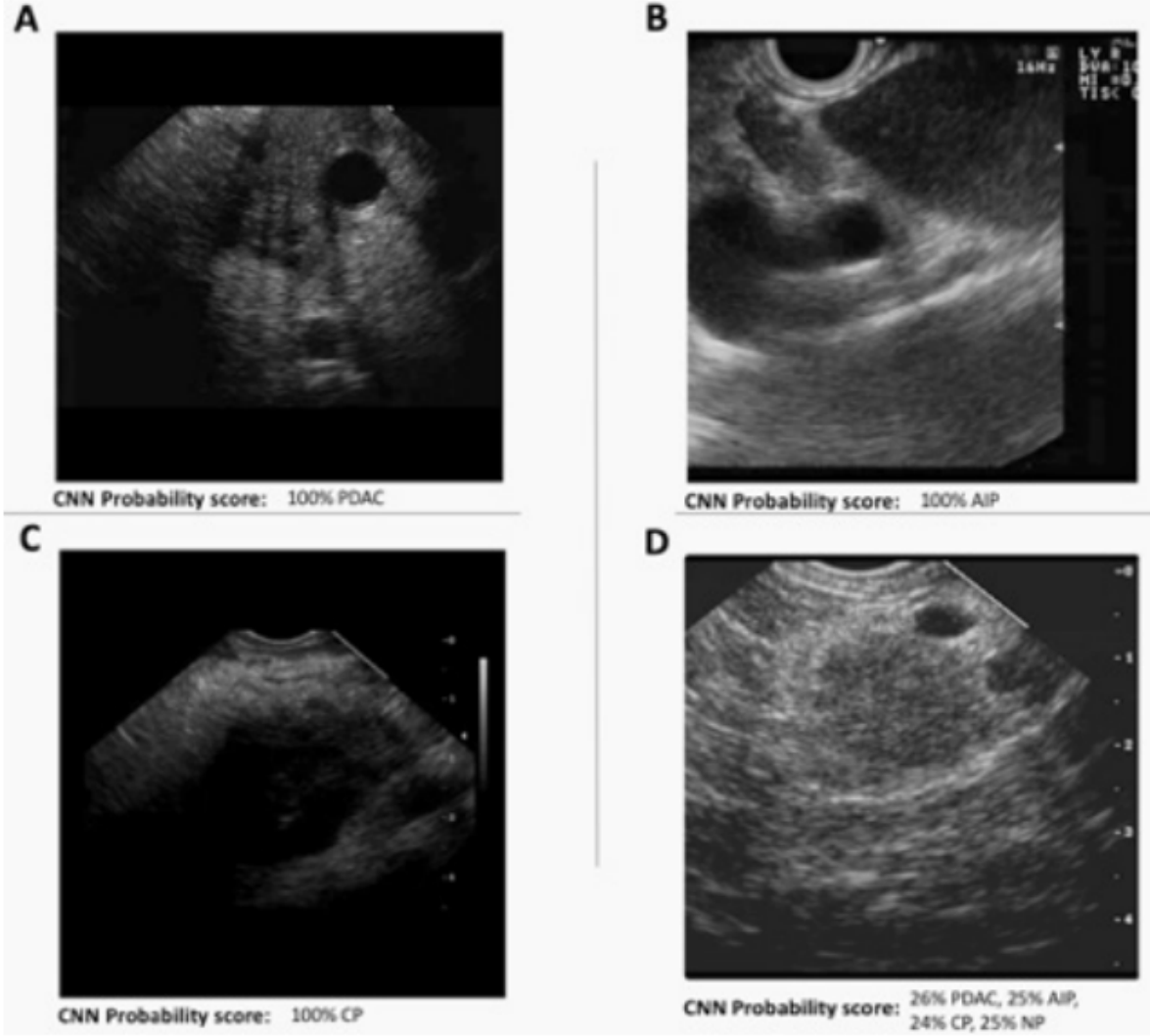
	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	PPV (95% CI)	NPV (95% CI)
Detection of aip from all other conditions					
Human Endosonographers	0.54 (0.44 to 0.63)	0.87 (0.83 to 0.90)	4.05 (3.00 to 5.47)	0.55 (0.47 to 0.62)	0.86 (0.84 to 0.88)
EUS CNN model	0.88 (0.64 to 0.99)	0.82 (0.70 to 0.91)	5.03 (2.79 to 9.06)	0.60 (0.45 to 0.73)	0.96 (0.86 to 0.99)
Detection of aip from pdac alone					
Human Endosonographers	0.54 (0.44 to 0.63)	0.82 (0.77 to 0.87)	3.05 (2.21 to 4.20)	0.60 (0.53 to 0.68)	0.78 (0.74 to 0.81)
EUS CNN model	0.88 (0.64 to 0.99)	0.88 (0.73 to 0.97)	7.50 (2.94 to 19.14)	0.79 (0.60 to 0.91)	0.94 (0.80 to 0.98)

