Supplementary Methods

WSI Preprocessing and Feature Extraction

During WSI preprocessing, tissue regions were first segmented from the background on low-resolution slides. These regions were then tiled into 512×512 pixel patches at 20× magnification. Patches with minimal tissue content were discarded. The remaining patches were resized to 224×224 pixels and underwent standard color normalization to mitigate staining variations across different centers and scanners.

For feature extraction, we selected the H-optimus-1 model, a Vision Transformer (ViT-G/14) foundation model pre-trained on approximately one million WSIs. Its pre-trained encoder was utilized to generate a 1024-dimensional feature vector for each tissue patch, effectively converting each WSI into a bag of feature instances.

DeepPMD Model Architecture

To enhance classification performance, we integrated patch-level features with patient-level clinicopathological data (gender and age) to create multimodal inputs. Our classification framework, DeepPMD, implements an attention-based multi-task aggregator. This aggregator first projects the 1024-dimensional feature embeddings to 384 dimensions through two fully connected layers with ReLU activation. The architecture then incorporates two parallel processing branches, each employing a gated attention mechanism: one branch is dedicated to the binary task of primary versus metastatic cancer classification, and the other is for the six-class primary site prediction. This multi-task attention pooling allows the model to dynamically weight patch features according to their diagnostic relevance for each specific task. The weighted features from each branch are aggregated into a comprehensive slide-level representation, which is then concatenated with the clinical information for the final prediction via a terminal fully connected layer.

Supplementary Criteria

Inclusion Criteria:

(1) Histologically confirmed primary LUAD or metastatic lung cancer (primary site unrestricted) via surgical resection or biopsy;

(2) Patients had not received any prior treatment before surgical resection or biopsy, including radiotherapy, chemotherapy, targeted therapy, or immunotherapy;

(3) Availability of complete clinicopathological data, formalin-fixed and paraffin-embedded (FFPE) H&E stained slides, clinical pathological information including diagnosis, patients’ basic information like gender, age, and so on.

Exclusion Criteria:

(1) Inconsistent Class Representation or Insufficient Sample Size: The pathological diagnosis was a type not represented across all four cohorts.

(2) Poor Image Quality: The WSI failed during the scanning process or was of poor quality unsuitable for diagnosis due to factors such as significant artifacts, blurring, or suboptimal staining.

(3) Minors: Patients under the age of 18.

