

Appendix A Supplementary Materials

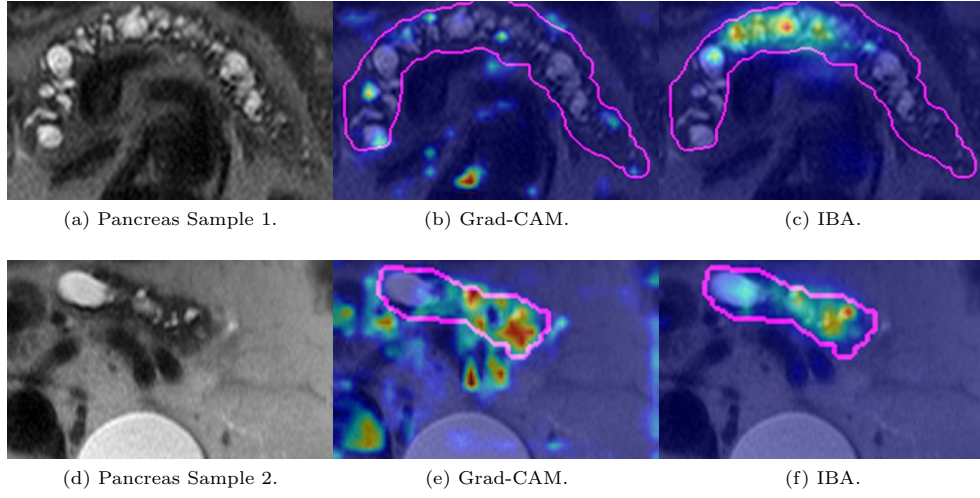


Fig. A1: Visual explanations using Grad-CAM and IBA.

Table A1: Data Distribution for IPMN Classification.

Data Centers	T1W Modality				T2W Modality			
	No Risk	Low Risk	High Risk	Total	No Risk	Low Risk	High Risk	Total
Center 1: NYU	48	79	23	150	48	79	24	151
Center 2: MCF	29	42	63	134	25	42	63	130
Center 3: NU	43	126	17	186	44	127	16	187
Center 4: AHN	1	11	4	16	1	13	4	18
Center 5: MCA	0	10	14	24	0	7	16	23
Center 6: IU	3	48	13	64	3	46	14	63
Center 7: EMC	40	23	15	78	38	30	15	83
Total	164	339	149	652	159	344	152	655

Table A2: PanSegNet’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	84.99±7.74	74.56±9.84	84.80±8.66	86.10±8.32	5.96±5.31	1.26±1.19
MCF	87.47±5.79	78.17±8.37	87.63±7.44	88.11±7.37	5.35±5.68	1.07±0.94
NU	88.33±5.89	79.52±7.98	87.49±7.69	89.72±5.74	4.53±6.86	1.15±0.89
AHN	81.22±6.11	68.79±8.16	84.54±8.92	79.15±7.85	8.27±3.75	1.61±0.78
MCA	78.72±10.98	66.03±12.55	75.29±14.90	85.04±6.52	8.37±10.79	2.04±2.21
IU	85.37±9.91	75.56±12.65	86.13±9.51	85.54±11.58	8.65±16.96	1.89±2.72
EMC	89.96±4.06	81.98±6.16	89.64±5.01	90.53±4.89	5.67±21.28	1.33±2.30
Global	86.81±7.30	77.32±9.74	86.56±8.59	87.84±7.80	5.81±10.90	1.30±1.55
T2W Modality						
NYU	87.17±4.39	77.52±6.62	87.98±6.22	86.88±6.16	5.03±3.17	0.93±0.49
MCF	87.85±5.58	78.72±7.86	89.13±6.01	87.34±8.02	4.94±6.61	0.96±1.13
NU	92.50±5.15	86.41±7.81	92.43±5.85	92.91±5.94	2.94±3.25	0.52±0.63
AHN	82.72±9.25	71.46±11.73	86.60±11.08	80.15±9.62	7.16±6.41	1.33±1.45
MCA	84.84±3.27	73.82±4.92	87.38±5.91	83.12±6.51	6.18±3.60	0.99±0.46
IU	92.46±3.73	86.19±6.18	94.16±4.17	91.02±5.06	2.46±1.97	0.40±0.28
EMC	91.05±8.64	84.46±11.30	93.38±4.78	89.74±11.02	4.39±7.38	0.70±1.18
Global	89.62±6.38	81.73±9.31	90.74±6.45	89.11±8.06	4.19±4.99	0.75±0.87