AIP, autoimmune pancreatitis; CNN, convolutional neural network; EUS, endoscopic ultrasound; LR+, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

Representative Images by CNN Algorithms

- Based on CNN probability scores:
  - A: most certain PDAC diagnosis
  - B: most certain AIP diagnosis
  - C: most certain CP diagnosis
  - D: overall most “confusing” image, with lowest overall probability of each diagnosis.



Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

Occlusion Heatmap Analysis

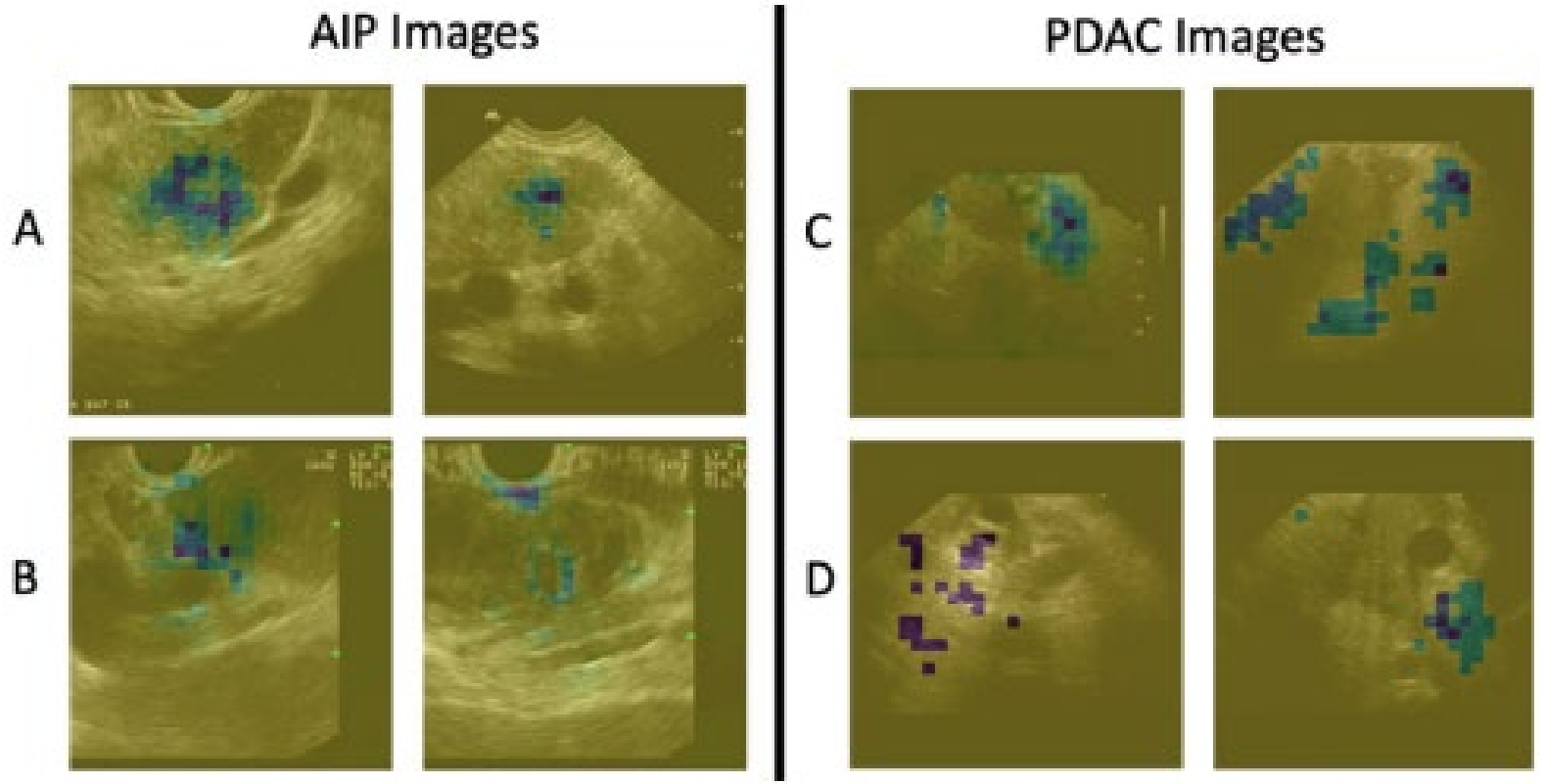
- Using the highest predicted diagnosis probabilities of the CNN model, representative images of AIP and PDAC were extracted.
- Each of the images were assigned a heatmap color based on whether features in that area positively or negatively contributed to disease class probability.