Supplementary Figures

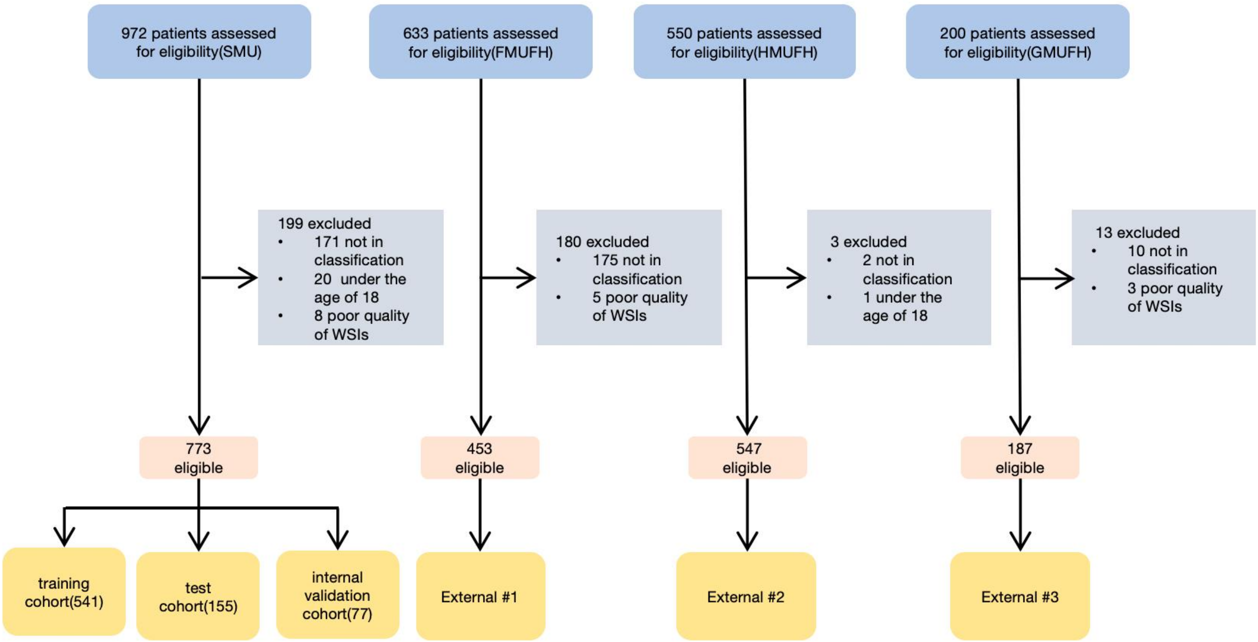


Figure S1: Data Screening flowchart. SMU = Southern Medical University. FMUFH = the First Affiliated Hospital of Shandong First Medical University. HMUFH = the Fourth Hospital of Hebei Medical University. GMUFH = the Sixth Affiliated Hospital of Guangzhou Medical University. WSI = whole-slide image.

