AI-Powered Fully automated Early Detection of Pancreatic Ductal Adenocarcinoma on Standard-of-care CT Scans

G. Suman[1], P. Korfiatis[1], A. Patra[1], H. Khasawneh[1], N. Rajamohan[1], C. Cook[1], J.G. Fletcher[1], S. Majumder[1], S.T. Chari[2], M. Truty[1], G. Petersen[1], A.H. Goenka[1]

[1]Mayo Clinic, Rochester, MN, USA

[2]MD Anderson Cancer Center, Houston, TX, USA

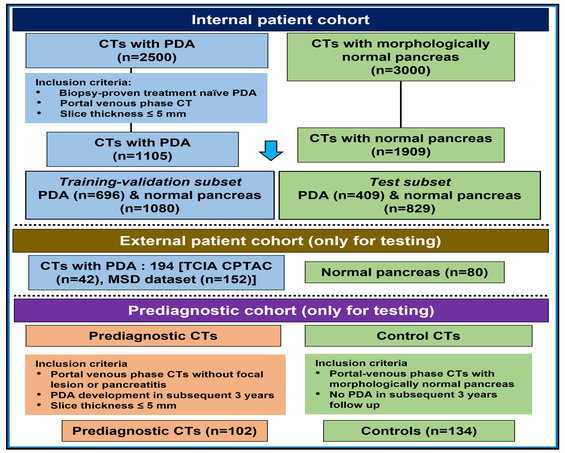
BACKGROUND

Up to 30% of tumors smaller than 2 cm can go undetected on CT scans. Unfortunately, pancreas cancer (PDA) can advance from being undetectable to widely metastatic in a matter of months which underscores the critical need to develop novel tools such as AI to improve its detection. Early detection of these small tumors is the key to improved outcomes. **In this study, we aimed to develop a fully automated convolutional neural network (CNN) for early detection of PDA on standard-of-care CT.**

METHODS

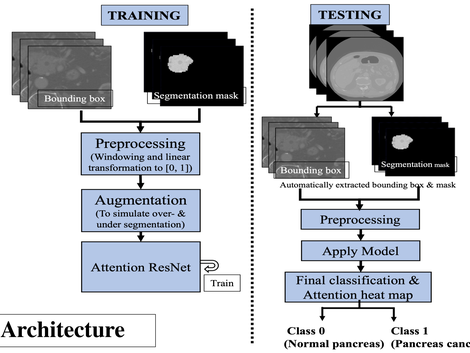
Study cohort:

- CT abdomen dataset with portal-venous phase & slice thickness (ST) ≤ 5 mm was curated. CTs with biliary stents & suboptimal quality were excluded. A 3D-CNN was trained on of 696 diagnostic CTs containing PDA & 1080 control CTs with morphologically normal pancreas.
- Model was evaluated on (a) **test subset** of 409 PDA & 829 control CTs; (b) **simulated cohort** with high pre-test probability of PDA (1-5% over 3-years), which matched the case-control distribution in high-risk groups (glycemically-defined new-onset diabetes & high ENDPAC score); (c) **multi-institutional public dataset** of 194 PDA & 80 control CTs; and (d) **Prediagnostic cohort** of 102 CTs (i.e. CTs with normal-appearing pancreas incidentally acquired 3–36 months before clinical diagnosis of PDA) & 134 control CTs.



CNN Architecture:

- CNN utilized two inputs: bounding box containing the pancreas (including the tumor in CTs with PDA) as well as peripancreatic tissue; & bounding box with pancreas only.
- Pancreatic mask (including the tumor in the CTs with PDA) was generated using our previously published fully automated modified CNN for volumetric pancreas segmentation[1].
- Image augmentation to simulate segmentation errors, inter-reader variability, & to compensate for the absence of radiologists' segmentations as input.
- Final prediction output generated was either class '0' (normal pancreas), or class '1' (PDA) for each patient & Heat map analysis was done to check tumor localization.



CNN Architecture

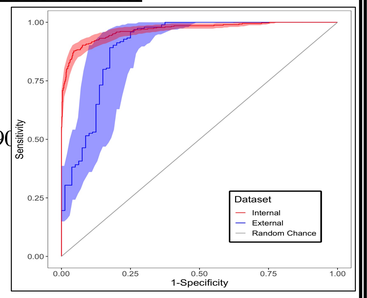
RESULTS

Internal test subset:

- Mean (SD) tumor diameter: 3·4 (1·3) cm.
- Majority of the PDA lesions were T2 (n=291, 71%) followed by T3 (n=60, 15%), T1 (n=35, 9%), & T4 (n=23, 6%).
- Model correctly classified 360 (88%) PDA CTs & 783 (94%) normal CTs (Table).
- The model correctly classified 360 of the 409 (88%) CTs with PDA & 783 of the 829 (94%) control CTs.
- Overall, 1143 of 1238 CTs in the test subset were correctly classified yielding an accuracy (CI) of 0·92 (0·91-0·94), AUROC of 0·97 (0·96-0·98), sensitivity of 0·88 (0·85-0·91), specificity of 0·95 (0·93-0·96).