  - **Cold color** regions (blue, purple) localized features which positively contributed to disease class probability
  - **Warm color** regions (orange, yellow) localized features which negatively contributed to disease class probability

Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

Occlusion heatmap results from CNN model

- AIP high value regions:
  - Pancreatic parenchyma
  - hyperechoic plane between pancreas and nearby blood vessels
- PDAC high value regions:
  - Retroperitoneal space
  - Posterior acoustic enhancement deep to the pancreatic vessels and dilated pancreatic duct

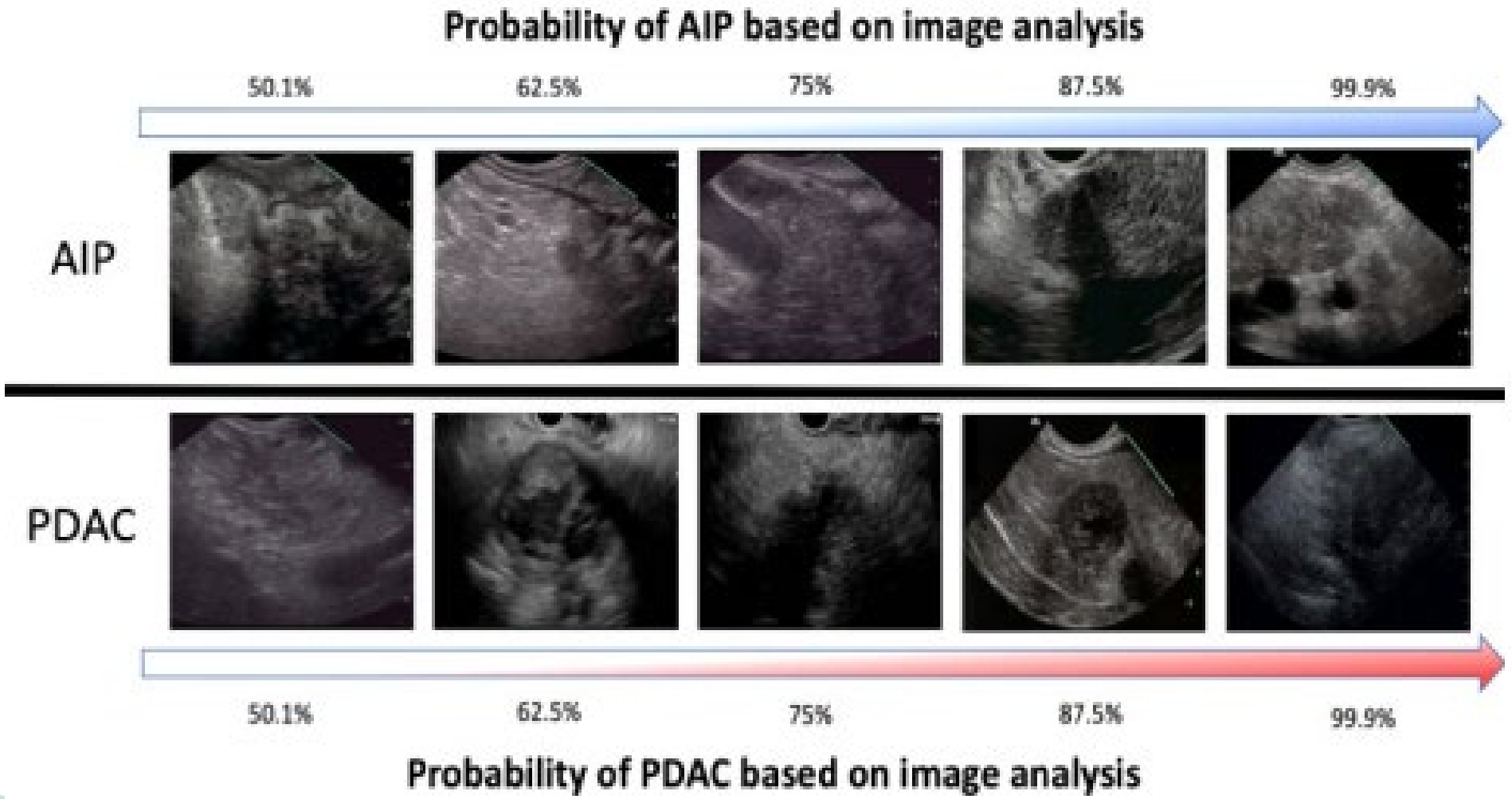


Cold color regions (blue, purple) localized features which positively contributed to disease class probability