Table A3: Swin-UNETR’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	77.78±2.40	65.26±2.51	77.48±3.78	80.72±1.68	23.76±11.55	6.93±5.58
MCF	81.63±2.92	70.17±3.51	82.97±2.06	82.05±3.48	11.27±2.84	2.68±1.71
NU	82.89±2.06	72.00±2.89	82.21±2.55	84.90±1.77	13.77±4.98	3.34±0.96
AHN	65.14±11.05	50.90±10.48	74.43±6.51	61.04±12.47	22.51±15.68	4.23±2.94
MCA	67.50±6.05	52.45±6.56	66.26±4.72	70.99±7.94	63.22±28.78	13.67±6.59
IU	79.17±2.31	67.09±3.06	78.74±3.73	81.87±3.43	30.53±6.75	8.03±2.95
EMC	74.30±2.49	61.89±3.31	73.22±5.13	79.46±2.72	72.44±48.10	24.83±19.55
Global	79.09±1.40	67.19±1.63	79.09±1.67	81.37±0.74	26.55±8.48	7.58±3.76
T2W Modality						
NYU	78.81±1.67	66.10±2.09	80.73±2.84	78.61±2.42	13.21±2.28	2.72±0.89
MCF	79.43±2.06	67.22±2.50	83.70±2.83	77.72±1.45	13.06±2.93	1.75±0.26
NU	81.18±1.02	69.80±1.04	81.48±0.04	82.19±1.95	17.38±3.96	4.20±1.52
AHN	64.21±6.26	50.03±6.59	72.19±8.17	64.18±9.03	31.07±13.51	8.69±5.93
MCA	59.63±4.03	44.37±3.81	63.87±5.44	61.36±9.01	81.04±22.10	24.77±8.46
IU	78.47±1.47	65.74±2.06	81.70±2.05	76.57±3.74	23.25±11.22	3.77±1.91
EMC	63.39±5.87	49.69±5.98	65.73±8.97	69.20±5.54	87.79±40.81	28.59±14.81
Global	76.29±0.66	63.77±0.82	78.70±2.14	76.86±1.21	28.53±7.01	7.57±2.60

Table A4: Swin-UNETR+FedAvg’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	71.84±2.51	58.25±2.21	71.06±3.45	76.67±1.26	34.81±8.00	12.12±6.18
MCF	77.94±3.36	65.40±3.85	79.16±3.10	79.07±3.57	15.64±3.96	3.60±1.74
NU	77.61±2.78	65.25±3.38	74.06±2.82	83.66±1.95	31.89±5.17	8.36±1.83
AHN	51.07±10.96	37.61±8.85	61.60±4.80	48.41±13.08	34.78±12.52	12.68±8.57
MCA	46.13±11.55	32.28±10.35	40.38±10.29	59.55±12.69	110.16±36.15	44.79±18.87
IU	71.09±3.26	58.03±4.01	67.60±3.86	80.11±4.29	53.95±6.37	15.16±3.77
EMC	55.67±8.40	42.23±8.18	53.61±11.02	67.57±0.69	106.46±50.34	43.83±26.68
Global	71.26±2.59	58.30±2.74	69.74±3.09	77.13±1.34	43.65±8.60	14.77±5.01
T2W Modality						
NYU	74.35±1.48	60.40±1.65	76.41±3.60	74.20±2.62	23.85±7.88	5.02±2.26
MCF	74.69±2.19	61.36±2.24	80.06±3.04	72.72±1.71	15.10±2.44	2.64±0.62
NU	73.99±2.16	60.71±2.15	73.26±2.12	76.54±2.20	27.46±2.41	6.97±1.70
AHN	59.41±4.71	45.24±3.78	64.31±3.14	60.20±6.89	43.94±7.98	15.80±5.17
MCA	43.20±4.86	29.99±3.94	40.48±4.54	55.05±8.71	129.24±14.32	46.21±10.93
IU	65.74±3.60	51.15±3.71	66.98±4.91	67.52±3.44	59.26±14.00	13.29±4.07
EMC	54.50±7.95	40.46±7.54	53.28±10.43	63.44±5.58	104.14±30.44	34.60±14.15
Global	69.19±1.51	55.48±1.44	70.52±2.98	71.31±0.95	41.87±5.84	11.74±3.02

