
A universal visual foundation model for computational cytopathology

In the format provided by the
authors and unedited

A Universal Visual Foundation Model for Computational Cytopathology

Supplementary Information

Supplementary Tables

Supplementary Table 1. Overview of 202 Downstream Tasks

Num.	Type of Tasks	Cohort	Downstream Tasks	Abbr.	Class	Num. of patches/slides	Task level
1	Zero-shot	SMH cohort	Pleural Effusion Malignancy	Pleural-BMC	2	2739	Patch
2	Zero-shot	SMH cohort	Ascites Malignancy	Ascites-BMC	2	2866	Patch
3	Zero-shot	SMH cohort	Pericardial Effusion Malignancy	Pericardial-BMC	2	177	Patch
4	Zero-shot	SMH cohort	Thyroid Nodule Malignancy	Thyroid-BMC	2	3568	Patch
5	Zero-shot	SMH cohort	Thyroid Cancer Subtype	Thyroid-CSC	2	1262	Patch
6	Zero-shot	SMH cohort	Cerebrospinal Fluid Malignancy	CSF-BMC	2	573	Patch
7	Zero-shot	SMH cohort	Bronchial Brushing Malignancy	Bronchial-BMC	2	2290	Patch
8	Zero-shot	SMH cohort	Bronchial Brushing Cancer Subtype	Bronchial-CSC	2	202	Patch
9	Zero-shot	SMH cohort	Cervical Cell Abnormality	Cervix-AD	2	2398	Patch
10	Zero-shot	SMH cohort	Cervical Cell Bacterial Infection	Cervix-BID	2	1720	Patch
11	Zero-shot	SMH cohort	Cervical Cell Candida Infection	Cervix-CID	2	1770	Patch
12	Zero-shot	SMH cohort	Salivary Gland Tumor Malignancy	Salivary-BMC	2	746	Patch
13	Zero-shot	SMH cohort	Vocal Cord Lesions Malignancy	VocalCord-BMC	2	62	Patch
14	Zero-shot	SMH cohort	Breast Lesions Malignancy	Breast-BMC	2	171	Patch
15	Zero-shot	SMH cohort	Pancreatic Lesions Malignancy	Pancreas-BMC	2	201	Patch
16	Zero-shot	SMH cohort	Liver Lesions Malignancy	Liver-BMC	2	229	Patch
17	Zero-shot	SMH cohort	Bile Duct Malignancy	BileDuct-BMC	2	165	Patch
18	Zero-shot	SMH cohort	Bile Cell Malignancy	Bile-BMC	2	673	Patch
19	Zero-shot	SMH cohort	Lymph Node Malignancy	LN-BMC	2	4386	Patch
20	Zero-shot	SMH cohort	Urinary Cell Abnormality	Urine-AD	2	2860	Patch
21	Zero-shot	SMH cohort	TIS System Classification	TIS-CLS	5	1462	Patch
22	Zero-shot	SMH cohort	Thyroid Bethesda Categories	Thyroid-BC	6	4145	Patch
23	Zero-shot	SMH cohort	Urinary Cell Grading	Urine-G	3	2846	Patch
24	Zero-shot	SMH cohort	Cervical Cell Grading	Cervix-G	6	2400	Patch
25	Zero-shot	SMH cohort	Cervical Cell HPV Infection	Cervix-HPVInf	4	2200	Patch
26	Zero-shot	SMH cohort	Salivary Gland Tumor Subtype	Salivary-CSC	3	541	Patch
27	Zero-shot	SMH cohort	Milan System Classification	Milan-CLS	3	120	Patch
28	Zero-shot	SMH cohort	Liver Lesion Subtype	Liver-CSC	4	166	Patch
29	Zero-shot	SMH cohort	Bile Duct Grading	BileDuct-G	4	205	Patch
30	Zero-shot	SMH cohort	Bile Cell Grading	Bile-G	4	894	Patch
31	Zero-shot	SMH cohort	Parasite Egg and Fungal Detection in Bile	Bile-PFD	3	830	Patch

32	Zero-shot	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer	LN-Origin	5	911	Patch
33	Zero-shot	SMH cohort	Lymph Node Cancer Subtype	LN-CSC	5	1041	Patch
34	Zero-shot	SMH cohort	Benign Lymph Node Classification	LN-Benign	3	864	Patch
35	Zero-shot	SCC cohort	Cervical Cell Abnormality	Cervix-AD	2	2338	Patch
36	Zero-shot	SCC cohort	Cervical Cell Candida Infection	Cervix-CID	2	2570	Patch
37	Zero-shot	SCC cohort	Cervical Cell Bacterial Infection	Cervix-BID	2	2090	Patch
38	Zero-shot	SCC cohort	Pleural Effusion Malignancy	Pleural-BMC	2	2114	Patch
39	Zero-shot	SCC cohort	Ascites Malignancy	Ascites-BMC	2	2591	Patch
40	Zero-shot	SCC cohort	Pericardial Effusion Malignancy	Pericardial-BMC	2	330	Patch
41	Zero-shot	SCC cohort	Thyroid Nodule Malignancy	Thyroid-BMC	2	4793	Patch
42	Zero-shot	SCC cohort	Thyroid Cancer Subtype	Thyroid-CSC	2	2186	Patch
43	Zero-shot	SCC cohort	Urinary Cell Abnormality	Urine-AD	2	2593	Patch
44	Zero-shot	SCC cohort	Cerebrospinal Fluid Malignancy	CSF-BMC	2	1009	Patch
45	Zero-shot	SCC cohort	Bronchial Brushing Malignancy	Bronchial-BMC	2	2736	Patch
46	Zero-shot	SCC cohort	Breast Lesions Malignancy	Breast-BMC	2	114	Patch
47	Zero-shot	SCC cohort	Pancreatic Lesions Malignancy	Pancreas-BMC	2	980	Patch
48	Zero-shot	SCC cohort	Bile Duct Malignancy	BileDuct-BMC	2	107	Patch
49	Zero-shot	SCC cohort	Bile Cell Malignancy	Bile-BMC	2	283	Patch
50	Zero-shot	SCC cohort	Parasite Egg Detection in Bile	Bile-PED	2	301	Patch
51	Zero-shot	SCC cohort	Lymph Node Malignancy	LN-BMC	2	1777	Patch
52	Zero-shot	SCC cohort	Cervical Cell Grading	Cervix-G	6	2325	Patch
53	Zero-shot	SCC cohort	Thyroid Bethesda Categories	Thyroid-BC	6	5361	Patch
54	Zero-shot	SCC cohort	Urinary Cell Grading	Urine-G	3	2417	Patch
55	Zero-shot	SCC cohort	Bronchial Brushing Cancer Subtype	Bronchial-CSC	3	978	Patch
56	Zero-shot	SCC cohort	Lymph Node Cancer Subtype	LN-CSC	3	915	Patch
57	Supervised	SMH cohort	Pleural Effusion Malignancy	Pleural-BMC	2	2739	Patch
58	Supervised	SMH cohort	Ascites Malignancy	Ascites-BMC	2	2866	Patch
59	Supervised	SMH cohort	Pericardial Effusion Malignancy	Pericardial-BMC	2	177	Patch
60	Supervised	SMH cohort	Thyroid Nodule Malignancy	Thyroid-BMC	2	3568	Patch
61	Supervised	SMH cohort	Thyroid Cancer Subtype	Thyroid-CSC	2	1262	Patch
62	Supervised	SMH cohort	Cerebrospinal Fluid Malignancy	CSF-BMC	2	573	Patch
63	Supervised	SMH cohort	Bronchial Brushing Malignancy	Bronchial-BMC	2	2290	Patch
64	Supervised	SMH cohort	Bronchial Brushing Cancer Subtype	Bronchial-CSC	2	202	Patch
65	Supervised	SMH cohort	Cervical Cell Abnormality	Cervix-AD	2	2398	Patch
66	Supervised	SMH cohort	Cervical Cell Bacterial Infection	Cervix-BID	2	1720	Patch
67	Supervised	SMH cohort	Cervical Cell Candida Infection	Cervix-CID	2	1770	Patch
68	Supervised	SMH cohort	Salivary Gland Tumor Malignancy	Salivary-BMC	2	746	Patch
69	Supervised	SMH cohort	Vocal Cord Lesions Malignancy	VocalCord-BMC	2	62	Patch

70	Supervised	SMH cohort	Breast Lesions Malignancy	Breast-BMC	2	171	Patch
71	Supervised	SMH cohort	Pancreatic Lesions Malignancy	Pancreas-BMC	2	201	Patch
72	Supervised	SMH cohort	Liver Lesions Malignancy	Liver-BMC	2	229	Patch
73	Supervised	SMH cohort	Bile Duct Malignancy	BileDuct-BMC	2	165	Patch
74	Supervised	SMH cohort	Bile Cell Malignancy	Bile-BMC	2	673	Patch
75	Supervised	SMH cohort	Lymph Node Malignancy	LN-BMC	2	4386	Patch
76	Supervised	SMH cohort	Urinary Cell Abnormality	Urine-AD	2	2860	Patch
77	Supervised	SMH cohort	TIS System Classification	TIS-CLS	5	1462	Patch
78	Supervised	SMH cohort	Thyroid Bethesda Categories	Thyroid-BC	6	4145	Patch
79	Supervised	SMH cohort	Urinary Cell Grading	Urine-G	3	2846	Patch
80	Supervised	SMH cohort	Cervical Cell Grading	Cervix-G	6	2400	Patch
81	Supervised	SMH cohort	Cervical Cell HPV Infection	Cervix-HPVInf	4	2200	Patch
82	Supervised	SMH cohort	Salivary Gland Tumor Subtype	Salivary-CSC	3	541	Patch
83	Supervised	SMH cohort	Milan System Classification	Milan-CLS	3	120	Patch
84	Supervised	SMH cohort	Liver Lesion Subtype	Liver-CSC	4	166	Patch
85	Supervised	SMH cohort	Bile Duct Grading	BileDuct-G	4	205	Patch
86	Supervised	SMH cohort	Bile Cell Grading	Bile-G	4	894	Patch
87	Supervised	SMH cohort	Parasite Egg and Fungal Detection in Bile	Bile-PFD	3	830	Patch
88	Supervised	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer	LN-Origin	5	911	Patch
89	Supervised	SMH cohort	Lymph Node Cancer Subtype	LN-CSC	5	1041	Patch
90	Supervised	SMH cohort	Benign Lymph Node Classification	LN-Benign	3	864	Patch
91	Supervised	SCC cohort	Cervical Cell Abnormality	Cervix-AD	2	2338	Patch
92	Supervised	SCC cohort	Cervical Cell Candida Infection	Cervix-CID	2	2570	Patch
93	Supervised	SCC cohort	Cervical Cell Bacterial Infection	Cervix-BID	2	2090	Patch
94	Supervised	SCC cohort	Pleural Effusion Malignancy	Pleural-BMC	2	2114	Patch
95	Supervised	SCC cohort	Ascites Malignancy	Ascites-BMC	2	2591	Patch
96	Supervised	SCC cohort	Pericardial Effusion Malignancy	Pericardial-BMC	2	330	Patch
97	Supervised	SCC cohort	Thyroid Nodule Malignancy	Thyroid-BMC	2	4793	Patch
98	Supervised	SCC cohort	Thyroid Cancer Subtype	Thyroid-CSC	2	2186	Patch
99	Supervised	SCC cohort	Urinary Cell Abnormality	Urine-AD	2	2593	Patch
100	Supervised	SCC cohort	Cerebrospinal Fluid Malignancy	CSF-BMC	2	1009	Patch
101	Supervised	SCC cohort	Bronchial Brushing Malignancy	Bronchial-BMC	2	2736	Patch
102	Supervised	SCC cohort	Breast Lesions Malignancy	Breast-BMC	2	114	Patch
103	Supervised	SCC cohort	Pancreatic Lesions Malignancy	Pancreas-BMC	2	980	Patch
104	Supervised	SCC cohort	Bile Duct Malignancy	BileDuct-BMC	2	107	Patch
105	Supervised	SCC cohort	Bile Cell Malignancy	Bile-BMC	2	283	Patch
106	Supervised	SCC cohort	Parasite Egg Detection in Bile	Bile-PED	2	301	Patch

107	Supervised	SCC cohort	Lymph Node Malignancy	LN-BMC	2	1777	Patch
108	Supervised	SCC cohort	Cervical Cell Grading	Cervix-G	6	2325	Patch
109	Supervised	SCC cohort	Thyroid Bethesda Categories	Thyroid-BC	6	5361	Patch
110	Supervised	SCC cohort	Urinary Cell Grading	Urine-G	3	2417	Patch
111	Supervised	SCC cohort	Bronchial Brushing Cancer Subtype	Bronchial-CSC	3	978	Patch
112	Supervised	SCC cohort	Lymph Node Cancer Subtype	LN-CSC	3	915	Patch
113	Supervised	Public Dataset	SipakMed	Cervix 1	5	966	Patch
114	Supervised	Public Dataset	SipakMed (Cropped)	Cervix 2	5	3532	Patch
115	Supervised	Public Dataset	ComparisonDetector	Cervix 3	11	1563	Patch
116	Supervised	Public Dataset	Mendeley_aug	Cervix 4	4	3232	Patch
117	Supervised	Public Dataset	Mendeley_ori	Cervix 5	4	962	Patch
118	Supervised	Public Dataset	Body Cavity Fluid Cytology Images	Effusion 1	2	693	Patch
119	Supervised	Public Dataset	Ascites Cytopathology	Effusion 2	6	3219	Patch
120	Supervised	Public Dataset	Thyroid Cytology	Thyroid 1	2	2364	Patch
121	Supervised	Public Dataset	Thyroid Dataset	Thyroid 2	3	2875	Patch
122	Supervised	Public Dataset	C-NMC_Leukemia	Blood 1	2	12528	Patch
123	Supervised	Public Dataset	Blood Cell Images	Blood 2	4	12444	Patch
124	Supervised	Public Dataset	Raabin-WBC	Blood 3	5	14514	Patch
125	Site Retrieval	SMH cohort	Ascites	Ascites	14	286	Patch
126	Site Retrieval	SMH cohort	Thyroid	Thyroid	14	356	Patch
127	Site Retrieval	SMH cohort	Urine	Urine	14	284	Patch
128	Site Retrieval	SMH cohort	Cerebrospinal Fluid	CSF	14	56	Patch
129	Site Retrieval	SMH cohort	Bronchial Brushing	Bronchial	14	20	Patch
130	Site Retrieval	SMH cohort	Cervix	Cervix	14	237	Patch
131	Site Retrieval	SMH cohort	Salivary	Salivary	14	52	Patch
132	Site Retrieval	SMH cohort	Vocal cord	VocalCord	14	5	Patch
133	Site Retrieval	SMH cohort	Breast	Breast	14	16	Patch
134	Site Retrieval	SMH cohort	Pancreas	Pancreas	14	19	Patch
135	Site Retrieval	SMH cohort	Liver	Liver	14	16	Patch
136	Site Retrieval	SMH cohort	Bile Duct	BileDuct	14	15	Patch
137	Site Retrieval	SMH cohort	Bile	Bile	14	88	Patch
138	Site Retrieval	SMH cohort	Lymph Node	LN	14	437	Patch
139	Site Retrieval	SCC cohort	Cervix	Cervix	11	230	Patch
140	Site Retrieval	SCC cohort	Ascites	Ascites	11	258	Patch
141	Site Retrieval	SCC cohort	Thyroid	Thyroid	11	218	Patch
142	Site Retrieval	SCC cohort	Urine	Urine	11	240	Patch
143	Site Retrieval	SCC cohort	Cerebrospinal Fluid	CSF	11	100	Patch
144	Site Retrieval	SCC cohort	Bronchial Brushing	Bronchial	11	97	Patch

145	Site Retrieval	SCC cohort	Breast	Breast	11	10	Patch
146	Site Retrieval	SCC cohort	Pancreas	Pancreas	11	97	Patch
147	Site Retrieval	SCC cohort	Bile Duct	BileDuct	11	10	Patch
148	Site Retrieval	SCC cohort	Bile	Bile	11	27	Patch
149	Site Retrieval	SCC cohort	Lymph node	LN	11	91	Patch
150	Fine-grained Retrieval	SMH cohort	Ascites Malignancy	Ascites-BMC	38	286	Patch
151	Fine-grained Retrieval	SMH cohort	Thyroid Nodule Malignancy	Thyroid-BMC	38	356	Patch
152	Fine-grained Retrieval	SMH cohort	Urinary Epithelial Cell Grading	Urine-G	38	284	Patch
153	Fine-grained Retrieval	SMH cohort	Cerebrospinal Fluid Malignancy	CSF-BMC	38	56	Patch
154	Fine-grained Retrieval	SMH cohort	Bronchial Brushing Cancer Subtype	Bronchial-CSC	38	20	Patch
155	Fine-grained Retrieval	SMH cohort	Cervical Cytology Grading	Cervix-G	38	237	Patch
156	Fine-grained Retrieval	SMH cohort	Salivary Gland Tumor Subtype	Salivary-CSC	38	52	Patch
157	Fine-grained Retrieval	SMH cohort	Vocal Cord Lesions Malignancy	VocalCord-BMC	38	5	Patch
158	Fine-grained Retrieval	SMH cohort	Breast Lesions Malignancy	Breast-BMC	38	16	Patch
159	Fine-grained Retrieval	SMH cohort	Pancreatic Lesions Malignancy	Pancreas-BMC	38	19	Patch
160	Fine-grained Retrieval	SMH cohort	Liver Lesion Subtype	Liver-CSC	38	16	Patch
161	Fine-grained Retrieval	SMH cohort	Bile Duct Malignancy	BileDuct-BMC	38	15	Patch
162	Fine-grained Retrieval	SMH cohort	Bile Cell Grading	Bile-G	38	88	Patch
163	Fine-grained Retrieval	SMH cohort	Lymph Node Malignancy	LN-BMC	38	437	Patch
164	Fine-grained Retrieval	SCC cohort	Cervical Cytology Abnormality Detection and Grading	Cervix-G	29	230	Patch
165	Fine-grained Retrieval	SCC cohort	Ascites Malignancy Classification	Ascites-BMC	29	258	Patch
166	Fine-grained Retrieval	SCC cohort	Thyroid Cancer Subtype Classification	Thyroid-CSC	29	218	Patch
167	Fine-grained Retrieval	SCC cohort	Urinary Epithelial Cell Abnormality Detection and Grading	Urine-G	29	240	Patch
168	Fine-grained Retrieval	SCC cohort	Cerebrospinal Fluid Malignancy Classification	CSF-BMC	29	100	Patch