Internal test set vs. external public dataset:

- 236 of 274 CTs (170/194 PDA CTs and 66/80 controls) from external dataset were correctly classified yielding accuracy of 0·86 (0·82-0·90), AUROC of 0·90 (0·86-0·95), sensitivity of 0·88 (0·83-0·92), specificity of 0·83 (0·74-0·90).



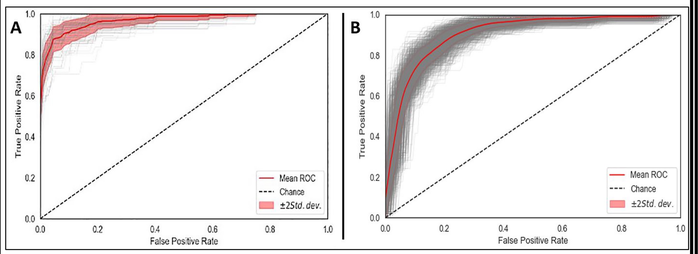
	Internal test subset (409 CTs with PDA & 829 control CTs)		Metric	Public dataset (194 CTs with PDA & 80 control CTs)		Metric
Tumor stage	Total CTs with PDA (n=409)	Misclassified CTs with PDA (n=49)	Sensitivity (95% CI)	Total CTs with PDA (n=194)	Misclassified CTs with PDA (n=24)	Sensitivity (95% CI)
1	35 (9%)	7 (14%)	0.80 (0.66-0.91)	25 (13%)	6 (25%)	0.76 (0.60-0.92)
2	291 (71%)	39 (80%)	0.87 (0.82-0.90)	113 (58%)	15 (63%)	0.87 (0.81-0.93)
3	60 (15%)	3 (<0.1%)	0.95 (0.88-1)	30 (15%)	2 (8%)	0.93 (0.83-1.00)
4	23 (6%)	0	1.00 (1.0-1.0)	26 (13%)	1 (4%)	0.96 (0.88-1.00)

Simulated cohort with high pre-test probability of PDA:

Accuracy (CI) of 0·95 (0·94-0·95) & AUROC 0·97 (0·94-0·99).

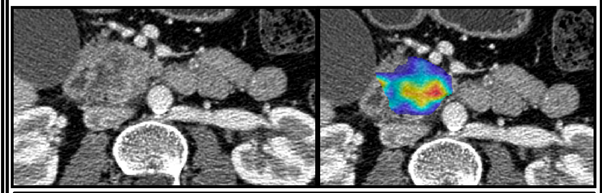
Prediagnostic cohort:

- **Median interval of 475 days** (range:93-1082 days) **before diagnosis.**
- 76 of 102 prediagnostic CTs were correctly classified. Overall accuracy of 0·84 (0·79-0·88), AUROC of 0·90 (0·87-0·94), sensitivity of 0·75 (0·66-0·83), specificity of 0·90 (0·85-0·95).



AUROC of 0·97 in simulated cohort with high pre-test probability of PDA, & 0·90 in prediagnostic cohort (A & B in figure, respectively).

Heat map analysis: In 350 of the 360 correctly classified PDA CTs in internal test set, & 65 of the 76 correctly classified prediagnostic CTs, the area of activation in the heat map overlapped with the tumor location or predicted site of tumor.



Axial CT image shows a 4. 3 cm hypodense pancreas cancer (stage T3) in pancreatic head. Heat activation map shows that model's classification of the CT as a tumor-containing CT was based on accurate localization of the tumor by the model.

Failure Mode Analysis:

- Most of the misclassified CTs in internal test set had T2 tumors (n=39/49, 80%).
- Mean tumor size of misclassified PDA CTs in internal test set (2.8 cm) was smaller than the correctly classified CTs (3.5 cm) (p=0·001).
- No difference in performance between different T stages, tumor location, age, sex, CT ST, vendor (p>0.05) in diagnostic or prediagnostic CTs.

CONCLUSION

- A 3D-CNN developed on a large dataset shows high accuracy and generalizability for fully-automated detection of PDA (including small & isodense tumors) on diagnostic CTs as well as for early detection of PDA from normal pancreas on prediagnostic CTs at a substantial lead time prior to clinical diagnosis.

REFERENCES

1. Panda, A., et al., Two-stage deep learning model for fully automated pancreas segmentation on computed tomography: Comparison with intra-reader and inter-reader reliability at full and reduced radiation dose on an external dataset. Med Phys, 2021. 48(5): p. 2468-2481.