Table A5: Swin-UNETR+FedProx($\mu = 0.3$)’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	60.31±2.27	46.02±2.12	59.92±2.67	66.25±0.97	81.99±9.69	26.72±5.39
MCF	71.63±4.18	57.63±4.54	73.65±4.55	72.56±3.43	30.03±11.76	6.65±3.25
NU	65.07±3.77	51.06±3.74	62.81±4.73	70.57±2.54	91.69±8.79	24.60±3.51
AHN	36.24±10.71	24.91±8.01	37.94±9.71	38.25±13.66	101.79±4.39	35.16±10.74
MCA	26.91±15.07	18.28±11.69	23.06±13.27	41.17±16.93	141.22±30.07	74.43±24.62
IU	49.87±5.22	36.51±4.64	43.31±5.09	64.78±6.79	123.78±8.26	47.66±5.68
EMC	28.87±14.46	20.21±10.66	34.43±19.59	33.06±13.66	132.43±43.78	74.32±36.88
Global	57.24±3.41	44.16±2.85	56.85±4.33	62.93±2.80	87.27±8.34	31.99±6.19
T2W Modality						
NYU	66.32±1.57	51.23±1.45	68.43±3.06	66.90±3.30	41.34±6.78	9.66±2.52
MCF	66.00±3.16	51.19±3.06	73.18±4.34	63.36±3.00	31.08±4.63	6.04±1.30
NU	62.60±3.64	47.76±3.39	60.81±4.80	67.36±2.27	57.78±11.94	15.84±4.55
AHN	52.45±8.06	38.52±7.26	55.24±7.96	55.00±7.23	54.94±14.03	18.18±8.25
MCA	30.01±5.61	18.73±4.48	25.81±5.46	44.53±9.74	162.60±17.22	64.85±10.62
IU	52.23±2.97	38.08±3.03	50.30±3.28	58.34±3.88	90.75±7.97	27.16±3.23
EMC	37.09±11.12	25.26±8.83	35.36±12.71	46.12±10.80	138.44±27.07	55.54±20.05
Global	58.11±2.61	43.85±2.22	59.01±4.38	61.46±1.11	66.68±9.10	20.87±4.81

Table A6: Swin-UNETR+FedProx($\mu = 0.01$)’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	70.16±2.67	56.32±2.50	68.93±2.96	75.42±1.94	48.86±6.08	15.39±5.71
MCF	77.56±3.35	64.78±3.86	78.55±2.42	78.98±4.17	16.27±5.43	3.90±1.93
NU	76.02±3.38	63.70±4.09	72.70±4.04	82.47±2.04	38.92±12.04	10.19±2.66
AHN	47.10±10.95	33.87±8.76	57.15±7.38	44.02±13.92	48.54±16.35	17.19±9.42
MCA	38.75±13.79	26.78±11.58	33.39±12.75	52.80±14.80	118.66±40.33	58.79±22.81
IU	67.67±3.76	54.48±4.85	64.12±4.80	77.69±4.72	70.14±9.79	21.67±4.17
EMC	50.07±12.37	37.68±11.39	47.84±14.68	61.21±7.92	119.22±48.27	55.12±31.83
Global	68.95±3.30	56.09±3.37	67.33±3.37	75.11±2.68	52.82±9.97	18.76±5.41
T2W Modality						
NYU	73.58±1.54	59.53±1.75	75.26±3.12	74.01±2.91	24.61±7.36	6.05±2.33
MCF	73.70±2.45	60.23±2.54	78.80±2.98	71.85±2.17	18.39±4.02	3.23±0.79
NU	72.90±2.49	59.35±2.47	71.49±1.54	76.13±3.82	31.58±4.63	8.19±1.97
AHN	57.18±7.15	43.33±5.93	65.22±8.08	59.25±10.56	53.58±13.61	16.15±7.26
MCA	39.65±4.32	26.78±3.34	35.97±4.44	52.74±7.16	143.75±17.54	53.97±13.63
IU	62.72±2.67	47.93±2.82	63.94±3.16	65.58±3.93	70.23±10.51	16.80±3.20
EMC	48.97±7.65	35.23±6.79	46.97±11.17	60.22±3.52	122.75±28.92	45.19±15.92
Global	67.25±1.86	53.47±1.77	68.22±2.99	70.23±1.21	48.36±6.51	14.52±3.64