169	Fine-grained Retrieval	SCC cohort	Cancer Subtype Classification for Bronchial Brushing Cytology	Bronchial-CSC	29	97	Patch
170	Fine-grained Retrieval	SCC cohort	Benign and Malignant Classification of Breast Lesions	Breast-BMC	29	10	Patch
171	Fine-grained Retrieval	SCC cohort	Benign and Malignant Classification of Pancreatic Lesions	Pancreas-BMC	29	97	Patch
172	Fine-grained Retrieval	SCC cohort	Bile Duct Brushing Abnormality Detection	BileDuct-BMC	29	10	Patch
173	Fine-grained Retrieval	SCC cohort	Benign and Malignant Classification of Bile Cells	Bile-BMC	29	27	Patch
174	Fine-grained Retrieval	SCC cohort	Lymph Node Cancer Subtype Classification	LN-CSC	29	91	Patch
175	Cell Segmentation	Public Dataset	Thyroid Dataset	Thyroid Dataset	NA	1473	Patch
176	Cell Segmentation	Public Dataset	Cervix Dataset	Cervix Dataset	NA	3565	Patch
177	Object Detection	Public Dataset	Blood Detect.1	Blood Detect.1	3	874	Patch
178	Object Detection	Public Dataset	Blood Detect.2	Blood Detect.2	3	364	Patch
179	Weakly-Supervised Classification	SMH cohort	Lymph Node Malignancy Classification	LN-BMC	2	791	Slide
180	Weakly-Supervised Classification	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer	LN-Origin	2	317	Slide
181	Weakly-Supervised Classification	SMH cohort	Thyroid Nodule Malignancy Classification	Thyroid-BMC	2	849	Slide
182	Weakly-Supervised Classification	SMH cohort	Cervical Cell Abnormality Detection	Cervix-AD	2	468	Slide
183	Weakly-Supervised Classification	SMH cohort	Cervical Cell HPV Infection	Cervix-HPVInf	2	508	Slide
184	Weakly-Supervised Classification	SMH cohort	Cervical Cell Candida Infection Detection	Cervix-CID	2	598	Slide
185	Weakly-Supervised Classification	SMH cohort	Cervical Cell Bacterial Infection Detection	Cervix-BID	2	502	Slide
186	Weakly-Supervised Classification	ZPH Cohort	Thyroid Nodule Malignancy (Cell Block)	Thyroid-BMC (CB)	2	100	Slide

187	Weakly-Supervised Classification	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer (Detailed)	LN-Origin (Detailed)	5	317	Slide
188	Weakly-Supervised Classification	SMH cohort	Lymph Node Cancer Subtype	LN-CSC	3	136	Slide
189	Weakly-Supervised Classification	SMH cohort	Cervical Cell HPV-Subtype Infection	Cervix-HPVInf (Detailed)	4	508	Slide
190	Weakly-Supervised Classification	ZPH Cohort	TIS System Classification	TIS-CLS (CB)	4	96	Slide
191	Few-shot Classification	SMH cohort	Lymph Node Malignancy Classification	LN-BMC	2	791	Slide
192	Few-shot Classification	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer	LN-Origin	2	317	Slide
193	Few-shot Classification	SMH cohort	Thyroid Nodule Malignancy Classification	Thyroid-BMC	2	849	Slide
194	Few-shot Classification	SMH cohort	Cervical Cell Abnormality Detection	Cervix-AD	2	468	Slide
195	Few-shot Classification	SMH cohort	Cervical Cell HPV Infection	Cervix-HPVInf	2	508	Slide
196	Few-shot Classification	SMH cohort	Cervical Cell Candida Infection Detection	Cervix-CID	2	598	Slide
197	Few-shot Classification	SMH cohort	Cervical Cell Bacterial Infection Detection	Cervix-BID	2	502	Slide
198	Few-shot Classification	ZPH Cohort	Thyroid Nodule Malignancy (Cell Block)	Thyroid-BMC (CB)	2	100	Slide
199	Few-shot Classification	SMH cohort	Primary Origin of Lymph Node Metastatic Cancer (Detailed)	LN-Origin (Detailed)	5	317	Slide
200	Few-shot Classification	SMH cohort	Lymph Node Cancer Subtype	LN-CSC	3	136	Slide
201	Few-shot Classification	SMH cohort	Cervical Cell HPV-Subtype Infection	Cervix-HPVInf (Detailed)	4	508	Slide
202	Few-shot Classification	ZPH Cohort	TIS System Classification	TIS-CLS (CB)	4	96	Slide

Note: SMH/SCC/ZPH are private cohorts. Num., number. Abbr., abbreviation. NA, not applicable.

Section I. Performance of Patch-level KNN Classifier in the SMH cohort

This section presents the performance of the patch-level KNN classifier across 34 downstream cytology tasks in the SMH cohort. These tasks include 20 binary and 14 multiclass classification problems, spanning diverse anatomical sites and pathological conditions. A cosine similarity-based KNN classifier (K=20) was applied: for each test sample, cosine similarities were computed against all training features, and the top-K neighbors were selected. Their labels were aggregated through similarity-weighted soft voting to generate the final prediction.

Supplementary Table 2. Performance of Patch-level KNN Classifier for Pleural Effusion Malignancy Classification in the SMH cohort

Task	Pleural-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged accuracy
CROWN	0.912	0.905	0.922	0.905	0.925	0.914
DINOv2	0.807	0.761	0.761	0.770	0.788	0.777
CTransPath	0.838	0.854	0.838	0.825	0.835	0.838
ResNet-BT	0.659	0.637	0.648	0.653	0.645	0.648
UNI	0.929	0.911	0.892	0.885	0.887	0.901
Virchow	0.891	0.889	0.891	0.887	0.905	0.892

Note: A total of 2739 pleural effusion cytology smears were included for a binary classification task: 963 malignant and 1776 benign cases.

Supplementary Table 3. Performance of Patch-level KNN Classifier for Ascites Malignancy Classification in the SMH cohort

Task	Ascites-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.930	0.929	0.956	0.927	0.935	0.935
DINOv2	0.808	0.817	0.831	0.807	0.796	0.812
CTransPath	0.855	0.861	0.887	0.871	0.884	0.872
ResNet-BT	0.740	0.760	0.763	0.709	0.772	0.749
UNI	0.943	0.948	0.920	0.913	0.937	0.932
Virchow	0.927	0.932	0.913	0.897	0.926	0.919

Note: A total of 2866 ascitic fluid cytology smears were included for a binary classification task: 720 malignant and 2146 benign cases.

Supplementary Table 4. Performance of Patch-level KNN Classifier for Pericardial Effusion Malignancy Classification in the SMH cohort

Task	Pericardial-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.861	0.972	0.806	0.806	0.939	0.877
DINOv2	0.833	0.833	0.750	0.694	0.606	0.743
CTransPath	0.750	0.861	0.778	0.861	0.758	0.802
ResNet-BT	0.528	0.694	0.694	0.639	0.636	0.638
UNI	0.806	0.861	0.861	0.889	0.818	0.847
Virchow	0.917	0.833	0.889	0.889	0.909	0.887

Note: A total of 177 pericardial effusion cytology smears were included for a binary classification task: 66 malignant and 111 benign cases.

Supplementary Table 5. Performance of Patch-level KNN Classifier for Thyroid Nodule Malignancy Classification in the SMH cohort

Task	Thyroid-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.964	0.962	0.961	0.964	0.944	0.959
DINOv2	0.770	0.793	0.810	0.800	0.802	0.795
CTransPath	0.919	0.922	0.920	0.930	0.941	0.926
ResNet-BT	0.821	0.808	0.835	0.836	0.840	0.828
UNI	0.964	0.959	0.965	0.971	0.968	0.965
Virchow	0.947	0.954	0.954	0.959	0.941	0.951

Note: A total of 3568 thyroid cytology smears were included for a binary classification task: 1034 malignant and 2534 benign cases.

Supplementary Table 6. Performance of Patch-level KNN Classifier for Thyroid Cancer Subtype Classification in the SMH cohort

Task	Thyroid-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.925	0.953	0.949	0.957	0.936	0.944
DINOv2	0.933	0.925	0.893	0.953	0.888	0.918
CTransPath	0.964	0.949	0.941	0.960	0.952	0.953
ResNet-BT	0.925	0.909	0.921	0.901	0.936	0.918
UNI	0.972	0.933	0.953	0.984	0.968	0.962
Virchow	0.953	0.945	0.957	0.933	0.940	0.945

Note: A total of 1262 thyroid cancer cytology smears were included for a binary classification task: 1122 papillary carcinoma and 140 other carcinoma cases.

Supplementary Table 7. Performance of Patch-level KNN Classifier for Cerebrospinal Fluid Malignancy Classification in the SMH cohort

Task	CSF-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.974	0.991	0.974	0.904	0.991	0.967
DINOv2	0.913	0.913	0.922	0.887	0.894	0.906
CTransPath	0.939	0.922	0.939	0.939	0.965	0.941
ResNet-BT	0.887	0.843	0.870	0.878	0.876	0.871
UNI	0.965	0.922	0.991	0.983	0.982	0.969
Virchow	0.930	0.974	0.965	0.974	0.965	0.962

Note: A total of 573 cerebrospinal fluid cytology smears were included for a binary classification task: 74 malignant and 499 benign cases.

Supplementary Table 8. Performance of Patch-level KNN Classifier for Bronchial Brushing Malignancy Classification in the SMH cohort

Task	Bronchial-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.952	0.934	0.930	0.945	0.952	0.943
DINOv2	0.934	0.908	0.913	0.919	0.897	0.914
CTransPath	0.928	0.932	0.924	0.897	0.913	0.919
ResNet-BT	0.893	0.915	0.902	0.937	0.930	0.915
UNI	0.919	0.917	0.930	0.932	0.948	0.929
Virchow	0.924	0.921	0.934	0.937	0.932	0.930

Note: A total of 2290 bronchial brushing and bronchoalveolar lavage smears were included for a binary classification task: 194 malignant and 2096 benign cases.

Supplementary Table 9. Performance of Patch-level KNN Classifier for Cancer Subtype Classification of Bronchial Brushing Cytology in the SMH cohort

Task	Bronchial-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.951	0.829	0.951	0.902	0.895	0.906
DINOv2	0.829	0.902	0.927	0.732	0.842	0.846
CTransPath	0.829	0.829	0.805	0.878	0.921	0.853
ResNet-BT	0.829	0.805	0.854	0.854	0.895	0.847
UNI	0.927	0.927	0.854	0.878	0.947	0.907
Virchow	0.854	0.854	0.878	0.951	0.921	0.892

Note: A total of 202 malignant bronchial cytology smears were included for a binary classification task: 171 non-small cell carcinoma and 31 small cell carcinoma cases.

Supplementary Table 10. Performance of Patch-level KNN Classifier for Cervical Cell Abnormality Detection in the SMH cohort

Task	Cervix-AD					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.971	0.967	0.971	0.952	0.971	0.966
DINOv2	0.877	0.892	0.888	0.892	0.893	0.888
CTransPath	0.898	0.908	0.925	0.892	0.897	0.904
ResNet-BT	0.925	0.875	0.898	0.910	0.904	0.902
UNI	0.942	0.944	0.965	0.952	0.950	0.950
Virchow	0.915	0.925	0.948	0.902	0.929	0.924

Note: A total of 2398 cervical cytology smears were included for a binary classification task: 1013 normal and 1385 abnormal cases.

Supplementary Table 11. Performance of Patch-level KNN Classifier for Cervical Cell Bacterial Infection Detection in the SMH cohort

Task	Cervix-BID					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.948	0.968	0.933	0.953	0.953	0.951
DINOv2	0.756	0.823	0.765	0.785	0.765	0.778
CTransPath	0.834	0.840	0.820	0.846	0.808	0.830
ResNet-BT	0.791	0.840	0.773	0.826	0.828	0.812
UNI	0.898	0.892	0.927	0.927	0.916	0.912
Virchow	0.855	0.881	0.843	0.875	0.895	0.870

Note: A total of 1720 cervical cytology smears were included for a binary classification task: 1018 without bacterial infection and 702 with bacterial vaginosis.

Supplementary Table 12. Performance of Patch-level KNN Classifier for Cervical Cell Candida Infection Detection in the SMH cohort

Task	Cervix-CID					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.935	0.958	0.958	0.963	0.944	0.951
DINOv2	0.864	0.873	0.870	0.879	0.864	0.870
CTransPath	0.901	0.893	0.915	0.932	0.901	0.908
ResNet-BT	0.901	0.881	0.862	0.867	0.867	0.876
UNI	0.969	0.932	0.932	0.972	0.938	0.949
Virchow	0.921	0.879	0.901	0.912	0.895	0.902

Note: A total of 1770 cervical cytology smears were included for a binary classification task: 1018 without fungal infection and 752 with Candida infection.

Supplementary Table 13. Performance of Patch-level KNN Classifier for Salivary Gland Tumor Malignancy Classification in the SMH cohort

Task	Salivary-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.880	0.847	0.873	0.900	0.808	0.862
DINOv2	0.793	0.820	0.773	0.767	0.842	0.799
CTransPath	0.807	0.860	0.853	0.787	0.870	0.835
ResNet-BT	0.793	0.813	0.833	0.827	0.801	0.814
UNI	0.867	0.887	0.853	0.873	0.815	0.859
Virchow	0.920	0.860	0.860	0.893	0.849	0.877

Note: A total of 746 salivary gland cytology smears were included for a binary classification task: 148 malignant and 598 benign cases.

Supplementary Table 14. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Vocal Cord Lesions in the SMH cohort

Task	VocalCord-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	1.000	0.923	0.692	0.769	0.800	0.837
DINOv2	0.846	0.538	0.846	0.692	0.600	0.705
CTransPath	0.538	0.846	0.692	0.769	0.700	0.709
ResNet-BT	0.615	0.769	0.615	0.769	0.700	0.694
UNI	0.692	0.692	0.692	0.846	0.900	0.765
Virchow	0.615	0.769	0.769	0.846	0.400	0.680

Note: A total of 62 vocal cord mass cytology smears were included for a binary classification task: 19 malignant and 43 benign cases.

Supplementary Table 15. Performance of Patch-level KNN Classifier for Breast Lesion Malignancy Classification in the SMH cohort

Task	Breast-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.829	0.829	0.857	0.943	0.742	0.840
DINOv2	0.771	0.686	0.829	0.771	0.613	0.734
CTransPath	0.629	0.771	0.771	0.714	0.774	0.732
ResNet-BT	0.771	0.714	0.686	0.714	0.742	0.726
UNI	0.943	0.771	0.771	0.743	0.839	0.813
Virchow	0.886	0.714	0.914	0.829	0.839	0.836

Note: A total of 171 breast mass cytology smears were included for a binary classification task: 47 malignant and 124 benign cases.

Supplementary Table 16. Performance of Patch-level KNN Classifier for Benign and Malignant Pancreatic Lesions in the SMH cohort

Task	Pancreas-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.780	0.854	0.610	0.780	0.703	0.745
DINOv2	0.805	0.634	0.756	0.683	0.784	0.732
CTransPath	0.756	0.683	0.780	0.780	0.730	0.746
ResNet-BT	0.683	0.732	0.707	0.756	0.784	0.732
UNI	0.829	0.683	0.829	0.756	0.730	0.765
Virchow	0.634	0.756	0.854	0.878	0.730	0.770

Note: A total of 201 pancreatic mass cytology smears were included for a binary classification task: 147 malignant and 54 benign cases.

Supplementary Table 17. Performance of Patch-level KNN Classifier for Benign and Malignant Liver Lesions in the SMH cohort

Task	Liver-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.826	0.848	0.783	0.696	0.800	0.790
DINOv2	0.739	0.717	0.674	0.609	0.689	0.686
CTransPath	0.804	0.826	0.674	0.717	0.644	0.733
ResNet-BT	0.783	0.696	0.609	0.652	0.756	0.699
UNI	0.696	0.826	0.891	0.891	0.822	0.825
Virchow	0.870	0.826	0.761	0.848	0.867	0.834

Note: A total of 229 liver mass cytology smears were included for a binary classification task: 125 malignant and 104 benign cases.

Supplementary Table 18. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Bile Duct Brushings in the SMH cohort

Task	BileDuct-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.939	0.909	0.879	0.879	0.848	0.891
DINOv2	0.485	0.515	0.606	0.576	0.545	0.545
CTransPath	0.879	0.697	0.636	0.758	0.848	0.764
ResNet-BT	0.606	0.636	0.515	0.576	0.697	0.606
UNI	0.879	0.879	0.879	0.909	0.848	0.879
Virchow	0.970	0.788	0.909	0.879	0.758	0.861

Note: A total of 165 bile duct brushing cytology smears were included for a binary classification task: 67 malignant and 98 benign cases.

Supplementary Table 19. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Bile Cells in the SMH cohort

Task	Bile-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.756	0.852	0.815	0.763	0.842	0.805
DINOv2	0.704	0.681	0.696	0.652	0.737	0.694
CTransPath	0.733	0.770	0.830	0.778	0.835	0.789
ResNet-BT	0.800	0.763	0.770	0.711	0.662	0.741
UNI	0.800	0.763	0.793	0.830	0.797	0.796
Virchow	0.822	0.844	0.741	0.800	0.835	0.808

Note: A total of 673 bile cytology smears were included for a binary classification task: 320 severe atypical hyperplasia and 353 benign cases.