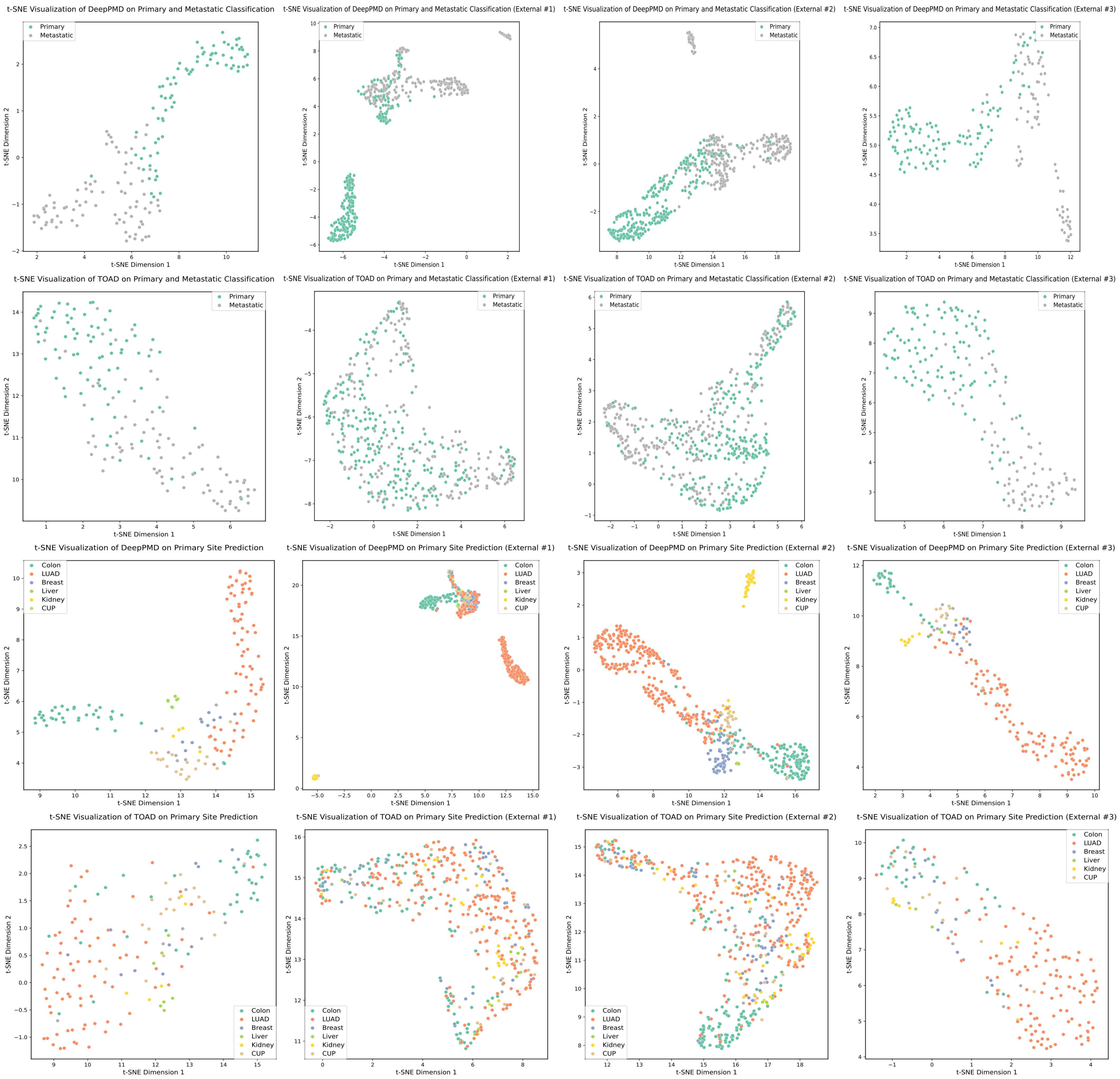


Figure S2: Comprehensive Visualization of Feature Space Across Internal and Multiple External Test Cohorts