Table A7: Swin-UNETR+FedProx($\mu = 0.005$)’s performance

Center	Dice(%)	Jaccard(%)	Precision(%)	Recall(%)	HD95(mm)	ASSD(mm)
T1W Modality						
NYU	70.36±2.53	56.56±2.52	69.05±3.13	75.88±1.13	45.50±9.37	14.81±5.80
MCF	78.06±3.01	65.45±3.43	78.71±2.11	79.71±3.99	16.84±4.69	3.89±1.76
NU	77.12±3.06	64.79±3.77	73.90±3.36	83.19±2.12	34.39±8.18	8.86±2.00
AHN	49.40±10.96	36.07±8.83	58.90±8.21	46.13±12.58	44.40±25.11	14.65±9.00
MCA	41.21±13.62	28.85±11.78	37.23±12.92	54.25±14.72	118.45±41.69	53.98±23.47
IU	68.46±3.17	55.10±3.98	64.54±3.92	77.89±4.16	62.72±7.43	19.55±3.45
EMC	51.19±11.52	38.50±10.66	49.55±13.78	62.32±6.74	116.00±50.94	52.80±31.64
Global	69.77±2.90	56.88±2.98	68.16±2.98	75.82±2.39	49.61±8.55	17.51±5.39
T2W Modality						
NYU	73.95±1.71	59.96±1.95	75.27±2.96	74.70±2.76	23.38±6.26	5.50±1.87
MCF	74.07±2.93	60.69±2.97	79.08±3.05	72.32±2.92	17.86±3.41	3.12±0.79
NU	73.45±2.47	60.05±2.53	71.86±1.80	77.06±3.60	35.49±4.63	8.57±1.95
AHN	56.13±5.71	42.54±4.85	60.49±4.12	57.76±8.25	44.34±11.90	16.72±5.56
MCA	41.39±5.59	28.18±4.58	37.43±5.67	54.20±8.71	138.35±23.02	48.29±13.09
IU	63.61±1.86	49.21±2.25	65.14±2.76	65.75±3.16	69.52±11.55	16.15±2.49
EMC	50.03±8.99	36.43±8.09	48.70±11.38	61.28±4.41	123.65±25.87	43.32±15.18
Global	67.82±1.44	54.16±1.43	68.63±2.41	70.89±0.84	48.64±5.82	13.99±2.96

Table A8: Binary Classification (High Risk or Not) Results for Multi-Center IPMN MRI Using DenseNet-121.