Supplementary Table 20. Performance of Patch-level KNN Classifier for Presence or Absence of Malignancy in Lymph Nodes in the SMH cohort

Task	LN-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.941	0.941	0.942	0.918	0.927	0.934
DINOv2	0.752	0.749	0.776	0.756	0.749	0.756
CTransPath	0.912	0.916	0.911	0.927	0.906	0.914
ResNet-BT	0.853	0.841	0.856	0.831	0.862	0.849
UNI	0.931	0.938	0.925	0.959	0.941	0.939
Virchow	0.938	0.938	0.943	0.946	0.927	0.939

Note: A total of 4386 lymph node cytology smears were included for a binary classification task: 2127 malignant and 2259 benign cases.

Supplementary Table 21. Performance of Patch-level KNN Classifier for Urinary Epithelial Cell Abnormality Detection in the SMH cohort

Task	Urine-AD					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.907	0.921	0.925	0.932	0.892	0.915
DINOv2	0.795	0.823	0.808	0.813	0.806	0.809
CTransPath	0.825	0.862	0.809	0.841	0.827	0.833
ResNet-BT	0.811	0.757	0.757	0.825	0.808	0.792
UNI	0.893	0.902	0.885	0.890	0.888	0.892
Virchow	0.888	0.878	0.893	0.890	0.890	0.888

Note: A total of 2860 urine cytology smears were included for a binary classification task: 1931 normal and 929 abnormal cases.

Supplementary Table 22. Performance of Patch-level KNN Classifier for Effusion Cytology Classification under the TIS System in the SMH cohort

Task	TIS-CLS					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.812	0.863	0.833	0.857	0.800	0.833
DINOv2	0.727	0.730	0.737	0.720	0.717	0.726
CTransPath	0.778	0.771	0.782	0.802	0.745	0.776
ResNet-BT	0.587	0.655	0.666	0.631	0.679	0.644
UNI	0.860	0.816	0.819	0.853	0.852	0.840
Virchow	0.867	0.805	0.771	0.853	0.834	0.826

Note: A total of 1462 effusion cytology smears were included for a five-class classification under the TIS system: 11 TIS-I, 941 TIS-II, 44 TIS-III, 34 TIS-IV, and 432 TIS-V.

Supplementary Table 23. Performance of Patch-level KNN Classifier for Thyroid Cytology Bethesda Category Classification in the SMH cohort

Task	Thyroid-BC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.647	0.655	0.667	0.677	0.692	0.668
DINOv2	0.446	0.429	0.448	0.429	0.469	0.444
CTransPath	0.568	0.555	0.559	0.581	0.571	0.567
ResNet-BT	0.425	0.407	0.455	0.407	0.458	0.430
UNI	0.684	0.662	0.662	0.655	0.676	0.668
Virchow	0.600	0.613	0.601	0.615	0.607	0.607

Note: A total of 4145 thyroid cytology smears were included for a six-class classification based on the Bethesda System: TBS-I (869), TBS-II (1141), TBS-III (632), TBS-IV (62), TBS-V (391), and TBS-VI (1050).

Supplementary Table 24. Performance of Patch-level KNN Classifier for Urinary Epithelial Cell Grading in the SMH cohort

Task	Urine-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.895	0.891	0.888	0.895	0.880	0.890
DINOv2	0.768	0.795	0.800	0.788	0.797	0.790
CTransPath	0.804	0.818	0.800	0.805	0.793	0.804
ResNet-BT	0.768	0.774	0.835	0.730	0.733	0.768
UNI	0.858	0.867	0.870	0.854	0.862	0.862
Virchow	0.863	0.849	0.860	0.863	0.867	0.861

Note: A total of 2846 urine cytology smears were included for a three-class classification: 1935 benign, 120 atypical/low-grade, and 791 high-grade carcinoma cases.

Supplementary Table 25. Performance of Patch-level KNN Classifier for Cervical Cytology Grading in the SMH cohort

Task	Cervix-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.767	0.781	0.763	0.758	0.781	0.770
DINOv2	0.642	0.692	0.704	0.617	0.679	0.667
CTransPath	0.685	0.706	0.704	0.721	0.660	0.695
ResNet-BT	0.635	0.663	0.685	0.627	0.610	0.644
UNI	0.760	0.763	0.775	0.777	0.771	0.769
Virchow	0.748	0.756	0.756	0.731	0.735	0.745

Note: A total of 2400 cervical cytology smears were included for a six-class classification: 1015 normal, 536 LSIL, 428 HSIL, 307 ASC-US, 31 squamous cell carcinoma, and 83 adenocarcinoma.

Supplementary Table 26. Performance of Patch-level KNN Classifier for Cervical Cell HPV Infection Detection in the SMH cohort

Task	Cervix-HPVInf					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.641	0.670	0.677	0.643	0.639	0.654
DINOv2	0.589	0.602	0.584	0.605	0.591	0.594
CTransPath	0.557	0.577	0.593	0.545	0.570	0.569
ResNet-BT	0.518	0.534	0.518	0.550	0.468	0.518
UNI	0.616	0.686	0.650	0.686	0.645	0.657
Virchow	0.607	0.591	0.602	0.552	0.545	0.580

Note: A total of 2200 cervical cytology smears were included for a four-class classification: 1018 HPV-negative, 362 low-risk HPV, 454 high-risk HPV, and 366 both.

Supplementary Table 27. Performance of Patch-level KNN Classifier for Salivary Gland Tumor Subtype Classification in the SMH cohort

Task	Salivary-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.853	0.844	0.688	0.761	0.762	0.782
DINOv2	0.596	0.550	0.550	0.560	0.571	0.566
CTransPath	0.752	0.716	0.844	0.743	0.762	0.763
ResNet-BT	0.670	0.642	0.550	0.587	0.629	0.616
UNI	0.826	0.789	0.844	0.844	0.829	0.826
Virchow	0.761	0.798	0.862	0.771	0.762	0.791

Note: A total of 541 salivary gland cytology smears were included for a three-class classification: 166 Warthin tumors, 279 pleomorphic adenomas, and 96 carcinomas.

Supplementary Table 28. Performance of Patch-level KNN Classifier for Milan System Classification of Salivary Gland Cytology in the SMH cohort

Task	Milan-CLS					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.708	0.833	0.833	0.583	0.792	0.750
DINOv2	0.667	0.875	0.708	0.750	0.750	0.750
CTransPath	0.750	0.667	0.833	0.667	0.833	0.750
ResNet-BT	0.792	0.750	0.750	0.792	0.667	0.750
UNI	0.750	0.750	0.833	0.917	0.625	0.775
Virchow	0.792	0.833	0.708	0.750	0.708	0.758

Note: A total of 120 salivary gland cytology smears were included for a three-class classification: 25 (I-II), 90 (III-IV), and 5 (V-VI).

Supplementary Table 29. Performance of Patch-level KNN Classifier for Liver Lesion Subtype Classification in the SMH cohort

Task	Liver-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.735	0.706	0.706	0.676	0.733	0.711
DINOv2	0.529	0.647	0.588	0.618	0.667	0.610
CTransPath	0.706	0.706	0.588	0.647	0.600	0.649
ResNet-BT	0.618	0.618	0.647	0.618	0.700	0.640
UNI	0.765	0.676	0.794	0.853	0.833	0.784
Virchow	0.765	0.853	0.706	0.735	0.800	0.772

Note: A total of 166 liver cytology smears were included for a four-class classification: 9 HCC, 42 adenocarcinoma, 11 NETs, and 104 benign.

Supplementary Table 30. Performance of Patch-level KNN Classifier for Cytological Grading of Bile Duct Brushings in the SMH cohort

Task	BileDuct-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.585	0.610	0.634	0.659	0.561	0.610
DINOv2	0.439	0.488	0.317	0.390	0.439	0.415
CTransPath	0.561	0.537	0.463	0.537	0.512	0.522
ResNet-BT	0.488	0.463	0.512	0.439	0.463	0.473
UNI	0.707	0.634	0.610	0.659	0.488	0.620
Virchow	0.854	0.659	0.390	0.512	0.634	0.610

Note: A total of 205 bile duct brushing cytology smears were included for a four-class classification: 93 normal, 29 mild, 38 moderate, and 45 severe atypical hyperplasia.

Supplementary Table 31. Performance of Patch-level KNN Classifier for Cytological Grading of Bile Cells in the SMH cohort

Task	Bile-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.620	0.659	0.693	0.603	0.562	0.627
DINOv2	0.497	0.542	0.531	0.514	0.506	0.518
CTransPath	0.581	0.559	0.587	0.609	0.624	0.592
ResNet-BT	0.492	0.603	0.508	0.559	0.590	0.550
UNI	0.547	0.637	0.598	0.665	0.624	0.614
Virchow	0.631	0.592	0.547	0.631	0.646	0.610

Note: A total of 894 bile cytology smears were included for a four-class classification: 361 normal, 47 mild, 166 moderate, and 320 severe atypia/carcinoma.

Supplementary Table 32. Performance of Patch-level KNN Classifier for Parasite Egg and Fungal Detection in Bile in the SMH cohort

Task	Bile-PFD					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.916	0.892	0.916	0.952	0.904	0.916
DINOv2	0.867	0.922	0.922	0.898	0.904	0.902
CTransPath	0.892	0.904	0.922	0.904	0.922	0.908
ResNet-BT	0.910	0.867	0.940	0.922	0.880	0.904
UNI	0.928	0.940	0.916	0.916	0.910	0.922
Virchow	0.946	0.940	0.922	0.916	0.922	0.929

Note: A total of 830 bile cytology smears were included for a three-class classification: 47 parasitic ova, 33 fungal infections, and 750 normal cases.

Supplementary Table 33. Performance of Patch-level KNN Classifier for Primary Origin Classification of Lymph Node Metastatic Cancer in the SMH cohort

Task	LN-Origin					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.689	0.716	0.699	0.639	0.765	0.702
DINOv2	0.464	0.459	0.530	0.492	0.531	0.495
CTransPath	0.546	0.585	0.530	0.601	0.559	0.564
ResNet-BT	0.514	0.372	0.481	0.464	0.486	0.463
UNI	0.667	0.667	0.710	0.683	0.626	0.670
Virchow	0.727	0.617	0.656	0.699	0.659	0.672

Note: A total of 911 lymph node metastatic carcinoma smears were included for a five-class classification: 43 nasopharyngeal, 210 lung, 422 thyroid, 149 breast, and 87 lymphoma.

Supplementary Table 34. Performance of Patch-level KNN Classifier for Lymph Node Cancer Subtype Classification in the SMH cohort

Task	LN-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.780	0.703	0.713	0.689	0.766	0.730
DINOv2	0.526	0.483	0.488	0.526	0.483	0.501
CTransPath	0.574	0.622	0.584	0.589	0.585	0.591
ResNet-BT	0.234	0.311	0.220	0.239	0.327	0.266
UNI	0.775	0.761	0.627	0.708	0.693	0.713
Virchow	0.617	0.632	0.689	0.670	0.702	0.662

Note: A total of 1041 lymph node cancer cytology smears were included for a five-class classification: 87 lymphoma, 40 lymphoepithelial carcinoma, 348 squamous cell carcinoma, 434 adenocarcinoma, and 132 small cell neuroendocrine carcinoma.

Supplementary Table 35. Performance of Patch-level KNN Classifier for Benign Lymph Node Classification in the SMH cohort

Task	LN-Benign					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.913	0.902	0.942	0.919	0.924	0.920
DINOv2	0.867	0.890	0.908	0.896	0.884	0.889
CTransPath	0.908	0.942	0.867	0.919	0.924	0.912
ResNet-BT	0.913	0.821	0.890	0.861	0.901	0.877
UNI	0.896	0.919	0.896	0.936	0.901	0.910
Virchow	0.925	0.850	0.925	0.896	0.919	0.903

Note: A total of 864 benign lymph node cytology smears were included for a three-class classification: 740 common benign lymph node, 20 suppurative lymphadenitis, and 104 granulomatous lesions.

Section II. Performance of Patch-level KNN Classifier in the SCC cohort

This section reports the performance of the patch-level KNN classifier across 22 downstream cytology tasks in the SCC cohort. The tasks comprise 17 binary classification tasks and 5 multiclass classification tasks, covering a variety of anatomical sites and pathological categories. Detailed results highlight the classifier's effectiveness in discriminating cytological features at the patch level within the SCC cohort.

Supplementary Table 36. Performance of Patch-level KNN Classifier for Cervical Cell Abnormality Detection in the SCC cohort

Task	Cervix-AD					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.985	0.983	0.991	0.985	0.991	0.987
DINOv2	0.942	0.919	0.927	0.955	0.912	0.931
CTransPath	0.925	0.904	0.940	0.919	0.923	0.922
ResNet-BT	0.959	0.908	0.915	0.955	0.942	0.936
UNI	0.994	0.991	0.985	0.987	0.983	0.988
Virchow	0.983	0.970	0.968	0.981	0.989	0.978

Note: A total of 2338 cervical cytology smears were included for a binary classification task: 1019 normal and 1319 abnormal cases.

Supplementary Table 37. Performance of Patch-level KNN Classifier for Cervical Cell Candida Infection Detection in the SCC cohort

Task	Cervix-CID					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.992	0.990	0.988	0.986	0.984	0.988
DINOv2	0.924	0.914	0.914	0.912	0.887	0.911
CTransPath	0.957	0.938	0.934	0.955	0.949	0.947
ResNet-BT	0.916	0.924	0.887	0.912	0.889	0.906
UNI	0.981	0.988	0.971	0.984	0.981	0.981
Virchow	0.971	0.961	0.982	0.977	0.957	0.970

Note: A total of 2570 cervical cytology smears were included for a binary classification task: 247 with Candida infection and 2323 normal cases.

Supplementary Table 38. Performance of Patch-level KNN Classifier for Cervical Cell Bacterial Infection Detection in the SCC cohort

Task	Cervix-BID					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.978	0.971	0.986	0.990	0.976	0.980
DINOV2	0.952	0.945	0.952	0.933	0.928	0.942
CTransPath	0.967	0.926	0.962	0.950	0.957	0.952
ResNet-BT	0.959	0.967	0.967	0.967	0.974	0.967
UNI	0.988	0.981	0.976	0.986	0.983	0.983
Virchow	0.962	0.959	0.959	0.957	0.969	0.961

Note: A total of 2090 cervical cytology smears were included for a binary classification task: 872 with bacterial vaginosis and 1218 normal cases.

Supplementary Table 39. Performance of Patch-level KNN Classifier for Pleural Effusion Malignancy Classification in the SCC cohort

Task	Pleural-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.991	0.993	0.983	0.993	0.993	0.991
DINOV2	0.929	0.936	0.910	0.917	0.917	0.922
CTransPath	0.946	0.962	0.953	0.967	0.960	0.957
ResNet-BT	0.929	0.943	0.915	0.927	0.922	0.927
UNI	0.993	0.991	0.998	0.988	0.993	0.992
Virchow	0.981	0.988	0.972	0.988	0.991	0.984

Note: A total of 2114 pleural effusion smears were included for a binary classification task: 1166 malignant and 948 benign cases.

Supplementary Table 40. Performance of Patch-level KNN Classifier for Ascites Malignancy Classification in the SCC cohort

Task	Ascites-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.985	0.985	0.994	0.990	0.986	0.988
DINOV2	0.948	0.940	0.929	0.950	0.924	0.938
CTransPath	0.961	0.975	0.983	0.987	0.967	0.975
ResNet-BT	0.983	0.963	0.963	0.971	0.965	0.969
UNI	0.990	0.990	0.994	0.987	0.983	0.989
Virchow	0.983	0.988	0.992	0.990	0.981	0.987

Note: A total of 2591 ascitic fluid smears were included for a binary classification task: 1283 malignant and 1308 benign cases.

Supplementary Table 41. Performance of Patch-level KNN Classifier for Pericardial Effusion Malignancy Classification in the SCC cohort

Task	Pericardial-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.985	0.970	0.955	0.985	0.985	0.976
DINOV2	0.924	0.864	0.909	0.864	0.939	0.900
CTransPath	0.879	0.955	0.909	0.939	0.894	0.915
ResNet-BT	0.955	0.879	0.909	0.879	0.788	0.882
UNI	0.939	0.970	1.000	1.000	1.000	0.982
Virchow	1.000	0.970	0.924	0.985	0.970	0.970

Note: A total of 330 pericardial effusion smears were included for a binary classification task: 239 malignant and 91 benign cases.

Supplementary Table 42. Performance of Patch-level KNN Classifier for Thyroid Nodule Malignancy Classification in the SCC cohort

Task	Thyroid-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.967	0.976	0.970	0.977	0.972	0.972
DINOV2	0.881	0.883	0.889	0.882	0.886	0.884
CTransPath	0.966	0.972	0.967	0.972	0.967	0.968
ResNet-BT	0.856	0.831	0.837	0.842	0.834	0.840
UNI	0.974	0.971	0.976	0.969	0.978	0.974
Virchow	0.969	0.967	0.953	0.969	0.972	0.966

Note: A total of 4793 thyroid cytology smears were included for a binary classification task: 3212 malignant and 1581 benign cases.

Supplementary Table 43. Performance of Patch-level KNN Classifier for Thyroid Cancer Subtype Classification in the SCC cohort

Task	Thyroid-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.952	0.966	0.963	0.952	0.965	0.960
DINOV2	0.957	0.970	0.959	0.945	0.968	0.960
CTransPath	0.970	0.961	0.961	0.963	0.965	0.964
ResNet-BT	0.961	0.943	0.945	0.959	0.975	0.957
UNI	0.970	0.945	0.968	0.943	0.963	0.958
Virchow	0.975	0.957	0.959	0.947	0.961	0.960

Note: A total of 2186 thyroid cancer cytology smears were included for a binary classification task: 2091 papillary carcinoma and 95 other types.