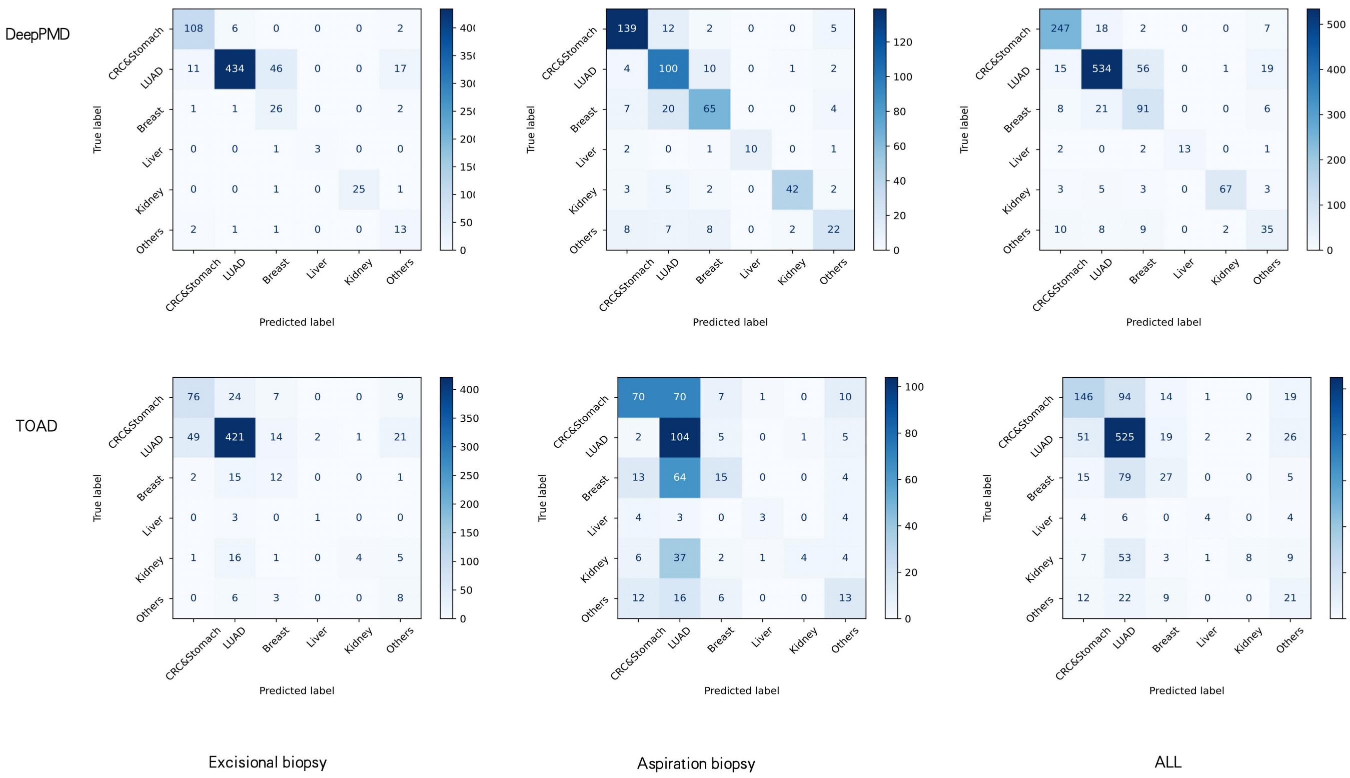
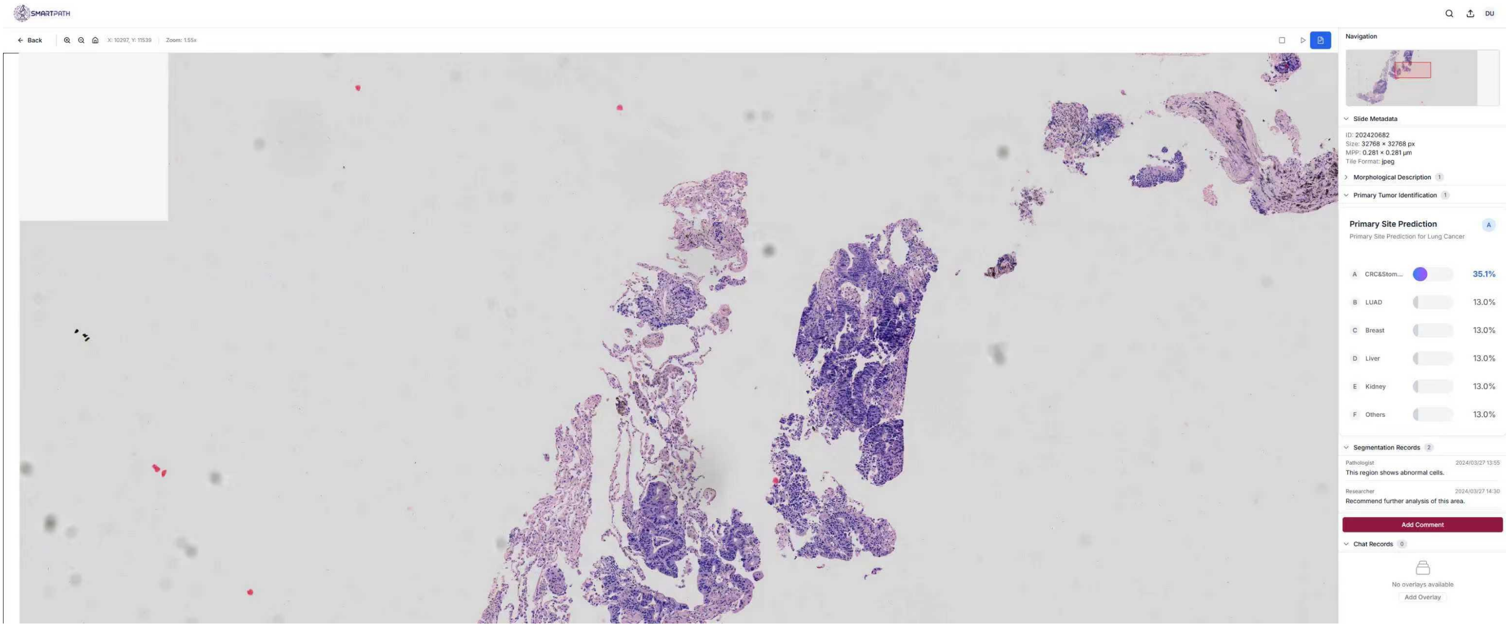


Figure S3. Confusion matrices for DeepPMD and TOAD models on primary site prediction, stratified by biopsy type. The performance of the DeepPMD (top row) and TOAD (bottom row) models on the six-class primary tumor site prediction task is evaluated across different sample types: excisional biopsies, aspiration biopsies, and all combined samples. The matrices demonstrate the consistently superior classification accuracy of the DeepPMD model compared to the TOAD model across all three cohorts.



X: 10297, Y: 11539

Zoom: 1.55x

ID: 202420682
Size: 32768 × 32768 px
MPP: 0.281 × 0.281 µm
Tile Format: jpeg

Primary Site Prediction

Primary Site Prediction for Lung Cancer

A CRC&Stom... 35.1%
B LUAD 13.0%
C Breast 13.0%
D Liver 13.0%
E Kidney 13.0%
F Others 13.0%

Pathologist 2024/03/27 13:55
This region shows abnormal cells.