Warm color regions (orange, yellow) localized features which negatively contributed to disease class probability

Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

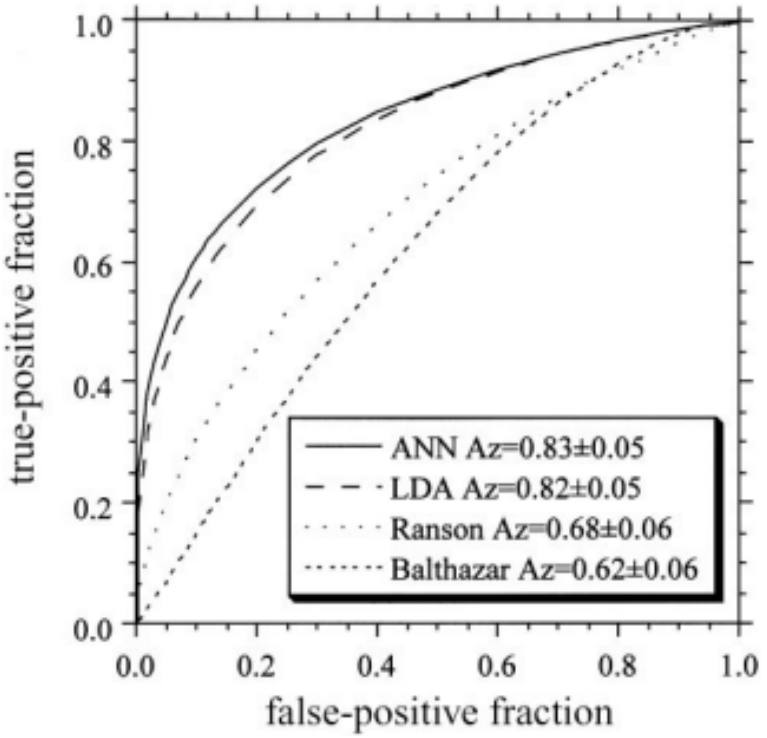
CNN Generated Probability Gradient



Marya NB, Powers PD, Chari ST, et al. Utilisation of artificial intelligence for the development of an eus-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut*. 2020;70(7):1335–44.

ANN performed better than Ranson and Balthazar for severity of AP

- With *Length of Stay* as proxy for severity of acute pancreatitis, ANN was compared against linear discrimination analysis, Ranson score, and Balthazar score
  - **ANN algorithms were proven to be the best for pancreatitis severity prediction based on the area under the ROC curve.**
  - However, there was no statistical difference between the performance of the ANN algorithm and the simpler linear discriminatory analysis.



ROC curve for ANN, linear discriminate analysis (LDA), the Ranson Score, and the Balthazar score.

Keogan MT, Lo JY, Freed KS, et al. Outcome analysis of patients with acute pancreatitis by using an artificial neural network. *Academic Radiology*. 2002;9(4):410–9.

Current Limitations

- Radiomic comparisons across different scanners and imaging centers are difficult due to greater imaging variation.
  - MRI and ultrasound, as more complex modalities, are more difficult than CT to develop machine learning tools with
- For radiomic analysis, the large initial effort needed to appropriately demarcate and segment the regions of interest prior to incorporation into machine learning algorithms can be time consuming and tedious.

- Studies were trained and tested on the same data-sets, thus possibly decreasing external validity of results.
- Current studies also utilize small sample sets derived from one medical center. Number of multi-center studies are small but growing.
- Ongoing data maintenance of algorithms can also be time consuming and cumbersome.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

Future Directions

- Incorporation of more multi-center studies with larger sample sizes.
- Testing of machine learning algorithms on external data sets to test external validity.
- For radiomic analysis, optimizing A.I. which also assists in the identification and segmentation of the regions of interest (i.e pancreas) with a high degree of accuracy and in less time.
- Increase accessibility and knowledge of clinicians to validated A.I. tools for use as clinical decision support tools.

Gorris M, Hoogenboom SA, Wallace MB, et al. Artificial Intelligence for the management of pancreatic diseases. *Digestive Endoscopy*. 2020;33(2):231–41.

Zhang Z, Sejdić E. Radiological images and Machine Learning: Trends, Perspectives, and prospects. *Computers in Biology and Medicine*. 2019;108:354–70.

SAR·DFP

Society *of* Abdominal Radiology
Disease-Focused Panels