Supplementary Table 44. Performance of Patch-level KNN Classifier for Urinary Cell Abnormality Detection in the SCC cohort

Task	Urine-AD					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.977	0.983	0.975	0.988	0.986	0.982
DINOV2	0.850	0.884	0.869	0.875	0.849	0.865
CTransPath	0.965	0.956	0.973	0.958	0.977	0.966
ResNet-BT	0.917	0.938	0.944	0.902	0.892	0.919
UNI	0.983	0.983	0.983	0.987	0.983	0.983
Virchow	0.983	0.981	0.987	0.975	0.963	0.978

Note: A total of 2593 urine cytology smears were included for a binary classification task: 1154 normal and 1439 abnormal cases.

Supplementary Table 45. Performance of Patch-level KNN Classifier for Cerebrospinal Fluid Malignancy Classification in the SCC cohort

Task	CSF-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.970	0.990	0.980	0.965	0.960	0.973
DINOV2	0.936	0.926	0.911	0.921	0.920	0.923
CTransPath	0.950	0.950	0.980	0.970	0.935	0.957
ResNet-BT	0.871	0.926	0.926	0.941	0.950	0.923
UNI	0.970	0.980	0.985	0.970	0.970	0.975
Virchow	0.980	0.990	0.980	0.950	1.000	0.980

Note: A total of 1009 cerebrospinal fluid smears were included for a binary classification task: 78 malignant and 931 benign cases.

Supplementary Table 46. Performance of Patch-level KNN Classifier for Malignancy Classification of Bronchial Brushing in the SCC cohort

Task	Bronchial-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.984	0.976	0.982	0.984	0.983	0.982
DINOV2	0.878	0.849	0.854	0.858	0.836	0.855
CTransPath	0.949	0.958	0.960	0.958	0.954	0.956
ResNet-BT	0.958	0.940	0.938	0.932	0.884	0.930
UNI	0.984	0.973	0.978	0.971	0.978	0.977
Virchow	0.971	0.984	0.964	0.974	0.963	0.971

Note: A total of 2736 bronchial brushing cytology smears were included for a binary classification task: 979 malignant and 1757 benign cases.

Supplementary Table 47. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Breast Lesions in the SCC cohort

Task	Breast-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.913	0.870	0.913	0.913	0.955	0.913
DINOV2	0.913	0.870	0.870	0.696	0.864	0.842
CTransPath	0.913	0.870	0.826	0.913	0.773	0.859
ResNet-BT	0.826	0.826	0.957	0.783	0.818	0.842
UNI	0.826	0.870	0.913	0.870	0.864	0.868
Virchow	0.696	0.826	0.957	0.957	0.955	0.878

Note: A total of 114 breast lesion cytology smears were included for a binary classification task: 18 malignant and 96 benign cases.

Supplementary Table 48. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Pancreatic Lesions in the SCC cohort

Task	Pancreas-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.969	0.969	0.954	0.959	0.954	0.961
DINOV2	0.811	0.918	0.898	0.862	0.862	0.870
CTransPath	0.949	0.959	0.939	0.949	0.954	0.950
ResNet-BT	0.801	0.816	0.724	0.765	0.806	0.783
UNI	0.969	0.954	0.969	0.954	0.964	0.962
Virchow	0.944	0.934	0.959	0.964	0.939	0.948

Note: A total of 980 pancreatic lesion smears were included for a binary classification task: 737 malignant and 243 benign cases.

Supplementary Table 49. Performance of Patch-level KNN Classifier for Abnormality Detection in Bile Duct Brushings in the SCC cohort

Task	BileDuct-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	1.000	0.818	0.955	0.864	0.789	0.885
DINOV2	0.864	0.818	0.773	1.000	0.789	0.849
CTransPath	0.955	0.864	0.818	0.818	0.895	0.870
ResNet-BT	0.727	0.773	0.909	0.864	0.947	0.844
UNI	0.955	0.909	0.773	0.864	0.947	0.889
Virchow	0.909	1.000	0.864	0.909	1.000	0.936

Note: A total of 107 bile duct brushing smears were included for a binary classification task: 90 abnormal and 17 normal cases.

Supplementary Table 50. Performance of Patch-level KNN Classifier for Benign and Malignant Classification of Bile Cells in the SCC cohort

Task	Bile-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.947	0.947	0.982	0.965	0.964	0.961
DINOV2	0.877	0.930	0.825	0.842	0.818	0.858
CTransPath	0.807	0.912	0.930	0.912	0.964	0.905
ResNet-BT	0.825	0.842	0.825	0.825	0.636	0.790
UNI	0.930	0.930	0.912	0.965	0.945	0.936
Virchow	0.947	0.912	0.947	0.947	0.873	0.925

Note: A total of 283 bile cytology smears were included for a binary classification task: 209 abnormal and 74 normal cases.

Supplementary Table 51. Performance of Patch-level KNN Classifier for Parasite Egg Detection in Bile in the SCC cohort

Task	Bile-PED					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.885	0.902	0.984	0.984	0.982	0.947
DINOV2	0.967	0.885	0.967	0.934	0.912	0.933
CTransPath	0.902	0.984	0.902	0.951	0.930	0.934
ResNet-BT	0.934	0.918	0.918	0.934	0.965	0.934
UNI	0.951	0.934	0.967	0.984	0.912	0.950
Virchow	0.951	0.967	0.918	0.902	0.982	0.944

Note: A total of 301 bile cytology smears were included for a binary classification task: 20 with parasite eggs and 281 normal cases.

Supplementary Table 52. Performance of Patch-level KNN Classifier for Malignancy Classification in Lymph Nodes in the SCC cohort

Task	LN-BMC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.992	0.986	0.986	0.983	0.980	0.985
DINOV2	0.899	0.919	0.885	0.910	0.929	0.908
CTransPath	0.963	0.944	0.969	0.966	0.972	0.963
ResNet-BT	0.865	0.916	0.916	0.843	0.892	0.886
UNI	0.983	0.986	0.975	0.989	0.989	0.984
Virchow	0.983	0.969	0.989	0.975	0.969	0.977

Note: A total of 1777 lymph node cytology smears were included for a binary classification task: 959 malignant and 818 benign cases.

Supplementary Table 53. Performance of Patch-level KNN Classifier for Cervical Cytology Grading in the SCC cohort

Task	Cervix-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.794	0.811	0.789	0.811	0.813	0.803
DINOv2	0.628	0.680	0.669	0.669	0.669	0.663
CTransPath	0.748	0.733	0.718	0.725	0.708	0.726
ResNet-BT	0.531	0.609	0.574	0.600	0.563	0.575
UNI	0.802	0.806	0.828	0.794	0.841	0.814
Virchow	0.781	0.781	0.791	0.776	0.789	0.784

Note: A total of 2325 cervical cytology smears were included for a multiclass classification task: 1019 normal, 486 LSIL, 432 HSIL, 363 ASC-US, 8 SCC, and 17 adenocarcinoma cases.

Supplementary Table 54. Performance of Patch-level KNN Classifier for Thyroid Cytology Bethesda Category Classification in the SCC cohort

Task	Thyroid-BC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.613	0.638	0.654	0.611	0.629	0.629
DINOv2	0.533	0.538	0.528	0.511	0.524	0.527
CTransPath	0.582	0.595	0.587	0.632	0.615	0.602
ResNet-BT	0.435	0.486	0.434	0.494	0.467	0.463
UNI	0.612	0.638	0.612	0.629	0.625	0.623
Virchow	0.599	0.575	0.581	0.580	0.591	0.585

Note: A total of 5361 thyroid cytology smears were included for a 6-class classification task: 100 Category I, 859 Category II, 1013 Category III, 177 Category IV, 2066 Category V, and 1146 Category VI cases.

Supplementary Table 55. Performance of Patch-level KNN Classifier for Urinary Epithelial Cell Grading in the SCC cohort

Task	Urine-G					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.919	0.921	0.928	0.932	0.923	0.925
DINOv2	0.820	0.841	0.860	0.791	0.842	0.831
CTransPath	0.893	0.888	0.890	0.913	0.890	0.895
ResNet-BT	0.874	0.911	0.841	0.890	0.881	0.880
UNI	0.924	0.901	0.926	0.928	0.909	0.917
Virchow	0.907	0.901	0.893	0.917	0.925	0.909

Note: A total of 2417 urine cytology smears were included for a 3-class classification task: 1158 normal, 182 atypical urothelial cells/low-grade tumors, and 1077 high-grade urothelial carcinoma cases.

Supplementary Table 56. Performance of Patch-level KNN Classifier for Cancer Subtype Classification of Bronchial Brushing Cytology in the SCC cohort

Task	Bronchial-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.760	0.770	0.735	0.724	0.799	0.758
DINOv2	0.429	0.515	0.551	0.444	0.510	0.490
CTransPath	0.582	0.638	0.714	0.694	0.706	0.667
ResNet-BT	0.531	0.464	0.566	0.571	0.521	0.531
UNI	0.750	0.735	0.786	0.776	0.742	0.758
Virchow	0.699	0.653	0.668	0.740	0.732	0.698

Note: A total of 978 malignant bronchial brushing smears were included for a 3-class classification task: 403 squamous cell carcinoma, 383 adenocarcinoma, and 192 small cell carcinoma cases.

Supplementary Table 57. Performance of Patch-level KNN Classifier for Lymph Node Cancer Subtype Classification in the SCC cohort

Task	LN-CSC					
Fold	Fold0	Fold1	Fold2	Fold3	Fold4	Averaged Accuracy
CROWN	0.749	0.831	0.798	0.798	0.787	0.792
DINOv2	0.727	0.672	0.765	0.760	0.765	0.738
CTransPath	0.814	0.809	0.721	0.792	0.803	0.788
ResNet-BT	0.732	0.694	0.760	0.727	0.776	0.738
UNI	0.820	0.787	0.792	0.814	0.814	0.805
Virchow	0.776	0.787	0.716	0.770	0.765	0.763

Note: A total of 915 lymph node cancer smears were included for a 3-class classification task: 230 squamous cell carcinoma, 653 adenocarcinoma, and 32 small cell carcinoma cases.

Section III. Performance of Patch-level Linear Classifier in the SMH cohort

This section reports the performance of the patch-level linear classifier across 34 downstream cytology tasks in the SMH cohort. These tasks comprise 20 binary and 14 multiclass classification problems, covering diverse anatomical sites and pathological conditions. The results provide a comprehensive assessment of the linear classifier's ability to discriminate cytological features at the patch level.

Supplementary Table 58. Performance of Patch-level Linear Classifier for Pleural Effusion Malignancy Classification in the SMH cohort

Task	Pleural-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.948	0.948	0.941	0.965	0.986
DINOv2	0.755	0.742	0.701	0.885	0.840
CTransPath	0.870	0.868	0.845	0.927	0.941
ResNet-BT	0.909	0.908	0.893	0.949	0.968
UNI	0.931	0.931	0.926	0.942	0.980
Virchow	0.934	0.933	0.925	0.955	0.981

Note: A total of 2739 pleural effusion cytology smears were included for a binary classification task: 963 malignant and 1776 benign cases.

Supplementary Table 59. Performance of Patch-level Linear Classifier for Ascites Malignancy Classification in the SMH cohort

Task	Ascites-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.954	0.954	0.937	0.972	0.989
DINOv2	0.802	0.785	0.676	0.929	0.842
CTransPath	0.878	0.873	0.804	0.952	0.945
ResNet-BT	0.924	0.922	0.884	0.964	0.975
UNI	0.948	0.947	0.920	0.976	0.986
Virchow	0.945	0.944	0.918	0.973	0.986

Note: A total of 2866 ascitic fluid cytology smears were included for a binary classification task: 720 malignant and 2146 benign cases.

Supplementary Table 60. Performance of Patch-level Linear Classifier for Pericardial Effusion Malignancy Classification in the SMH cohort

Task	Pericardial-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.903	0.903	0.895	0.928	0.956
DINOv2	0.700	0.663	0.626	0.919	0.791
CTransPath	0.723	0.665	0.636	0.982	0.834
ResNet-BT	0.751	0.750	0.738	0.793	0.873
UNI	0.859	0.858	0.844	0.902	0.941
Virchow	0.904	0.904	0.895	0.928	0.964

Note: A total of 177 pericardial effusion cytology smears were included for a binary classification task: 66 malignant and 111 benign cases.

Supplementary Table 61. Performance of Patch-level Linear Classifier for Thyroid Nodule Malignancy Classification in the SMH cohort

Task	Thyroid-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.978	0.978	0.971	0.989	0.997
DINOv2	0.767	0.751	0.675	0.893	0.799
CTransPath	0.920	0.918	0.880	0.975	0.971
ResNet-BT	0.929	0.926	0.891	0.981	0.982
UNI	0.980	0.980	0.975	0.987	0.997
Virchow	0.972	0.972	0.968	0.978	0.996

Note: A total of 3568 thyroid cytology smears were included for a binary classification task: 1034 malignant and 2534 benign cases.

Supplementary Table 62. Performance of Patch-level Linear Classifier for Thyroid Cancer Subtype Classification in the SMH cohort

Task	Thyroid-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.976	0.975	0.905	0.814	0.992
DINOv2	0.917	0.895	0.638	0.279	0.917
CTransPath	0.911	0.884	0.600	0.200	0.957
ResNet-BT	0.963	0.961	0.860	0.729	0.974
UNI	0.975	0.974	0.911	0.829	0.988
Virchow	0.971	0.970	0.896	0.800	0.990

Note: A total of 1262 thyroid cancer cytology smears were included for a binary classification task: 1122 papillary carcinoma and 140 other carcinoma cases.

Supplementary Table 63. Performance of Patch-level Linear Classifier for Cerebrospinal Fluid Malignancy Classification in the SMH cohort

Task	CSF-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.977	0.977	0.936	0.992	0.994
DINOv2	0.871	0.811	0.500	1.000	0.722
CTransPath	0.871	0.811	0.500	1.000	0.905
ResNet-BT	0.899	0.874	0.643	0.988	0.972
UNI	0.974	0.973	0.928	0.990	0.994
Virchow	0.981	0.981	0.961	0.988	0.996

Note: A total of 573 cerebrospinal fluid cytology smears were included for a binary classification task: 74 malignant and 499 benign cases.

Supplementary Table 64. Performance of Patch-level Linear Classifier for Bronchial Brushing Malignancy Classification in the SMH cohort

Task	Bronchial-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.960	0.957	0.810	0.990	0.974
DINOv2	0.913	0.878	0.513	0.995	0.722
CTransPath	0.915	0.875	0.500	1.000	0.850
ResNet-BT	0.924	0.896	0.560	0.998	0.858
UNI	0.952	0.949	0.782	0.987	0.942
Virchow	0.951	0.946	0.765	0.989	0.959

Note: A total of 2290 bronchial brushing and bronchoalveolar lavage smears were included for a binary classification task: 194 malignant and 2096 benign cases.

Supplementary Table 65. Performance of Patch-level Linear Classifier for Cancer Subtype Classification of Bronchial Brushing Cytology in the SMH cohort

Task	Bronchial-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.941	0.939	0.871	0.971	0.983
DINOv2	0.847	0.776	0.500	1.000	0.655
CTransPath	0.847	0.776	0.500	1.000	0.790
ResNet-BT	0.847	0.776	0.500	1.000	0.803
UNI	0.921	0.906	0.805	0.971	0.970
Virchow	0.911	0.912	0.856	0.936	0.971

Note: A total of 202 malignant bronchial cytology smears were included for a binary classification task: 171 non-small cell carcinoma and 31 small cell carcinoma cases.

Supplementary Table 66. Performance of Patch-level Linear Classifier for Cervical Cell Abnormality Detection in the SMH cohort

Task	Cervix-AD				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.970	0.970	0.969	0.972	0.995
DINOv2	0.898	0.898	0.902	0.878	0.948
CTransPath	0.923	0.923	0.924	0.919	0.977
ResNet-BT	0.947	0.947	0.948	0.940	0.983
UNI	0.967	0.967	0.967	0.966	0.995
Virchow	0.967	0.968	0.968	0.967	0.993

Note: A total of 2398 cervical cytology smears were included for a binary classification task: 1013 normal and 1385 abnormal cases.

Supplementary Table 67. Performance of Patch-level Linear Classifier for Cervical Cell Bacterial Infection Detection in the SMH cohort

Task	Cervix-BID				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.966	0.966	0.964	0.953	0.994
DINOv2	0.758	0.752	0.737	0.621	0.832
CTransPath	0.835	0.832	0.817	0.718	0.924
ResNet-BT	0.878	0.878	0.878	0.876	0.952
UNI	0.961	0.961	0.959	0.946	0.990
Virchow	0.927	0.927	0.923	0.902	0.971

Note: A total of 1720 cervical cytology smears were included for a binary classification task: 1018 without bacterial infection and 702 with bacterial vaginosis.

Supplementary Table 68. Performance of Patch-level Linear Classifier for Cervical Cell Candida Infection Detection in the SMH cohort

Task	Cervix-CID				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.966	0.966	0.964	0.957	0.993
DINOv2	0.868	0.868	0.864	0.834	0.943
CTransPath	0.913	0.913	0.909	0.879	0.973
ResNet-BT	0.902	0.901	0.900	0.890	0.972
UNI	0.962	0.962	0.962	0.957	0.993
Virchow	0.937	0.937	0.935	0.924	0.987

Note: A total of 1770 cervical cytology smears were included for a binary classification task: 1018 without fungal infection and 752 with Candida infection.

Supplementary Table 69. Performance of Patch-level Linear Classifier for Salivary Gland Tumor Malignancy Classification in the SMH cohort

Task	Salivary-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.869	0.855	0.726	0.963	0.896
DINOv2	0.802	0.713	0.500	1.000	0.609
CTransPath	0.802	0.713	0.500	1.000	0.792
ResNet-BT	0.811	0.760	0.572	0.968	0.810
UNI	0.878	0.871	0.764	0.953	0.908
Virchow	0.890	0.880	0.764	0.973	0.898

Note: A total of 746 salivary gland cytology smears were included for a binary classification task: 148 malignant and 598 benign cases.