Researcher 2024/03/27 14:30
Recommend further analysis of this area.

FigureS4: Window interface diagram of the DeepPMD model system.

Supplementary Tables

	Internal set (Nanfang, n=773)		External set (n=1187)		External set #1 (Qianfoshan, n=453)		External set #2 (Hebei, n=547)		External set #3 (Qingyuan, n=187)	
Number of whole slide images	1314		1778		453		1007		318	
Median age, years	58(18-86)		61(19-86)		61(31-86)		61(19-82)		60(30-83)	
Biopsy type										
Excision	420	54.33%	701	59.06%	281	62.03%	301	55.03%	119	63.67%
Aspiration	353	45.67%	486	40.94%	172	37.97%	246	44.97%	68	36.36%
Sex										
Male	415	53.69%	611	51.47%	245	54.08%	266	48.63%	100	53.48%
Female	358	46.31%	576	48.53%	208	45.92%	281	51.37%	87	46.52%
Histological diagnosis										
Lung adenocarcinoma	391	50.58%	625	52.65%	233	51.43%	273	49.91%	119	63.64%
Gastric & Colorectal Cancer	183	23.67%	274	23.08%	104	22.96%	140	25.59%	30	16.04%
Breast Cancer	51	6.60%	126	10.61%	50	11.04%	63	11.52%	13	6.95%
Renal Cancer	25	3.23%	81	6.82%	30	6.62%	43	7.86%	8	4.28%
Liver Cancer	34	4.40%	18	1.52%	10	2.21%	5	0.91%	3	1.60%
Others	89	11.51%	63	5.31%	26	5.74%	23	4.20%	14	7.49%
Scanner	Leica				Ibingli		Leica		3DHISTECH	

Table S1: Baseline characteristics of participants

Data set	Model	Macro-AUC (95% CI)	Macro-ACC (95% CI)	Macro-F1 (95% CI)	Micro-AUC (95% CI)	Micro-ACC (95% CI)	Weighted-F1 (95% CI)
Internal set	TOAD	0.896 (0.845,0.940)	0.828 (0.765,0.883)	0.827 (0.765,0.883))	0.890 (0.838,0.937)	0.828 (0.766,0.886)	0.828 (0.766,0.886)
	DeepPMD	0.993 (0.984,0.999)	0.936 (0.894,0.970)	0.936 (0.892,0.968)	0.992 (0.982,0.998)	0.936 (0.892,0.968)	0.936 (0.892,0.968)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #1	TOAD	0.648 (0.597,0.700)	0.606 (0.561,0.652)	0.606 (0.561,0.651)	0.645 (0.593,0.698)	0.607 (0.561,0.653)	0.607 (0.561,0.653)
	DeepPMD	0.850 (0.813,0.884)	0.807 (0.772,0.842)	0.802 (0.764,0.840)	0.865 (0.831,0.898)	0.803 (0.766,0.841)	0.801 (0.763,0.840)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #2	TOAD	0.785 (0.746,0.824)	0.712 (0.671,0.748)	0.709 (0.667,0.746)	0.788 (0.749,0.826)	0.711 (0.669,0.748)	0.709 (0.667,0.747)
	DeepPMD	0.977 (0.966,0.986)	0.903 (0.877,0.928)	0.903 (0.877,0.928)	0.974 (0.961,0.984)	0.903 (0.878,0.929)	0.903 (0.877,0.928)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #3	TOAD	0.922 (0.877,0.963)	0.844 (0.791,0.894)	0.823 (0.767,0.877)	0.876 (0.821,0.928)	0.829 (0.775,0.882)	0.832 (0.779,0.884)
	DeepPMD	0.981 (0.962,0.995)	0.918 (0.877,0.954)	0.899 (0.850,0.943)	0.967 (0.939,0.990)	0.904 (0.856,0.947)	0.905 (0.859,0.947)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*

Table S2: Performance of the DeepPMD and TOAD models in distinguishing primary and metastatic cancer in the indicated data set. Macro-AUC = area under the curve for each class averaged. Macro-ACC = accuracy for each class averaged. Macro-F1 = F1 score for each class averaged. Micro-AUC = area under the curve calculated globally. Micro-ACC = accuracy calculated globally. Weighted-F1 = F1 score weighted by class sample size. CI = confidence interval.