Method	T1W Modality		T2W Modality	
	ACC(%)	AUC(%)	ACC(%)	AUC(%)
Center 1: New York University Langone Health (NYU)				
Centralized	86.65±3.37	77.09±12.25	83.50±7.70	87.54±3.89
FedAvg	81.28±4.45	68.42±13.05	87.43±2.12	90.05±4.79
FedProx ($\mu = 0.1$)	81.95±5.31	68.92±14.56	85.40±3.11	88.98±5.36
FedProx ($\mu = 0.3$)	73.19±15.91	70.19±14.82	86.79±3.14	92.67±3.60
Center 2: Mayo Clinic Florida (MCF)				
Centralized	65.71±8.49	74.80±13.04	72.44±12.66	81.46±10.11
FedAvg	64.95±9.86	74.31±10.05	70.79±4.94	80.85±8.25
FedProx ($\mu = 0.1$)	65.75±11.13	73.66±12.14	67.80±8.71	80.75±9.88
FedProx ($\mu = 0.3$)	65.02±13.48	74.94±13.22	63.87±10.16	81.10±6.57
Center 3: Northwestern University (NU)				
Centralized	84.96±5.82	67.66±10.41	82.93±9.11	57.34±16.99
FedAvg	90.33±2.36	74.66±13.61	90.89±2.43	57.94±17.95
FedProx ($\mu = 0.1$)	88.15±2.52	73.46±10.93	84.93±10.16	60.18±22.42
FedProx ($\mu = 0.3$)	88.16±2.46	73.54±14.27	89.29±2.20	58.24±27.04
Center 4: Allegheny Health Network (AHN)				
Centralized	87.50±12.50	66.67±40.82	67.50±19.20	70.83±29.76
FedAvg	50.00±17.68	33.33±23.57	76.25±17.81	85.42±14.88
FedProx ($\mu = 0.1$)	56.25±32.48	58.33±43.30	70.00±27.61	70.83±18.16
FedProx ($\mu = 0.3$)	56.25±36.98	58.33±36.32	88.75±11.39	79.17±21.65
Center 5: Mayo Clinic Arizona (MCA)				
Centralized	66.67±20.41	73.26±23.88	55.00±20.21	75.00±25.00
FedAvg	75.00±14.43	85.07±15.88	48.33±9.57	75.00±15.31
FedProx ($\mu = 0.1$)	75.00±14.43	88.19±11.85	44.17±11.40	68.75±27.24
FedProx ($\mu = 0.3$)	62.50±13.82	88.19±11.85	39.17±6.82	81.25±10.83
Center 6: Istanbul University Faculty of Medicine (IU)				
Centralized	78.12±6.99	64.50±24.95	79.48±4.87	83.36±11.06
FedAvg	79.69±5.18	63.02±20.20	76.35±9.00	78.79±13.25
FedProx ($\mu = 0.1$)	75.00±7.65	59.46±20.63	79.58±9.09	77.42±13.58
FedProx ($\mu = 0.3$)	70.31±19.45	62.54±16.40	77.81±2.85	79.86±11.47
Center 7: Erasmus Medical Center (EMC)				
Centralized	83.22±7.76	84.45±7.30	74.82±13.06	83.95±11.71
FedAvg	80.79±2.02	66.87±17.54	77.20±11.72	68.75±12.10
FedProx ($\mu = 0.1$)	79.47±3.76	83.31±15.08	77.20±6.84	81.74±16.10
FedProx ($\mu = 0.3$)	78.16±2.62	81.12±15.36	77.14±6.06	78.68±11.32
Global				
Centralized	79.90±1.52	78.13±2.06	78.23±5.09	82.37±1.02
FedAvg	79.29±2.04	76.19±1.02	81.08±1.69	81.24±2.98
FedProx ($\mu = 0.1$)	78.52±1.23	76.46±1.21	78.29±2.87	81.35±1.77
FedProx ($\mu = 0.3$)	75.26±4.90	76.67±1.09	79.27±2.41	81.81±2.15

Table A9: Comparison of IPMN Classification Accuracy: Radiologists vs. DenseNet-121.

Method	Input Modality		Three-Class Classification			Binary Classification	
	T1W	T2W	No-Risk	Low-Risk	High-Risk	No/Low-Risk	High-Risk
Radiologist 1	✓	✓	89.03	58.43	64.08	89.94	64.08
Radiologist 2	✓	✓	93.55	30.12	32.39	97.13	32.39
Radiologist 3	✓	✓	91.61	56.33	41.55	94.66	41.55
Average	✓	✓	91.40	48.29	46.01	93.91	46.01
DenseNet-121	✓	✗	37.42	81.02	33.80	91.38	40.85
DenseNet-121	✗	✓	54.55	79.13	52.52	86.11	53.96
DenseNet-121(feature fusion)	✓	✓	45.16	83.13	42.96	95.69	43.66
DenseNet-121(probability fusion)	✓	✓	45.16	84.94	45.77	94.66	47.18