Supplementary Table 70. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Vocal Cord Lesions in the SMH cohort

Task	VocalCord-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.823	0.795	0.739	0.978	0.938
DINOv2	0.677	0.560	0.489	0.978	0.598
CTransPath	0.694	0.568	0.500	1.000	0.556
ResNet-BT	0.537	0.444	0.547	0.494	0.615
UNI	0.724	0.662	0.608	0.933	0.813
Virchow	0.740	0.701	0.646	0.908	0.790

Note: A total of 62 vocal cord mass cytology smears were included for a binary classification task: 19 malignant and 43 benign cases.

Supplementary Table 71. Performance of Patch-level Linear Classifier for Breast Lesion Malignancy Classification in the SMH cohort

Task	Breast-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.848	0.849	0.828	0.871	0.925
DINOv2	0.725	0.705	0.604	0.871	0.764
CTransPath	0.725	0.610	0.500	1.000	0.751
ResNet-BT	0.778	0.730	0.643	0.935	0.838
UNI	0.818	0.804	0.723	0.936	0.899
Virchow	0.854	0.841	0.768	0.960	0.875

Note: A total of 171 breast mass cytology smears were included for a binary classification task: 47 malignant and 124 benign cases.

Supplementary Table 72. Performance of Patch-level Linear Classifier for Benign and Malignant Pancreatic Lesions in the SMH cohort

Task	Pancreas-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.856	0.846	0.766	0.573	0.909
DINOv2	0.721	0.613	0.493	0.000	0.592
CTransPath	0.731	0.618	0.500	0.000	0.759
ResNet-BT	0.726	0.702	0.614	0.371	0.792
UNI	0.861	0.852	0.787	0.629	0.901
Virchow	0.831	0.827	0.773	0.649	0.915

Note: A total of 201 pancreatic mass cytology smears were included for a binary classification task: 147 malignant and 54 benign cases.

Supplementary Table 73. Performance of Patch-level Linear Classifier for Benign and Malignant Liver Lesions in the SMH cohort

Task	Liver-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.877	0.877	0.875	0.846	0.932
DINOv2	0.633	0.629	0.638	0.693	0.724
CTransPath	0.729	0.709	0.707	0.471	0.829
ResNet-BT	0.689	0.670	0.693	0.714	0.822
UNI	0.838	0.838	0.836	0.808	0.910
Virchow	0.869	0.868	0.866	0.828	0.938

Note: A total of 229 liver mass cytology smears were included for a binary classification task: 125 malignant and 104 benign cases.

Supplementary Table 74. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Bile Duct Brushings in the SMH cohort

Task	BileDuct-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.873	0.870	0.864	0.919	0.942
DINOv2	0.539	0.438	0.478	0.819	0.506
CTransPath	0.630	0.512	0.543	1.000	0.900
ResNet-BT	0.727	0.716	0.761	0.580	0.852
UNI	0.836	0.835	0.830	0.866	0.928
Virchow	0.836	0.833	0.840	0.837	0.937

Note: A total of 165 bile duct brushing cytology smears were included for a binary classification task: 67 malignant and 98 benign cases.

Supplementary Table 75. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Bile Cells in the SMH cohort

Task	Bile-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.857	0.857	0.858	0.844	0.941
DINOv2	0.704	0.701	0.700	0.790	0.803
CTransPath	0.787	0.787	0.790	0.736	0.848
ResNet-BT	0.792	0.790	0.795	0.736	0.868
UNI	0.841	0.841	0.842	0.830	0.933
Virchow	0.848	0.848	0.851	0.793	0.924

Note: A total of 673 bile cytology smears were included for a binary classification task: 320 severe atypical hyperplasia and 353 benign cases.

Supplementary Table 76. Performance of Patch-level Linear Classifier for Presence or Absence of Malignancy in Lymph Nodes in the SMH cohort

Task	LN-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.952	0.952	0.952	0.952	0.989
DINOv2	0.736	0.734	0.734	0.803	0.835
CTransPath	0.914	0.914	0.914	0.906	0.968
ResNet-BT	0.936	0.936	0.936	0.941	0.980
UNI	0.956	0.956	0.956	0.951	0.991
Virchow	0.951	0.951	0.950	0.952	0.990

Note: A total of 4386 lymph node cytology smears were included for a binary classification task: 2127 malignant and 2259 benign cases.

Supplementary Table 77. Performance of Patch-level Linear Classifier for Urinary Epithelial Cell Abnormality Detection in the SMH cohort

Task	Urine-AD				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.934	0.933	0.918	0.874	0.981
DINOv2	0.771	0.749	0.683	0.432	0.852
CTransPath	0.834	0.829	0.790	0.665	0.913
ResNet-BT	0.841	0.836	0.801	0.685	0.932
UNI	0.921	0.921	0.909	0.874	0.973
Virchow	0.920	0.920	0.910	0.883	0.973

Note: A total of 2860 urine cytology smears were included for a binary classification task: 1931 normal and 929 abnormal cases.

Supplementary Table 78. Performance of Patch-level Linear Classifier for Effusion Cytology Classification under the TIS System in the SMH cohort

Task	TIS-CLS				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.853	0.830	0.364	0.937	0.892
DINOv2	0.720	0.683	0.277	0.866	0.764
CTransPath	0.777	0.748	0.310	0.895	0.781
ResNet-BT	0.832	0.807	0.350	0.926	0.841
UNI	0.858	0.833	0.368	0.936	0.876
Virchow	0.858	0.834	0.369	0.936	0.884

Note: A total of 1462 effusion cytology smears were included for a five-class classification under the TIS system: 11 TIS-I, 941 TIS-II, 44 TIS-III, 34 TIS-IV, and 432 TIS-V.

Supplementary Table 79. Performance of Patch-level Linear Classifier for Thyroid Cytology Bethesda Category Classification in the SMH cohort

Task	Thyroid-BC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.710	0.702	0.577	0.939	0.931
DINOv2	0.420	0.375	0.297	0.873	0.700
CTransPath	0.598	0.562	0.431	0.913	0.845
ResNet-BT	0.624	0.593	0.468	0.920	0.882
UNI	0.717	0.703	0.575	0.941	0.930
Virchow	0.701	0.689	0.561	0.938	0.918

Note: A total of 4145 thyroid cytology smears were included for a six-class classification based on the Bethesda System: TBS-I (869), TBS-II (1141), TBS-III (632), TBS-IV (62), TBS-V (391), and TBS-VI (1050).

Supplementary Table 80. Performance of Patch-level Linear Classifier for Urinary Epithelial Cell Grading in the SMH cohort

Task	Urine-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.911	0.906	0.723	0.936	0.963
DINOv2	0.753	0.719	0.463	0.784	0.812
CTransPath	0.810	0.786	0.518	0.837	0.873
ResNet-BT	0.827	0.809	0.557	0.860	0.902
UNI	0.901	0.892	0.671	0.927	0.953
Virchow	0.890	0.880	0.655	0.918	0.950

Note: A total of 2846 urine cytology smears were included for a three-class classification: 1935 benign, 120 atypical/low-grade, and 791 high-grade carcinoma cases.

Supplementary Table 81. Performance of Patch-level Linear Classifier for Cervical Cytology Grading in the SMH cohort

Task	Cervix-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.811	0.799	0.595	0.960	0.953
DINOv2	0.673	0.624	0.389	0.924	0.822
CTransPath	0.704	0.641	0.405	0.930	0.874
ResNet-BT	0.741	0.722	0.509	0.943	0.912
UNI	0.807	0.794	0.603	0.959	0.949
Virchow	0.789	0.782	0.573	0.955	0.939

Note: A total of 2400 cervical cytology smears were included for a six-class classification: 1015 normal, 536 LSIL, 428 HSIL, 307 ASC-US, 31 squamous cell carcinoma, and 83 adenocarcinoma.

Supplementary Table 82. Performance of Patch-level Linear Classifier for Cervical Cell HPV Infection Detection in the SMH cohort

Task	Cervix-HPVInf				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.691	0.687	0.613	0.895	0.869
DINOv2	0.595	0.536	0.451	0.845	0.782
CTransPath	0.566	0.519	0.425	0.828	0.800
ResNet-BT	0.571	0.538	0.485	0.856	0.803
UNI	0.668	0.658	0.578	0.885	0.862
Virchow	0.610	0.582	0.505	0.862	0.826

Note: A total of 2200 cervical cytology smears were included for a four-class classification: 1018 HPV-negative, 362 low-risk HPV, 454 high-risk HPV, and 366 both.

Supplementary Table 83. Performance of Patch-level Linear Classifier for Salivary Gland Tumor Subtype Classification in the SMH cohort

Task	Salivary-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.852	0.849	0.806	0.919	0.946
DINOv2	0.551	0.491	0.423	0.730	0.657
CTransPath	0.629	0.542	0.456	0.747	0.871
ResNet-BT	0.737	0.694	0.624	0.846	0.897
UNI	0.861	0.856	0.808	0.921	0.955
Virchow	0.843	0.835	0.777	0.912	0.941

Note: A total of 541 salivary gland cytology smears were included for a three-class classification: 166 Warthin tumors, 279 pleomorphic adenomas, and 96 carcinomas.

Supplementary Table 84. Performance of Patch-level Linear Classifier for Milan System Classification of Salivary Gland Cytology in the SMH cohort

Task	Milan-CLS				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.750	0.692	0.381	0.705	0.522
DINOv2	0.742	0.639	0.330	0.663	0.552
CTransPath	0.750	0.643	0.333	0.667	0.698
ResNet-BT	0.750	0.643	0.333	0.667	0.610
UNI	0.808	0.774	0.465	0.775	0.670
Virchow	0.750	0.707	0.420	0.735	0.634

Note: A total of 120 salivary gland cytology smears were included for a three-class classification: 25 (I-II), 90 (III-IV), and 5 (V-VI).

Supplementary Table 85. Performance of Patch-level Linear Classifier for Liver Lesion Subtype Classification in the SMH cohort

Task	Liver-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.699	0.645	0.355	0.834	0.684
DINOv2	0.603	0.526	0.261	0.779	0.604
CTransPath	0.627	0.483	0.250	0.750	0.585
ResNet-BT	0.663	0.561	0.292	0.781	0.645
UNI	0.723	0.687	0.406	0.876	0.778
Virchow	0.747	0.706	0.414	0.881	0.811

Note: A total of 166 liver cytology smears were included for a four-class classification: 9 HCC, 42 adenocarcinoma, 11 NETs, and 104 benign.

Supplementary Table 86. Performance of Patch-level Linear Classifier for Cytological Grading of Bile Duct Brushings in the SMH cohort

Task	BileDuct-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.620	0.588	0.501	0.864	0.787
DINOv2	0.444	0.317	0.263	0.759	0.595
CTransPath	0.454	0.283	0.250	0.750	0.737
ResNet-BT	0.493	0.439	0.390	0.824	0.743
UNI	0.590	0.564	0.490	0.853	0.758
Virchow	0.590	0.555	0.470	0.849	0.756

Note: A total of 205 bile duct brushing cytology smears were included for a four-class classification: 93 normal, 29 mild, 38 moderate, and 45 severe atypical hyperplasia.

Supplementary Table 87. Performance of Patch-level Linear Classifier for Cytological Grading of Bile Cells in the SMH cohort

Task	Bile-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.624	0.591	0.447	0.858	0.784
DINOv2	0.547	0.488	0.363	0.819	0.695
CTransPath	0.596	0.519	0.392	0.839	0.684
ResNet-BT	0.551	0.497	0.375	0.827	0.738
UNI	0.638	0.601	0.446	0.864	0.773
Virchow	0.613	0.578	0.433	0.854	0.761

Note: A total of 894 bile cytology smears were included for a four-class classification: 361 normal, 47 mild, 166 moderate, and 320 severe atypia/carcinoma.

Supplementary Table 88. Performance of Patch-level Linear Classifier for Parasite Egg and Fungal Detection in Bile in the SMH cohort

Task	Bile-PFD				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.937	0.924	0.579	0.795	0.927
DINOv2	0.900	0.856	0.332	0.665	0.755
CTransPath	0.904	0.858	0.333	0.667	0.784
ResNet-BT	0.919	0.890	0.437	0.728	0.864
UNI	0.952	0.945	0.672	0.841	0.933
Virchow	0.941	0.931	0.610	0.826	0.921

Note: A total of 830 bile cytology smears were included for a three-class classification: 47 parasitic ova, 33 fungal infections, and 750 normal cases.

Supplementary Table 89. Performance of Patch-level Linear Classifier for Primary Origin Classification of Lymph Node Metastatic Cancer in the SMH cohort

Task	LN-Origin				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.768	0.757	0.639	0.933	0.924
DINOv2	0.501	0.425	0.276	0.833	0.668
CTransPath	0.527	0.409	0.311	0.826	0.798
ResNet-BT	0.631	0.607	0.493	0.888	0.841
UNI	0.716	0.702	0.613	0.916	0.911
Virchow	0.707	0.691	0.577	0.915	0.900

Note: A total of 911 lymph node metastatic carcinoma smears were included for a five-class classification: 43 nasopharyngeal, 210 lung, 422 thyroid, 149 breast, and 87 lymphoma.

Supplementary Table 90. Performance of Patch-level Linear Classifier for Lymph Node Cancer Subtype Classification in the SMH cohort

Task	LN-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.775	0.763	0.635	0.933	0.921
DINOv2	0.494	0.423	0.273	0.836	0.668
CTransPath	0.579	0.537	0.383	0.865	0.798
ResNet-BT	0.605	0.561	0.454	0.878	0.860
UNI	0.786	0.776	0.647	0.936	0.928
Virchow	0.751	0.740	0.604	0.927	0.910

Note: A total of 1041 lymph node cancer cytology smears were included for a five-class classification: 87 lymphoma, 40 lymphoepithelial carcinoma, 348 squamous cell carcinoma, 434 adenocarcinoma, and 132 small cell neuroendocrine carcinoma.

Supplementary Table 91. Performance of Patch-level Linear Classifier for Benign Lymph Node Classification in the SMH cohort

Task	LN-Benign				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.944	0.938	0.693	0.891	0.950
DINOv2	0.862	0.826	0.391	0.716	0.774
CTransPath	0.859	0.796	0.340	0.672	0.804
ResNet-BT	0.915	0.897	0.510	0.833	0.906
UNI	0.929	0.921	0.652	0.866	0.950
Virchow	0.942	0.936	0.681	0.888	0.953

Note: A total of 864 benign lymph node cytology smears were included for a three-class classification: 740 common benign lymph node, 20 suppurative lymphadenitis, and 104 granulomatous lesions.

Section IV. Performance of Patch-level Linear Classifier in the SCC cohort

This section presents the performance of the patch-level linear classifier across 22 downstream cytology tasks in the SCC cohort. The tasks include 17 binary classification tasks and 5 multiclass classification tasks, spanning various anatomical sites and pathological categories. Detailed results demonstrate the linear classifier's effectiveness in patch-level cytological feature discrimination within the SCC cohort.

Supplementary Table 92. Performance of Patch-level Linear Classifier for Cervical Cell Normality and Abnormality Classification in the SCC cohort

Task	Cervix-AD				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.991	0.991	0.991	0.994	1.000
DINOv2	0.880	0.880	0.879	0.887	0.955
CTransPath	0.972	0.972	0.970	0.983	0.993
ResNet-BT	0.977	0.977	0.975	0.992	0.996
UNI	0.990	0.990	0.989	0.995	0.999
Virchow	0.985	0.985	0.984	0.992	0.998

Note: A total of 2338 cervical cytology smears were included for a binary classification task: 1019 normal and 1319 abnormal cases.

Supplementary Table 93. Performance of Patch-level Linear Classifier for Cervical Cell Candida Infection Detection in the SCC cohort

Task	Cervix-CID				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.993	0.993	0.973	0.947	0.996
DINOv2	0.912	0.889	0.602	0.219	0.902
CTransPath	0.932	0.915	0.648	0.296	0.970
ResNet-BT	0.965	0.962	0.854	0.717	0.981
UNI	0.988	0.988	0.952	0.907	0.996
Virchow	0.984	0.983	0.938	0.883	0.993

Note: A total of 2570 cervical cytology smears were included for a binary classification task: 247 with Candida infection and 2323 normal cases.

Supplementary Table 94. Performance of Patch-level Linear Classifier for Cervical Cell Bacterial Infection Detection in the SCC cohort

Task	Cervix-BID				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.990	0.990	0.990	0.985	0.999
DINOv2	0.922	0.922	0.918	0.891	0.974
CTransPath	0.969	0.969	0.967	0.950	0.995
ResNet-BT	0.976	0.976	0.977	0.979	0.995
UNI	0.991	0.991	0.990	0.986	0.998
Virchow	0.984	0.984	0.983	0.977	0.997

Note: A total of 2090 cervical cytology smears were included for a binary classification task: 872 with bacterial vaginosis and 1218 normal cases.

Supplementary Table 95. Performance of Patch-level Linear Classifier for Pleural Effusion Malignancy Classification in the SCC cohort

Task	Pleural-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.991	0.991	0.990	0.988	0.999
DINOv2	0.903	0.903	0.905	0.924	0.961
CTransPath	0.970	0.970	0.970	0.975	0.996
ResNet-BT	0.987	0.987	0.987	0.987	0.998
UNI	0.996	0.996	0.995	0.992	0.999
Virchow	0.993	0.993	0.993	0.992	0.999

Note: A total of 2114 pleural effusion smears were included for a binary classification task: 1166 malignant and 948 benign cases.