Data set	Model	Macro-AUC (95% CI)	Macro-ACC (95% CI)	Macro-F1 (95% CI)	Micro-AUC (95% CI)	Micro-ACC (95% CI)	Weighted-F1 (95% CI)
Internal set	TOAD	0.885 (0.838,0.928)	0.516 (0.408,0.635)	0.533 (0.417,0.654)	0.918 (0.884,0.948)	0.708 (0.633,0.778)	0.696 (0.615,0.770)
	DeepPMD	0.992 (0.985,0.998)	0.933 (0.884,0.974)	0.929 (0.881,0.968)	0.995 (0.991,0.998)	0.930 (0.886,0.968)	0.931 (0.890,0.968)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #1	TOAD	0.719 (0.671,0.761)	0.269 (0.232,0.307)	0.259 (0.218,0.304)	0.828 (0.803,0.851)	0.537 (0.492,0.581)	0.489 (0.438,0.537)
	DeepPMD	0.914 (0.889,0.935)	0.648 (0.575,0.716)	0.660 (0.594,0.717)	0.944 (0.929,0.958)	0.741 (0.702,0.779)	0.748 (0.708,0.786)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #2	TOAD	0.790 (0.722,0.838)	0.457 (0.354,0.544)	0.468 (0.369,0.546)	0.877 (0.857,0.896)	0.643 (0.603,0.682)	0.616 (0.572,0.659)
	DeepPMD	0.973 (0.958,0.986)	0.864 (0.822,0.903)	0.854 (0.818,0.890)	0.982 (0.974,0.989)	0.883 (0.856,0.909)	0.885 (0.859,0.911)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
External set #3	TOAD	0.899 (0.751,0.940)	0.476 (0.367,0.609)	0.429 (0.332,0.535)	0.926 (0.897,0.951)	0.723 (0.652,0.786)	0.723 (0.656,0.784)
	DeepPMD	0.981 (0.823,0.996)	0.860 (0.731,0.949)	0.827 (0.683,0.912)	0.990 (0.982,0.996)	0.893 (0.845,0.936)	0.900 (0.857,0.939)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*

Table S3: Performance of the DeepPMD and TOAD models in predicting primary tumor sites in the indicated data set.

Macro-AUC = area under the curve for each class averaged. Macro-ACC = accuracy for each class averaged. Macro-F1 = F1 score for each class averaged. Micro-AUC = area under the curve calculated globally. Micro-ACC = accuracy calculated globally. Weighted-F1 = F1 score weighted by class sample size. CI = confidence interval.

Data set	Model	Macro-AUC (95% CI)	Macro-ACC (95% CI)	Macro-F1 (95% CI)	Micro-AUC (95% CI)	Micro-ACC (95% CI)	Weighted-F1 (95% CI)
Excisional biopsy	TOAD	0.843 (0.814,0.870)	0.765 (0.732,0.798)	0.723 (0.689,0.757)	0.791 (0.757,0.825)	0.750 (0.719,0.781)	0.762 (0.733,0.790)
	DeepPMD	0.962 (0.950,0.973)	0.889 (0.869,0.910)	0.837 (0.808,0.865)	0.915 (0.892,0.936)	0.854 (0.828,0.879)	0.860 (0.836,0.884)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
Aspiration biopsy	TOAD	0.723 (0.673,0.770)	0.663 (0.617,0.708)	0.582 (0.537,0.626)	0.702 (0.656,0.746)	0.604 (0.560,0.646)	0.632 (0.591,0.671)
	DeepPMD	0.944 (0.920,0.966)	0.844 (0.801,0.883)	0.841 (0.799,0.877)	0.959 (0.941,0.973)	0.883 (0.852,0.909)	0.883 (0.853,0.911)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*

Table S4: Performance of DeepPMD and TOAD in distinguishing primary and metastatic cancer in the excisional biopsy and aspiration biopsy. Macro-AUC = area under the curve for each class averaged. Macro-ACC = accuracy for each class averaged. Macro-F1 = F1 score for each class averaged. Micro-AUC = area under the curve calculated globally. Micro-ACC = accuracy calculated globally. Weighted-F1 = F1 score weighted by class sample size. CI = confidence interval.