Supplementary Table 96. Performance of Patch-level Linear Classifier for Ascites Malignancy Classification in the SCC cohort

Task	Ascites-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.993	0.993	0.993	0.995	0.999
DINOv2	0.930	0.930	0.930	0.955	0.976
CTransPath	0.978	0.978	0.978	0.991	0.998
ResNet-BT	0.985	0.985	0.985	0.993	0.998
UNI	0.994	0.994	0.994	0.995	0.999
Virchow	0.992	0.992	0.992	0.993	0.999

Note: A total of 2591 ascitic fluid smears were included for a binary classification task: 1283 malignant and 1308 benign cases.

Supplementary Table 97. Performance of Patch-level Linear Classifier for Pericardial Effusion Malignancy Classification in the SCC cohort

Task	Pericardial-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.973	0.973	0.971	0.968	0.997
DINOv2	0.861	0.855	0.800	0.668	0.938
CTransPath	0.876	0.863	0.777	0.559	0.982
ResNet-BT	0.952	0.951	0.932	0.890	0.989
UNI	0.985	0.985	0.986	0.989	0.996
Virchow	0.967	0.967	0.960	0.945	0.995

Note: A total of 330 pericardial effusion smears were included for a binary classification task: 239 malignant and 91 benign cases.

Supplementary Table 98. Performance of Patch-level Linear Classifier for Thyroid Nodule Malignancy Classification in the SCC cohort

Task	Thyroid-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.980	0.980	0.975	0.960	0.997
DINOv2	0.879	0.878	0.864	0.820	0.936
CTransPath	0.970	0.970	0.963	0.940	0.988
ResNet-BT	0.974	0.974	0.966	0.942	0.992
UNI	0.980	0.980	0.974	0.958	0.996
Virchow	0.979	0.978	0.973	0.957	0.996

Note: A total of 4793 thyroid cytology smears were included for a binary classification task: 3212 malignant and 1581 benign cases.

Supplementary Table 99. Performance of Patch-level Linear Classifier for Thyroid Cancer Subtype Classification in the SCC cohort

Task	Thyroid-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.981	0.979	0.804	0.611	0.983
DINOv2	0.958	0.942	0.541	0.084	0.855
CTransPath	0.957	0.935	0.500	0.000	0.910
ResNet-BT	0.973	0.969	0.744	0.495	0.960
UNI	0.981	0.980	0.824	0.653	0.983
Virchow	0.976	0.973	0.751	0.505	0.979

Note: A total of 2186 thyroid cancer cytology smears were included for a binary classification task: 2091 papillary carcinoma and 95 other types.

Supplementary Table 100. Performance of Patch-level Linear Classifier for Urinary Cell Abnormality Detection in the SCC cohort

Task	Urine-AD				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.987	0.987	0.986	0.990	0.998
DINOv2	0.856	0.857	0.860	0.827	0.933
CTransPath	0.967	0.967	0.967	0.967	0.992
ResNet-BT	0.981	0.981	0.981	0.986	0.995
UNI	0.986	0.986	0.986	0.990	0.997
Virchow	0.983	0.983	0.982	0.985	0.997

Note: A total of 2593 urine cytology smears were included for a binary classification task: 1154 normal and 1439 abnormal cases.

Supplementary Table 101. Performance of Patch-level Linear Classifier for Cerebrospinal Fluid Malignancy Classification in the SCC cohort

Task	CSF-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.982	0.982	0.943	0.989	0.997
DINOv2	0.935	0.915	0.600	0.996	0.846
CTransPath	0.923	0.886	0.500	1.000	0.956
ResNet-BT	0.973	0.970	0.843	0.997	0.992
UNI	0.982	0.982	0.948	0.988	0.997
Virchow	0.990	0.990	0.959	0.996	0.999

Note: A total of 1009 cerebrospinal fluid smears were included for a binary classification task: 78 malignant and 931 benign cases.

Supplementary Table 102. Performance of Patch-level Linear Classifier for Malignancy Classification of Bronchial Brushing in the SCC cohort

Task	Bronchial-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.982	0.982	0.982	0.983	0.997
DINOv2	0.822	0.815	0.780	0.929	0.896
CTransPath	0.969	0.968	0.963	0.983	0.993
ResNet-BT	0.970	0.970	0.967	0.977	0.992
UNI	0.977	0.977	0.978	0.976	0.997
Virchow	0.981	0.981	0.980	0.982	0.996

Note: A total of 2736 bronchial brushing cytology smears were included for a binary classification task: 979 malignant and 1757 benign cases.

Supplementary Table 103. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Breast Lesions in the SCC cohort

Task	Breast-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.921	0.911	0.789	0.979	0.968
DINOv2	0.860	0.821	0.598	0.979	0.803
CTransPath	0.842	0.770	0.500	1.000	0.864
ResNet-BT	0.886	0.849	0.678	0.989	0.927
UNI	0.894	0.881	0.714	0.979	0.934
Virchow	0.913	0.913	0.868	0.937	0.953

Note: A total of 114 breast lesion cytology smears were included for a binary classification task: 18 malignant and 96 benign cases.

Supplementary Table 104. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Pancreatic Lesions in the SCC cohort

Task	Pancreas-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.964	0.964	0.941	0.893	0.979
DINOv2	0.854	0.847	0.776	0.621	0.870
CTransPath	0.916	0.912	0.844	0.700	0.955
ResNet-BT	0.956	0.956	0.929	0.877	0.964
UNI	0.960	0.960	0.934	0.881	0.980
Virchow	0.962	0.962	0.934	0.877	0.978

Note: A total of 980 pancreatic lesion smears were included for a binary classification task: 737 malignant and 243 benign cases.

Supplementary Table 105. Performance of Patch-level Linear Classifier for Abnormality Detection in Bile Duct Brushings in the SCC cohort

Task	BileDuct-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.935	0.922	0.800	1.000	0.970
DINOv2	0.832	0.776	0.514	0.978	0.720
CTransPath	0.842	0.769	0.500	1.000	0.902
ResNet-BT	0.860	0.814	0.589	0.978	0.953
UNI	0.953	0.950	0.878	0.989	0.967
Virchow	0.962	0.962	0.931	0.978	0.996

Note: A total of 107 bile duct brushing smears were included for a binary classification task: 90 abnormal and 17 normal cases.

Supplementary Table 106. Performance of Patch-level Linear Classifier for Benign and Malignant Classification of Bile Cells in the SCC cohort

Task	Bile-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.958	0.958	0.945	0.971	0.980
DINOv2	0.816	0.798	0.716	0.923	0.895
CTransPath	0.739	0.627	0.500	1.000	0.968
ResNet-BT	0.911	0.910	0.862	0.966	0.962
UNI	0.933	0.932	0.902	0.966	0.976
Virchow	0.940	0.940	0.916	0.967	0.975

Note: A total of 283 bile cytology smears were included for a binary classification task: 209 abnormal and 74 normal cases.

Supplementary Table 107. Performance of Patch-level Linear Classifier for Parasite Egg Detection in Bile in the SCC cohort

Task	Bile-PED				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.957	0.945	0.675	0.350	0.920
DINOv2	0.934	0.901	0.500	0.000	0.772
CTransPath	0.934	0.901	0.500	0.000	0.646
ResNet-BT	0.934	0.901	0.500	0.000	0.792
UNI	0.950	0.936	0.648	0.300	0.942
Virchow	0.963	0.960	0.795	0.600	0.930

Note: A total of 301 bile cytology smears were included for a binary classification task: 20 with parasite eggs and 281 normal cases.

Supplementary Table 108. Performance of Patch-level Linear Classifier for Presence or Absence of Malignancy in Lymph Nodes in the SCC cohort

Task	LN-BMC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.989	0.989	0.989	0.985	1.000
DINOv2	0.891	0.891	0.893	0.913	0.957
CTransPath	0.974	0.974	0.973	0.968	0.996
ResNet-BT	0.981	0.981	0.981	0.979	0.997
UNI	0.988	0.988	0.987	0.983	0.999
Virchow	0.986	0.986	0.986	0.985	0.999

Note: A total of 1777 lymph node cytology smears were included for a binary classification task: 959 malignant and 818 benign cases.

Supplementary Table 109. Performance of Patch-level Linear Classifier for Cervical Cytology Grading in the SCC cohort

Task	Cervix-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.847	0.841	0.536	0.968	0.953
DINOv2	0.640	0.587	0.349	0.911	0.786
CTransPath	0.768	0.754	0.463	0.950	0.869
ResNet-BT	0.796	0.770	0.489	0.957	0.922
UNI	0.842	0.834	0.530	0.967	0.952
Virchow	0.838	0.832	0.539	0.966	0.950

Note: A total of 2325 cervical cytology smears were included for a multiclass classification task: 1019 normal, 486 LSIL, 432 HSIL, 363 ASC-US, 8 SCC, and 17 adenocarcinoma cases.

Supplementary Table 110. Performance of Patch-level Linear Classifier for Thyroid Cytology Bethesda Category Classification in the SCC cohort

Task	Thyroid-BC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.659	0.649	0.626	0.919	0.909
DINOv2	0.510	0.480	0.456	0.886	0.812
CTransPath	0.611	0.584	0.495	0.904	0.873
ResNet-BT	0.588	0.574	0.540	0.906	0.885
UNI	0.645	0.637	0.623	0.916	0.903
Virchow	0.634	0.619	0.596	0.914	0.898

Note: A total of 5361 thyroid cytology smears were included for a 6-class classification task: 100 Category I, 859 Category II, 1013 Category III, 177 Category IV, 2066 Category V, and 1146 Category VI cases.

Supplementary Table 111. Performance of Patch-level Linear Classifier for Urinary Epithelial Cell Grading in the SCC cohort

Task	Urine-G				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.937	0.933	0.818	0.965	0.974
DINOv2	0.815	0.784	0.588	0.886	0.897
CTransPath	0.891	0.857	0.643	0.933	0.939
ResNet-BT	0.913	0.892	0.698	0.948	0.949
UNI	0.929	0.928	0.821	0.963	0.970
Virchow	0.921	0.916	0.788	0.958	0.971

Note: A total of 2417 urine cytology smears were included for a 3-class classification task: 1158 normal, 182 atypical urothelial cells/low-grade tumors, and 1077 high-grade urothelial carcinoma cases.

Supplementary Table 112. Performance of Patch-level Linear Classifier for Cancer Subtype Classification of Bronchial Brushing and Bronchoalveolar Lavage Cytology in the SCC cohort

Task	Bronchial-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.829	0.829	0.840	0.908	0.947
DINOv2	0.468	0.418	0.404	0.709	0.646
CTransPath	0.701	0.694	0.658	0.836	0.850
ResNet-BT	0.719	0.716	0.701	0.847	0.883
UNI	0.830	0.830	0.833	0.909	0.957
Virchow	0.825	0.825	0.830	0.907	0.941

Note: A total of 978 malignant bronchial brushing smears were included for a 3-class classification task: 403 squamous cell carcinoma, 383 adenocarcinoma, and 192 small cell carcinoma cases.

Supplementary Table 113. Performance of Patch-level Linear Classifier for Lymph Node Cancer Subtype Classification in the SCC cohort

Task	LN-CSC				
Metrics	Accuracy	F1	Sensitivity	Specificity	AUC
CROWN	0.860	0.854	0.715	0.869	0.934
DINOv2	0.719	0.610	0.342	0.675	0.702
CTransPath	0.737	0.645	0.364	0.695	0.819
ResNet-BT	0.811	0.782	0.533	0.811	0.880
UNI	0.875	0.870	0.727	0.880	0.951
Virchow	0.855	0.849	0.740	0.869	0.929

Note: A total of 915 lymph node cancer smears were included for a 3-class classification task: 230 squamous cell carcinoma, 653 adenocarcinoma, and 32 small cell carcinoma cases.

Section V. Performance of Patch-level Linear Classifier in the Public Dataset

This section reports the performance of the patch-level linear classifier on 12 downstream cytology tasks from public datasets. These include 9 routine anatomical site tasks (effusion, cervix, and thyroid) with 2 binary classification and 7 multiclass classification tasks, as well as 3 out-of-distribution (OOD) tasks involving blood smear samples, comprising 1 binary and 2 multiclass classification tasks. The evaluation highlights the classifier’s generalizability across both common and challenging cytological domains.

Supplementary Table 114. Performance of Patch-level Linear Classifier for Cervical Cell Type Classification in the Original SipakMed Dataset

Task	Cervix 1				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.908	0.908	0.989	0.919	0.976
DINOv2	0.514	0.482	0.813	0.444	0.874
CTransPath	0.725	0.724	0.940	0.696	0.925
ResNet-BT	0.736	0.733	0.944	0.759	0.930
UNI	0.911	0.911	0.990	0.921	0.977
Virchow	0.847	0.848	0.979	0.858	0.960

Note: A total of 966 cervical cytology images were included for a 5-class classification task: 223 dyskeratotic, 238 koilocytotic, 271 metaplastic, 108 parabasal, and 126 superficial-intermediate cells. Dataset available at:

<https://www.kaggle.com/datasets/prahladmehandiratta/cervical-cancer-largest-dataset-sipakmed>

Supplementary Table 115. Performance of Patch-level Linear Classifier for Cervical Cell Type Classification in the Cropped SipakMed Dataset

Task	Cervix 2				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.966	0.966	0.998	0.968	0.991
DINOv2	0.744	0.743	0.941	0.768	0.934
CTransPath	0.855	0.856	0.975	0.861	0.963
ResNet-BT	0.873	0.873	0.983	0.867	0.967
UNI	0.955	0.955	0.997	0.956	0.988
Virchow	0.936	0.936	0.994	0.939	0.984

Note: A total of 3,532 cropped cervical cytology images were included for a 5-class classification task: 296 dyskeratotic, 825 koilocytotic, 793 metaplastic, 787 parabasal, and 831 superficial-intermediate cells. Dataset available at:

<https://www.kaggle.com/datasets/prahladmehandiratta/cervical-cancer-largest-dataset-sipakmed>

Supplementary Table 116. Performance of Patch-level Linear Classifier for Cervical Lesion Type Classification in the ComparisonDetector Dataset

Task	Cervix 3				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.708	0.701	0.957	0.711	0.971
DINOV2	0.322	0.267	0.792	0.327	0.932
CTransPath	0.509	0.477	0.871	0.514	0.951
ResNet-BT	0.584	0.545	0.916	0.588	0.958
UNI	0.662	0.656	0.943	0.666	0.966
Virchow	0.629	0.622	0.931	0.634	0.963

Note: A total of 1,563 cervical cytology images were included for a multiclass classification task, with approximately 130–150 images per class across 11 lesion types: HSIL, ASC-US, ASC-H, LSIL, Actinomyces, SCC, AGC, Trichomonas, Candida, Flora, and Herpes.

Dataset available at: <https://github.com/CVIU-CSU/ComparisonDetector>

Supplementary Table 117. Performance of Patch-level Linear Classifier for Cervical Lesion Classification in the Augmented Mendeley Dataset

Task	Cervix 4				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.999	0.999	1.000	0.999	1.000
DINOV2	0.775	0.760	0.943	0.777	0.925
CTransPath	0.927	0.927	0.991	0.928	0.976
ResNet-BT	0.969	0.969	0.998	0.969	0.990
UNI	0.998	0.999	1.000	0.999	1.000
Virchow	0.988	0.988	1.000	0.988	0.996

Note: A total of 3,232 cervical cytology images were included for a 4-class classification task: 815 HSIL, 791 LSIL, 812 NILM, and 814 SCC cases. Dataset available at:

<https://www.kaggle.com/datasets/srijanshovit/mendelylbc-original-with-augmented>

Supplementary Table 118. Performance of Patch-level Linear Classifier for Cervical Lesion Classification in the Original Mendeley Dataset

Task	Cervix 5				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.987	0.986	0.999	0.968	0.996
DINOV2	0.758	0.731	0.898	0.589	0.885
CTransPath	0.781	0.735	0.965	0.506	0.873
ResNet-BT	0.926	0.917	0.993	0.825	0.975
UNI	0.992	0.992	1.000	0.982	0.997
Virchow	0.979	0.979	0.998	0.948	0.991

Note: A total of 962 cervical cytology images were included for a 4-class classification task: 163 HSIL, 113 LSIL, 612 NILM, and 74 SCC cases. Dataset available at:

<https://www.kaggle.com/datasets/srijanshovit/mendelylbc-original-with-augmented>

Supplementary Table 119. Performance of Patch-level Linear Classifier for Benign and Malignant Classification in Body Cavity Fluid Cytology Images

Task	Effusion 1				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.994	0.994	0.999	0.998	0.981
DINOv2	0.918	0.916	0.944	0.961	0.775
CTransPath	0.931	0.925	0.946	1.000	0.700
ResNet-BT	0.968	0.968	0.985	0.993	0.888
UNI	0.980	0.979	0.998	0.998	0.919
Virchow	0.971	0.971	0.982	0.993	0.900

Note: A total of 693 body cavity fluid cytology images were included for a binary classification task: 160 benign and 533 malignant cases. Dataset available at: <https://www.kaggle.com/datasets/cmacus/body-cavity-fluid-cytology-images>.