Data set	Model	Macro-AUC (95% CI)	Macro-ACC (95% CI)	Macro-F1 (95% CI)	Micro-AUC (95% CI)	Micro-ACC (95% CI)	Weighted-F1 (95% CI)
Excisional biopsy	TOAD	0.874 (0.829,0.906)	0.457 (0.372,0.572)	0.429 (0.346,0.519)	0.928 (0.915,0.941)	0.744 (0.714,0.775)	0.749 (0.718,0.782)
	DeepPMD	0.974 (0.949,0.991)	0.850 (0.747,0.925)	0.767 (0.659,0.827)	0.980 (0.973,0.986)	0.868 (0.843,0.893)	0.887 (0.867,0.908)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*
Aspiration biopsy	TOAD	0.731 (0.694,0.766)	0.342 (0.297,0.394)	0.332 (0.275,0.394)	0.773 (0.748,0.798)	0.430 (0.387,0.473)	0.391 (0.346,0.439)
	DeepPMD	0.935 (0.913,0.953)	0.728 (0.670,0.781)	0.756 (0.702,0.801)	0.952 (0.938,0.964)	0.777 (0.739,0.815)	0.775 (0.735,0.812)
	P value	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*	*P<0.001*

Table S5: Performance of the DeepPMD and TOAD models in predicting primary tumor sites in the excisional biopsy and aspiration biopsy. Macro-AUC = area under the curve for each class averaged. Macro-ACC = accuracy for each class averaged. Macro-F1 = F1 score for each class averaged. Micro-AUC = area under the curve calculated globally. Micro-ACC = accuracy calculated globally. Weighted-F1 = F1 score weighted by class sample size. CI = confidence interval.

Data set	Model	AUC (95% CI)	ACC (95% CI)	Weighted-F1 (95% CI)
Lung adenocarcinoma	TOAD	0.730 (0.684, 0.778)	0.889 (0.828, 0.944)	0.505 (0.447, 0.561)
	DeepPMD	0.936 (0.907, 0.959)	0.854 (0.787, 0.916)	0.765 (0.703, 0.822)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Gastric & Colorectal Cancer	TOAD	0.776 (0.733, 0.822)	0.443 (0.368, 0.522)	0.526 (0.456, 0.597)
	DeepPMD	0.965 (0.947, 0.979)	0.880 (0.824, 0.928)	0.865 (0.822, 0.904)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Breast Cancer	TOAD	0.685 (0.627, 0.744)	0.157 (0.090, 0.224)	0.229 (0.137, 0.317)
	DeepPMD	0.937 (0.913, 0.958)	0.678 (0.581, 0.770)	0.706 (0.629, 0.774)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Renal Cancer	TOAD	0.731 (0.649, 0.808)	0.074 (0.016, 0.153)	0.134 (0.031, 0.261)
	DeepPMD	0.966 (0.924, 0.992)	0.779 (0.660, 0.885)	0.848 (0.761, 0.917)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Liver Cancer	TOAD	0.779 (0.620, 0.908)	0.213 (0.000, 0.462)	0.301 (0.000, 0.583)
	DeepPMD	0.949 (0.868, 1.000)	0.716 (0.444, 0.938)	0.828 (0.615, 0.968)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Others	TOAD	0.678 (0.583, 0.763)	0.275 (0.151, 0.410)	0.295 (0.172, 0.423)
	DeepPMD	0.856 (0.787, 0.916)	0.466 (0.320, 0.609)	0.526 (0.388, 0.653)
	P value	*P<0.001*	*P<0.001*	*P<0.001*

Table S6: Performance of the DeepPMD and TOAD models in predicting primary tumor sites on aspiration biopsy samples.

AUC = area under the curve.ACC = accuracy. F1 = F1 score. CI = confidence interval.

Data set	Model	AUC (95% CI)	ACC (95% CI)	F1 (95% CI)
Lung adenocarcinoma	TOAD	0.849 (0.817, 0.879)	0.829 (0.796, 0.861)	0.848 (0.822, 0.871)
	DeepPMD	0.975 (0.964, 0.983)	0.855 (0.825, 0.884)	0.914 (0.896, 0.931)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Gastric & Colorectal Cancer	TOAD	0.896 (0.870, 0.922)	0.654 (0.566, 0.741)	0.621 (0.548, 0.691)
	DeepPMD	0.993 (0.987, 0.997)	0.931 (0.882, 0.973)	0.907 (0.868, 0.942)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Breast Cancer	TOAD	0.843 (0.741, 0.933)	0.399 (0.219, 0.583)	0.354 (0.197, 0.500)
	DeepPMD	0.966 (0.938, 0.989)	0.865 (0.727, 0.971)	0.490 (0.367, 0.602)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Renal Cancer	TOAD	0.895 (0.809, 0.962)	0.149 (0.036, 0.304)	0.245 (0.067, 0.452)
	DeepPMD	0.999 (0.996, 1.000)	0.924 (0.809, 1.000)	0.960 (0.895, 1.000)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Liver Cancer	TOAD	0.937 (0.855, 0.992)	0.253 (0.000, 0.750)	0.268 (0.000, 0.667)
	DeepPMD	1.000 (1.000, 1.000)	0.740 (0.250, 1.000)	0.828 (0.400, 1.000)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Others	TOAD	0.846 (0.746, 0.926)	0.472 (0.231, 0.714)	0.259 (0.113, 0.406)
	DeepPMD	0.933 (0.821, 0.994)	0.766 (0.545, 0.944)	0.499 (0.326, 0.655)
	P value	*P<0.001*	*P<0.001*	*P<0.001*

Table S7: Performance of the DeepPMD and TOAD models in predicting primary tumor sites on excisional biopsy samples.

AUC = area under the curve.ACC = accuracy. F1 = F1 score. CI = confidence interval.

Data set	Model	AUC (95% CI)	ACC (95% CI)	F1 (95% CI)
Lung adenocarcinoma	TOAD	0.763 (0.737, 0.789)	0.839 (0.812, 0.868)	0.748 (0.723, 0.773)
	DeepPMD	0.946 (0.934, 0.957)	0.855 (0.826, 0.881)	0.882 (0.863, 0.901)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Gastric & Colorectal Cancer	TOAD	0.837 (0.811, 0.862)	0.533 (0.474, 0.589)	0.573 (0.519, 0.624)
	DeepPMD	0.983 (0.975, 0.989)	0.901 (0.865, 0.934)	0.884 (0.855, 0.912)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Breast Cancer	TOAD	0.704 (0.650, 0.757)	0.215 (0.146, 0.290)	0.272 (0.191, 0.356)
	DeepPMD	0.934 (0.914, 0.952)	0.723 (0.647, 0.797)	0.629 (0.561, 0.692)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Renal Cancer	TOAD	0.762 (0.689, 0.826)	0.098 (0.037, 0.169)	0.174 (0.070, 0.281)
	DeepPMD	0.982 (0.960, 0.996)	0.826 (0.735, 0.905)	0.886 (0.825, 0.935)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Liver Cancer	TOAD	0.798 (0.654, 0.911)	0.218 (0.045, 0.438)	0.294 (0.071, 0.526)
	DeepPMD	0.975 (0.935, 1.000)	0.722 (0.500, 0.923)	0.834 (0.667, 0.960)
	P value	*P<0.001*	*P<0.001*	*P<0.001*
Others	TOAD	0.737 (0.666, 0.800)	0.329 (0.214, 0.452)	0.282 (0.188, 0.376)
	DeepPMD	0.890 (0.837, 0.936)	0.549 (0.427, 0.673)	0.517 (0.411, 0.619)
	P value	*P<0.001*	*P<0.001*	*P<0.001*

Table S8: Performance of the DeepPMD and TOAD models in predicting primary tumor sites on all combined biopsy samples.

AUC = area under the curve.ACC = accuracy. F1 = F1 score. CI = confidence interval.