Supplementary Table 120. Performance of Patch-level Linear Classifier for Cell Type Classification in Ascites Cytopathology Images

Task	Effusion 2				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.797	0.785	0.959	0.687	0.948
DINOv2	0.651	0.611	0.878	0.450	0.910
CTransPath	0.563	0.466	0.875	0.297	0.873
ResNet-BT	0.666	0.611	0.901	0.466	0.915
UNI	0.790	0.787	0.959	0.707	0.949
Virchow	0.761	0.755	0.949	0.630	0.942

Note: A total of 3,219 ascites cytology images were included for a 6-class classification task: 700 malignant suspicious, 1,483 malignant determined, 30 benign eosinophil granulocytes, 200 benign lymphocytes, 656 benign mesothelial cells, and 150 benign neutrophil granulocytes. Dataset available at:

https://pan.baidu.com/s/1r0cd0PVM5DiUmaNozMSxgg?_at_=1751855157154#list/path=%2F&parentPath=%2Fsharelink874836592-638759983842828

Supplementary Table 121. Performance of Patch-level Linear Classifier for Papillary Thyroid Carcinoma Classification in Thyroid Cytology Images

Task	Thyroid 1				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.867	0.864	0.937	0.726	0.934
DINOv2	0.740	0.706	0.763	0.338	0.934
CTransPath	0.790	0.768	0.846	0.442	0.957
ResNet-BT	0.848	0.844	0.910	0.704	0.917
UNI	0.871	0.869	0.940	0.771	0.919
Virchow	0.864	0.861	0.931	0.735	0.926

Note: A total of 2,364 thyroid cytology images were included for a binary classification task: 1,597 papillary thyroid carcinoma (PTC) and 767 non-PTC cases. Dataset available at:

<https://www.kaggle.com/datasets/maximsolopov/thyroid-cytology-224px>

Supplementary Table 122. Performance of Patch-level Linear Classifier for Thyroid Cell Type Classification in Thyroid Dataset

Task	Thyroid 2				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.867	0.865	0.966	0.852	0.930
DINOv2	0.709	0.690	0.859	0.671	0.845
CTransPath	0.826	0.818	0.954	0.796	0.906
ResNet-BT	0.801	0.793	0.936	0.775	0.894
UNI	0.870	0.868	0.968	0.855	0.931
Virchow	0.835	0.831	0.955	0.824	0.914

Note: A total of 2,875 thyroid cytology images were included for a 3-class classification task: 1,320 background, 1,016 colloid, and 539 follicular cells. The dataset provided a predefined independent test set comprising 2,400 images, which was used for performance evaluation. Dataset available at: <https://pan.baidu.com/s/1N8cid1wvvrjgHMR6aLQhRg>

Supplementary Table 123. Performance of Patch-level Linear Classifier for Leukemia Cell Classification in the C-NMC_Leukemia Dataset

Task	Blood 1				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.731	0.702	0.753	0.926	0.365
DINOv2	0.628	0.574	0.472	0.867	0.179
CTransPath	0.674	0.619	0.570	0.924	0.204
ResNet-BT	0.667	0.610	0.566	0.912	0.205
UNI	0.673	0.643	0.681	0.858	0.324
Virchow	0.671	0.623	0.643	0.898	0.242

Note: A total of 12,528 images were included for a binary classification task: 8,491 ALL and 4,037 healthy cells. Dataset available at:
<https://www.kaggle.com/datasets/andrewmvd/leukemia-classification>

Supplementary Table 124. Performance of Patch-level Linear Classifier for White Blood Cell Classification in the Blood Cell Images Dataset

Task	Blood 2				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.850	0.851	0.969	0.850	0.950
DINOv2	0.344	0.309	0.621	0.344	0.781
CTransPath	0.759	0.756	0.919	0.759	0.920
ResNet-BT	0.667	0.665	0.880	0.667	0.889
UNI	0.830	0.833	0.963	0.830	0.943
Virchow	0.775	0.774	0.945	0.775	0.925

Note: A total of 12,444 blood cell images were included for a 4-class classification task: 3120 eosinophils, 3123 neutrophils, 3098 monocytes, and 3103 lymphocytes. Dataset available at:
<https://www.kaggle.com/datasets/paultimothymooney/blood-cells>

Supplementary Table 125. Performance of Patch-level Linear Classifier for White Blood Cell Classification in the Raabin-WBC Dataset

Task	Blood 3				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.962	0.962	0.997	0.947	0.990
DINOv2	0.626	0.583	0.845	0.430	0.851
CTransPath	0.896	0.892	0.983	0.789	0.968
ResNet-BT	0.935	0.935	0.992	0.899	0.981
UNI	0.950	0.952	0.995	0.947	0.986
Virchow	0.923	0.927	0.992	0.901	0.980

Note: A total of 14,514 blood cell images were included for a 5-class classification task: 301 basophils, 1,066 eosinophils, 3,461 lymphocytes, 795 monocytes, and 8,891 neutrophils. Dataset available at:
<https://www.kaggle.com/datasets/raabindata/raabin-wbc>

Section VI. Image Retrieval Performance across Anatomical Sites in SMH cohort

This section presents image retrieval performance evaluated across 14 anatomical sites within the SMH cytology cohort, based on a candidate pool of 17,154 cytology images. Retrieval tasks are conducted at the site level, assessing the model's ability to accurately identify relevant images corresponding to specific anatomical origins. These experiments provide insights into model robustness and performance variability across diverse cytological specimens in a large-scale retrieval setting.

Supplementary Table 126. Image Retrieval Performance for Ascites Cytology in SMH cohort (n = 286)

Site	Ascites			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.825	0.920	0.937	0.762
DINOv2	0.311	0.524	0.675	0.189
CTransPath	0.601	0.776	0.857	0.570
ResNet-BT	0.629	0.808	0.850	0.552
UNI	0.808	0.909	0.948	0.745
Virchow	0.783	0.878	0.916	0.664

Supplementary Table 127. Image Retrieval Performance for Thyroid Cytology in SMH cohort (n = 356)

Site	Thyroid			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.913	0.989	0.992	0.921
DINOv2	0.413	0.713	0.817	0.301
CTransPath	0.767	0.930	0.958	0.787
ResNet-BT	0.834	0.955	0.980	0.846
UNI	0.848	0.930	0.952	0.854
Virchow	0.837	0.958	0.969	0.837

Supplementary Table 128. Image Retrieval Performance for Urine Cytology in SMH cohort (n = 284)

Site	Urine			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.820	0.912	0.940	0.739
DINOv2	0.401	0.630	0.757	0.306
CTransPath	0.725	0.842	0.908	0.581
ResNet-BT	0.711	0.849	0.912	0.637
UNI	0.778	0.884	0.947	0.715
Virchow	0.722	0.873	0.930	0.620

Supplementary Table 129. Image Retrieval Performance for Cerebrospinal Fluid Cytology in SMH cohort (n = 56)

Site	Cerebrospinal Fluid			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.875	0.929	0.964	0.732
DINOv2	0.571	0.679	0.732	0.393
CTransPath	0.732	0.821	0.839	0.554
ResNet-BT	0.696	0.804	0.821	0.589
UNI	0.750	0.857	0.893	0.571
Virchow	0.732	0.839	0.857	0.661

Supplementary Table 130. Image Retrieval Performance for Bronchial Cytology in SMH cohort (n = 20)

Site	Bronchial Brushing			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.400	0.500	0.650	0.100
DINOv2	0.000	0.050	0.050	0.000
CTransPath	0.050	0.250	0.500	0.000
ResNet-BT	0.150	0.200	0.300	0.050
UNI	0.350	0.600	0.600	0.150
Virchow	0.050	0.300	0.350	0.000

Supplementary Table 131. Image Retrieval Performance for Cervical Cytology in SMH cohort (n = 237)

Site	Cervix			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.937	0.966	0.975	0.928
DINOv2	0.561	0.738	0.806	0.532
CTransPath	0.852	0.949	0.975	0.890
ResNet-BT	0.861	0.954	0.962	0.831
UNI	0.928	0.962	0.970	0.911
Virchow	0.857	0.928	0.949	0.848

Supplementary Table 132. Image Retrieval Performance for Salivary Gland Cytology in SMH cohort (n = 52)

Site	Salivary			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.423	0.654	0.712	0.385
DINOv2	0.077	0.212	0.346	0.038
CTransPath	0.385	0.500	0.577	0.288
ResNet-BT	0.423	0.635	0.731	0.269
UNI	0.538	0.731	0.769	0.442
Virchow	0.500	0.692	0.750	0.404

Supplementary Table 133. Image Retrieval Performance for Vocal Cord Cytology in SMH cohort (n = 5)

Site	Vocal Cord			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.200	0.200	0.400	0.000
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.000	0.000	0.200	0.000
ResNet-BT	0.000	0.000	0.200	0.000
UNI	0.000	0.200	0.400	0.000
Virchow	0.200	0.200	0.400	0.000

Supplementary Table 134. Image Retrieval Performance for Breast Cytology in SMH cohort (n = 16)

Site	Breast			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.250	0.375	0.438	0.063
DINOv2	0.000	0.000	0.063	0.000
CTransPath	0.188	0.250	0.375	0.063
ResNet-BT	0.063	0.125	0.313	0.063
UNI	0.063	0.125	0.125	0.000
Virchow	0.250	0.250	0.375	0.000

Supplementary Table 135. Image Retrieval Performance for Pancreatic Cytology in SMH cohort (n = 19)

Site	Pancreas			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.158	0.316	0.368	0.053
DINOv2	0.000	0.053	0.053	0.000
CTransPath	0.105	0.316	0.368	0.000
ResNet-BT	0.000	0.000	0.105	0.000
UNI	0.105	0.158	0.211	0.053
Virchow	0.211	0.316	0.316	0.053

Supplementary Table 136. Image Retrieval Performance for Liver Cytology in SMH cohort (n = 16)

Site	Liver			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.500	0.563	0.625	0.063
DINOv2	0.000	0.000	0.063	0.000
CTransPath	0.188	0.250	0.313	0.063
ResNet-BT	0.063	0.125	0.188	0.063
UNI	0.438	0.563	0.625	0.125
Virchow	0.688	0.688	0.750	0.250

Supplementary Table 137. Image Retrieval Performance for Bile Duct Cytology in SMH cohort (n = 15)

Site	BileDuct			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.400	0.733	0.800	0.067
DINOv2	0.000	0.133	0.133	0.000
CTransPath	0.133	0.467	0.467	0.067
ResNet-BT	0.133	0.200	0.200	0.000
UNI	0.133	0.200	0.400	0.133
Virchow	0.067	0.267	0.400	0.067

Supplementary Table 138. Image Retrieval Performance for Bile Cytology in SMH cohort (n = 88)

Site	Bile			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.670	0.773	0.841	0.523
DINOv2	0.205	0.364	0.432	0.114
CTransPath	0.477	0.693	0.795	0.273
ResNet-BT	0.341	0.636	0.716	0.227
UNI	0.511	0.682	0.727	0.364
Virchow	0.364	0.534	0.705	0.227

Supplementary Table 139. Image Retrieval Performance for Lymph Node Cytology in SMH cohort (n = 437)

Site	Lymph node			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.753	0.913	0.961	0.741
DINOv2	0.332	0.622	0.792	0.192
CTransPath	0.648	0.851	0.915	0.577
ResNet-BT	0.629	0.828	0.908	0.554
UNI	0.730	0.890	0.947	0.673
Virchow	0.700	0.904	0.954	0.680

Section VII. Image Retrieval Performance across Anatomical Sites in the SCC cohort

This section summarizes image retrieval performance across 11 anatomical sites in the SCC cytology cohort, using a candidate pool of 12,527 cytology images. Each retrieval task is defined at the site level, where queries are constructed from target anatomical regions and evaluated for retrieval accuracy across the full candidate pool. The results provide an overview of model generalization across diverse cytological origins and highlight site-specific retrieval challenges in real-world settings.

Supplementary Table 140. Image Retrieval Performance for Cervical Cytology in SCC cohort (n = 230)

Site	Cervix			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.935	0.974	0.983	0.913
DINOv2	0.539	0.735	0.787	0.509
CTransPath	0.835	0.939	0.957	0.835
ResNet-BT	0.826	0.900	0.922	0.778
UNI	0.874	0.943	0.952	0.874
Virchow	0.848	0.948	0.961	0.843

Supplementary Table 141. Image Retrieval Performance for Ascites Cytology in SCC cohort (n = 258)

Site	Ascites			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.829	0.922	0.946	0.845
DINOv2	0.422	0.690	0.818	0.341
CTransPath	0.736	0.895	0.934	0.713
ResNet-BT	0.674	0.833	0.919	0.702
UNI	0.787	0.934	0.957	0.814
Virchow	0.756	0.950	0.961	0.795

Supplementary Table 142. Image Retrieval Performance for Thyroid Cytology in SCC cohort (n = 218)

Site	Thyroid			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.954	0.968	0.977	0.931
DINOv2	0.628	0.830	0.885	0.642
CTransPath	0.771	0.899	0.945	0.794
ResNet-BT	0.835	0.963	0.972	0.867
UNI	0.885	0.968	0.986	0.890
Virchow	0.858	0.936	0.977	0.839

Supplementary Table 143. Image Retrieval Performance for Urine Cytology in SCC cohort (n = 240)

Site	Urine			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.700	0.829	0.900	0.592
DINOv2	0.408	0.650	0.813	0.279
CTransPath	0.613	0.792	0.879	0.579
ResNet-BT	0.629	0.858	0.908	0.667
UNI	0.671	0.863	0.917	0.613
Virchow	0.688	0.833	0.871	0.608

Supplementary Table 144. Image Retrieval Performance for Cerebrospinal Fluid Cytology in SCC cohort (n = 100)

Site	Cerebrospinal Fluid			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.860	0.980	0.990	0.880
DINOv2	0.600	0.860	0.910	0.570
CTransPath	0.800	0.920	0.960	0.790
ResNet-BT	0.790	0.930	0.950	0.800
UNI	0.800	0.940	0.960	0.820
Virchow	0.800	0.910	0.950	0.810

Supplementary Table 145. Image Retrieval Performance for Bronchial Cytology in SCC cohort (n = 97)

Site	Bronchial brushing			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.474	0.691	0.814	0.227
DINOv2	0.155	0.381	0.557	0.021
CTransPath	0.309	0.567	0.742	0.186
ResNet-BT	0.361	0.577	0.711	0.155
UNI	0.402	0.649	0.773	0.258
Virchow	0.423	0.629	0.753	0.155

Supplementary Table 146. Image Retrieval Performance for Breast Cytology in SCC cohort (n = 10)

Site	Breast			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.100	0.100	0.100	0.100
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.200	0.200	0.200	0.000
ResNet-BT	0.100	0.200	0.200	0.000
UNI	0.300	0.400	0.400	0.100
Virchow	0.300	0.400	0.500	0.100

Supplementary Table 147. Image Retrieval Performance for Pancreatic Cytology in SCC cohort (n = 97)

Site	Pancreas			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.433	0.691	0.773	0.371
DINOv2	0.165	0.340	0.412	0.093
CTransPath	0.268	0.577	0.742	0.206
ResNet-BT	0.216	0.464	0.649	0.165
UNI	0.443	0.619	0.691	0.361
Virchow	0.371	0.577	0.670	0.299

Supplementary Table 148. Image Retrieval Performance for Bile Duct Cytology in SCC cohort (n = 10)

Site	BileDuct			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.200	0.400	0.500	0.000
DINOv2	0.000	0.100	0.100	0.000
CTransPath	0.100	0.200	0.200	0.000
ResNet-BT	0.000	0.000	0.100	0.000
UNI	0.000	0.100	0.200	0.000
Virchow	0.100	0.300	0.300	0.000

Supplementary Table 149. Image Retrieval Performance for Bile Cytology in SCC cohort (n = 27)

Site	Bile			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.519	0.704	0.741	0.185
DINOv2	0.111	0.333	0.333	0.074
CTransPath	0.407	0.519	0.556	0.111
ResNet-BT	0.333	0.519	0.667	0.074
UNI	0.519	0.519	0.593	0.296
Virchow	0.296	0.444	0.481	0.185

Supplementary Table 150. Image Retrieval Performance for Lymph Node Cytology in SCC cohort (n = 91)

Site	Lymph node			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.440	0.769	0.857	0.231
DINOv2	0.143	0.297	0.440	0.000
CTransPath	0.385	0.637	0.747	0.154
ResNet-BT	0.220	0.538	0.670	0.121
UNI	0.341	0.582	0.758	0.165
Virchow	0.341	0.659	0.791	0.165

Section VIII. Subtype-level Image Retrieval Performance in the SMH cohort

This section reports fine-grained (subtype-level) image retrieval performance across 14 anatomical sites encompassing 38 distinct subtypes within the SMH cytology cohort, based on a candidate pool of 17,154 cytology images. Each task involves either binary or multiclass retrieval according to clinically relevant subcategories, such as benign versus malignant or specific pathological subtypes. This evaluation offers insight into model performance under highly granular diagnostic conditions, reflecting the real-world cytological variability within anatomical regions.

Supplementary Table 151. Fine-grained Image Retrieval Performance for Ascites Malignancy Classification in the SMH cohort

Task	Ascites-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.787	0.909	0.934	0.717
DINOv2	0.266	0.434	0.587	0.150
CTransPath	0.549	0.724	0.801	0.514
ResNet-BT	0.594	0.773	0.825	0.500
UNI	0.762	0.878	0.916	0.682
Virchow	0.731	0.860	0.906	0.598

Note: This binary retrieval task includes 286 ascites cytology images (72 malignant, 214benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 152. Fine-grained Image Retrieval Performance for Thyroid Nodule Malignancy Classification in the SMH cohort

Task	Thyroid-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.879	0.966	0.983	0.876
DINOv2	0.329	0.607	0.716	0.216
CTransPath	0.728	0.882	0.935	0.716
ResNet-BT	0.772	0.916	0.944	0.758
UNI	0.823	0.916	0.941	0.823
Virchow	0.809	0.941	0.958	0.801

Note: This binary retrieval task includes 356 thyroid cytology images (103 malignant, 253 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 153. Fine-grained Image Retrieval Performance for Urinary Epithelial Cell Grading in the SMH cohort

Task	Urine-G			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.789	0.884	0.898	0.669
DINOv2	0.345	0.532	0.637	0.232
CTransPath	0.609	0.764	0.838	0.486
ResNet-BT	0.623	0.782	0.835	0.514
UNI	0.701	0.799	0.880	0.613
Virchow	0.662	0.810	0.863	0.546

Note: This multiclass retrieval task includes 284 urine cytology images (12 atypical urothelial cells & low-grade urothelial tumors, 79 high-grade urothelial carcinoma, 193 normal), with class-wise retrieval performed at the subtype level.

Supplementary Table 154. Fine-grained Image Retrieval Performance for Cerebrospinal Fluid Malignancy Classification in the SMH cohort

Task	CSF-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.875	0.911	0.946	0.714
DINOv2	0.554	0.643	0.696	0.393
CTransPath	0.732	0.804	0.839	0.554
ResNet-BT	0.661	0.786	0.804	0.589
UNI	0.732	0.839	0.875	0.571
Virchow	0.732	0.839	0.857	0.643

Note: This binary retrieval task includes 56 cerebrospinal fluid cytology images (7 malignant, 49 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 155. Fine-grained Image Retrieval Performance for Bronchial Brushing Cancer Subtype Classification in the SMH cohort

Task	Bronchial-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.350	0.450	0.650	0.100
DINOv2	0.000	0.050	0.050	0.000
CTransPath	0.050	0.250	0.500	0.000
ResNet-BT	0.150	0.200	0.300	0.050
UNI	0.300	0.600	0.600	0.100
Virchow	0.050	0.250	0.300	0.000

Note: This binary retrieval task includes 20 bronchial cytology images (17 non-small cell carcinoma, 3 small cell carcinoma), with class-wise retrieval performed at the subtype level.

Supplementary Table 156. Fine-grained Image Retrieval Performance for Cervical Cytology Grading in the SMH cohort

Task	Cervix-G			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.768	0.890	0.920	0.713
DINOv2	0.392	0.557	0.616	0.363
CTransPath	0.624	0.764	0.835	0.599
ResNet-BT	0.624	0.789	0.840	0.557
UNI	0.722	0.852	0.903	0.696
Virchow	0.692	0.810	0.844	0.667

Note: This multiclass retrieval task includes 237 cervical cytology images (101 normal, 30 ASC-US, 53 LSIL, 42 HSIL, 3 squamous cell carcinoma cases, and 8 adenocarcinoma), with class-wise retrieval performed at the subtype level.

Supplementary Table 157. Fine-grained Image Retrieval Performance for Salivary Gland Tumor Subtype Classification in the SMH cohort

Task	Salivary-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.404	0.596	0.673	0.365
DINOv2	0.077	0.173	0.231	0.019
CTransPath	0.385	0.462	0.519	0.288
ResNet-BT	0.365	0.558	0.654	0.250
UNI	0.519	0.673	0.712	0.442
Virchow	0.481	0.596	0.692	0.327

Note: This multiclass retrieval task includes 52 salivary gland cytology images (9 malignant, 27 pleomorphic adenoma, 16 Warthin's tumor), with class-wise retrieval performed at the subtype level.

Supplementary Table 158. Fine-grained Image Retrieval Performance for Vocal Cord Lesions Malignancy Classification in the SMH cohort

Task	VocalCord-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.200	0.200	0.400	0.000
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.000	0.000	0.000	0.000
ResNet-BT	0.000	0.000	0.200	0.000
UNI	0.000	0.200	0.400	0.000
Virchow	0.200	0.200	0.400	0.000

Note: This binary retrieval task includes 5 vocal cord cytology images (1 malignant, 4 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 159. Fine-grained Image Retrieval Performance for Breast Lesions Malignancy Classification in the SMH cohort

Task	Breast-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.188	0.375	0.438	0.000
DINOv2	0.000	0.000	0.063	0.000
CTransPath	0.188	0.250	0.375	0.063
ResNet-BT	0.063	0.125	0.250	0.063
UNI	0.063	0.125	0.125	0.000
Virchow	0.250	0.250	0.375	0.000

Note: This binary retrieval task includes 16 breast cytology images (4 malignant, 12 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 160. Fine-grained Image Retrieval Performance for Pancreatic Lesions Malignancy Classification in the SMH cohort

Task	Pancreas-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.158	0.211	0.368	0.053
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.105	0.211	0.263	0.000
ResNet-BT	0.000	0.000	0.105	0.000
UNI	0.105	0.105	0.158	0.053
Virchow	0.158	0.211	0.211	0.053

Note: This binary retrieval task includes 19 pancreatic cytology images (14 malignant, 5 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 161. Fine-grained Image Retrieval Performance for Liver Lesion Subtype Classification in the SMH cohort

Task	Liver-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.500	0.563	0.625	0.000
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.188	0.188	0.250	0.063
ResNet-BT	0.063	0.125	0.188	0.063
UNI	0.438	0.563	0.563	0.125
Virchow	0.688	0.688	0.750	0.250

Note: This multiclass retrieval task includes 16 liver cytology images (4 adenocarcinoma, 10 benign, 1 hepatocellular carcinoma, 1 neuroendocrine tumor), with class-wise retrieval performed at the subtype level.

Supplementary Table 162. Fine-grained Image Retrieval Performance for Bile Duct Malignancy Classification in the SMH cohort

Task	BileDuct-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.400	0.733	0.800	0.067
DINOv2	0.000	0.067	0.133	0.000
CTransPath	0.133	0.467	0.467	0.067
ResNet-BT	0.133	0.200	0.200	0.000
UNI	0.133	0.200	0.400	0.133
Virchow	0.067	0.267	0.400	0.067

Note: This binary retrieval task includes 15 bile duct cytology images (6 malignant, 9 benign), with class-wise retrieval performed at the subtype level.

Supplementary Table 163. Fine-grained Image Retrieval Performance for Bile Cell Grading in the SMH cohort

Task	Bile-G			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.432	0.591	0.659	0.227
DINOv2	0.091	0.170	0.227	0.045
CTransPath	0.307	0.477	0.602	0.102
ResNet-BT	0.193	0.364	0.477	0.080
UNI	0.341	0.523	0.625	0.182
Virchow	0.273	0.398	0.534	0.125

Note: This multiclass retrieval task includes 88 gallbladder/bile cytology images (36 benign, 4 mild atypical hyperplasia, 16 moderate atypical hyperplasia, 32 severe atypical hyperplasia), with class-wise retrieval performed at the subtype level.

Supplementary Table 164. Fine-grained Image Retrieval Performance for Lymph Node Malignancy Classification in the SMH cohort

Task	LN-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.725	0.888	0.936	0.686
DINOv2	0.256	0.481	0.616	0.128
CTransPath	0.622	0.817	0.886	0.540
ResNet-BT	0.600	0.785	0.876	0.515
UNI	0.709	0.870	0.922	0.629
Virchow	0.686	0.886	0.941	0.641

Note: This binary retrieval task includes 437 lymph node cytology images (212 malignant, 225 benign), with class-wise retrieval performed at the subtype level.

Section IX. Subtype-level Image Retrieval Performance in the SCC cohort

This section presents fine-grained (subtype-level) image retrieval performance across 11 anatomical sites encompassing 29 distinct subtypes within the SCC cohort, using a candidate pool of 12,527 cytology images. Both binary and multiclass retrieval tasks were evaluated based on subtype-level ground truth, providing a detailed comparison of model performance under clinically realistic conditions.

Supplementary Table 165. Fine-grained Image Retrieval Performance for Cervical Cytology Grading in the SCC cohort

Task	Cervix-G			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.730	0.926	0.948	0.722
DINOv2	0.430	0.626	0.674	0.365
CTransPath	0.661	0.817	0.878	0.617
ResNet-BT	0.643	0.770	0.817	0.578
UNI	0.709	0.852	0.887	0.661
Virchow	0.691	0.874	0.900	0.665

Note: This multiclass retrieval task includes 230 cervical cytology images (101 normal, 36 ASC-US, 48 LSIL, 43 HSIL, 1 squamous cell carcinoma, 1 adenocarcinoma) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 166. Fine-grained Image Retrieval Performance for Ascites Malignancy Classification in the SCC cohort

Task	Ascites-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.822	0.915	0.938	0.845
DINOv2	0.407	0.667	0.802	0.318
CTransPath	0.729	0.891	0.934	0.705
ResNet-BT	0.663	0.826	0.911	0.690
UNI	0.783	0.934	0.957	0.810
Virchow	0.756	0.950	0.961	0.791

Note: This binary retrieval task includes 258 ascites cytology images (128 malignant, 130 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 167. Fine-grained Image Retrieval Performance for Thyroid Cancer Subtype Classification in the SCC cohort

Task	Thyroid-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.927	0.950	0.959	0.904
DINOv2	0.596	0.803	0.858	0.615
CTransPath	0.757	0.894	0.931	0.780
ResNet-BT	0.835	0.945	0.959	0.858
UNI	0.881	0.963	0.982	0.872
Virchow	0.853	0.927	0.954	0.835

Note: This binary retrieval task includes 218 thyroid cytology images (209 papillary thyroid carcinoma, 9 other carcinoma) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 168. Fine-grained Image Retrieval Performance for Urinary Epithelial Cell Grading in the SCC cohort

Task	Urine-G			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.654	0.783	0.867	0.554
DINOv2	0.342	0.563	0.704	0.217
CTransPath	0.567	0.750	0.842	0.517
ResNet-BT	0.592	0.800	0.854	0.579
UNI	0.633	0.829	0.892	0.550
Virchow	0.617	0.763	0.817	0.517

Note: This multiclass retrieval task includes 240 urine cytology images (18 atypical urothelial cells & low-grade urothelial tumors, 107 high-grade urothelial carcinoma, 115 normal) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 169. Fine-grained Image Retrieval Performance for Cerebrospinal Fluid Malignancy Classification in the SCC cohort

Task	CSF-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.850	0.970	0.990	0.860
DINOv2	0.580	0.840	0.890	0.560
CTransPath	0.780	0.920	0.960	0.780
ResNet-BT	0.780	0.920	0.940	0.790
UNI	0.790	0.930	0.950	0.810
Virchow	0.780	0.910	0.940	0.770

Note: This binary retrieval task includes 100 CSF cytology images (7 malignant, 93 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 170. Fine-grained Image Retrieval Performance for

Cancer Subtype Classification for Bronchial Brushing in the SCC cohort

Task	Bronchial-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.433	0.588	0.701	0.144
DINOv2	0.052	0.227	0.340	0.000
CTransPath	0.268	0.464	0.598	0.093
ResNet-BT	0.309	0.495	0.598	0.093
UNI	0.330	0.505	0.598	0.186
Virchow	0.371	0.515	0.608	0.124

Note: This multiclass retrieval task includes 97 bronchial cytology images (40 squamous cell carcinoma, 38 adenocarcinoma, 19 small cell carcinoma) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 171. Fine-grained Image Retrieval Performance for Breast Lesion Malignant Classification in the SCC cohort

Task	Breast-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.100	0.100	0.100	0.100
DINOv2	0.000	0.000	0.000	0.000
CTransPath	0.200	0.200	0.200	0.000
ResNet-BT	0.100	0.200	0.200	0.000
UNI	0.300	0.400	0.400	0.100
Virchow	0.300	0.400	0.500	0.100

Note: This binary retrieval task includes 10 breast cytology images (1 malignant, 9 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 172. Fine-grained Image Retrieval Performance for Pancreatic Cytology in the SCC cohort

Task	Pancreas-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.412	0.660	0.742	0.351
DINOv2	0.134	0.299	0.371	0.082
CTransPath	0.258	0.557	0.711	0.196
ResNet-BT	0.216	0.454	0.619	0.144
UNI	0.412	0.588	0.660	0.340
Virchow	0.371	0.577	0.660	0.299

Note: This binary retrieval task includes 97 pancreatic cytology images (73 malignant, 24 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 173. Fine-grained Image Retrieval Performance for Bile Duct Malignant Classification in the SCC cohort

Task	BileDuct-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.200	0.400	0.500	0.000
DINOv2	0.000	0.100	0.100	0.000
CTransPath	0.100	0.200	0.200	0.000
ResNet-BT	0.000	0.000	0.100	0.000
UNI	0.000	0.100	0.200	0.000
Virchow	0.100	0.300	0.300	0.000

Note: This binary retrieval task includes 10 bile duct cytology images (9 malignant, 1 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 174. Fine-grained Image Retrieval Performance for Bile Malignant Classification in the SCC cohort

Task	Bile-BMC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.519	0.704	0.741	0.185
DINOv2	0.111	0.333	0.333	0.037
CTransPath	0.333	0.481	0.519	0.074
ResNet-BT	0.296	0.481	0.630	0.074
UNI	0.481	0.481	0.556	0.259
Virchow	0.259	0.407	0.444	0.111

Note: This binary retrieval task includes 27 bile cytology images (20 malignant, 7 benign) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Supplementary Table 175. Fine-grained Image Retrieval Performance for Lymph Node Cancer Subtype Classification in the SCC cohort

Task	LN-CSC			
Metrics	ACC@1	ACC@3	ACC@5	MVACC@5
CROWN	0.341	0.626	0.736	0.198
DINOv2	0.099	0.209	0.308	0.000
CTransPath	0.363	0.604	0.714	0.110
ResNet-BT	0.165	0.440	0.615	0.055
UNI	0.297	0.505	0.681	0.110
Virchow	0.275	0.549	0.648	0.143

Note: This multiclass retrieval task includes 91 lymph node cytology images (23 squamous cell carcinoma, 65 adenocarcinoma, 3 small cell carcinoma) in the SCC cohort, with class-wise retrieval performed at the subtype level.

Section X. Cell Segmentation Results on Public Datasets

Cell segmentation was performed on one publicly available thyroid cytology dataset and one cervical cytology dataset. Segmentation performance was evaluated using DICE Similarity Coefficient and Intersection over Union (IoU) scores, where higher values indicate better segmentation accuracy.

Supplementary Table 176. Cell Segmentation Performance on the Public Thyroid Cytology Dataset

Task	Thyroid Seg.					
Metrics	DICE			IoU		
	Background	Cell	Overall	Background	Cell	Overall
CROWN	0.973	0.718	0.846	0.948	0.560	0.754
DINOv2	0.975	0.711	0.843	0.952	0.551	0.751
CTransPath	0.950	0.409	0.680	0.905	0.257	0.581
ResNet-BT	0.957	0.449	0.703	0.917	0.289	0.603
UNI	0.972	0.687	0.830	0.946	0.523	0.735
Virchow	0.975	0.697	0.836	0.951	0.535	0.743

Note: The dataset is available at <https://github.com/Junchao-Zhu/TCSegNet>

Supplementary Table 177. Cell Segmentation Performance on the Public Cervical Cytology Dataset

Task	Cervix Seg.					
Metrics	DICE			IoU		
	Background	Cell	Overall	Background	Cell	Overall
CROWN	0.964	0.475	0.719	0.930	0.312	0.621
DINOv2	0.976	0.432	0.704	0.953	0.276	0.614
CTransPath	0.976	0.458	0.717	0.952	0.297	0.625
ResNet-BT	0.978	0.410	0.694	0.957	0.258	0.608
UNI	0.974	0.432	0.703	0.949	0.275	0.612
Virchow	0.973	0.450	0.712	0.948	0.290	0.619

Note: The dataset is available at <https://osf.io/cka2f/files/osfstorage#>

Section XI. Object Detection Results on Public Datasets

Object detection was performed on two publicly available blood smear datasets. Detection performance was evaluated using AP, AP50, and AP75. These metrics are standard indicators of object detection accuracy, with higher values indicating better performance.

Supplementary Table 178. Object Detection Performance on the Public Blood Detect.1 Dataset

Task	Blood Detect.1		
Metrics	AP	AP50	AP75
CROWN	41.6%	67.5%	44.1%
DINOv2	17.6%	44.1%	8.7%
CTransPath	17.5%	33.1%	16.7%
ResNet-BT	30.2%	53.3%	28.6%
UNI	40.5%	64.8%	47.8%
Virchow	19.3%	44.9%	11.4%

Note: The dataset is available at

<https://www.heywhale.com/mw/dataset/62c2af90913a54a66038165a/file>

Supplementary Table 179. Object Detection Performance on the Public Blood Detect.2 Dataset

Task	Blood Detect.2		
Metrics	AP	AP50	AP75
CROWN	12.5%	45.5%	0.5%
DINOv2	9.2%	36.3%	1.3%
CTransPath	3.2%	13.0%	0.5%
ResNet-BT	10.4%	37.1%	0.8%
UNI	11.9%	49.1%	0.2%
Virchow	11.2%	40.1%	2.8%

Note: The dataset is available at https://github.com/Shenggan/BCCD_Dataset

XII. Performance of Weakly Supervised Slide-level Classification

This section presents the performance of slide-level classification models across 12 downstream diagnostic tasks. The tasks include 8 binary classification tasks and 4 multiclass classification tasks, spanning multiple anatomical sites and pathological categories. To evaluate model performance, slides were randomly split into training and testing sets, with approximately 90% of slides used for model training and the remaining 10% reserved for independent testing. This random splitting procedure was repeated five times, yielding five separately trained models. Detailed results demonstrate the models' capability in slide-level feature discrimination across diverse cytology domains.

Supplementary Table 180. Performance of Slide-level Classification for Lymph Node Malignancy Detection

TASK	LN-BMC				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.919	0.919	0.965	0.920	0.920
DINOv2	0.712	0.706	0.799	0.715	0.715
CTransPath	0.834	0.833	0.882	0.834	0.834
ResNet-BT	0.831	0.831	0.906	0.831	0.831
UNI	0.894	0.894	0.949	0.895	0.895
Virchow	0.874	0.874	0.923	0.874	0.874

Note: A total of 791 lymph node slides were included for a binary classification task: 396 with malignancy and 395 without malignancy.

Supplementary Table 181. Performance of Slide-level Classification for Primary Origin of Lymph Node Metastatic Cancer

TASK	LN-Origin				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.888	0.794	0.912	0.768	0.768
DINOv2	0.837	0.722	0.868	0.732	0.732
CTransPath	0.862	0.769	0.850	0.790	0.790
ResNet-BT	0.848	0.742	0.860	0.741	0.741
UNI	0.882	0.812	0.863	0.829	0.829
Virchow	0.882	0.763	0.864	0.776	0.776

Note: A total of 317 lymph node metastasis slides were included for a binary classification task: 258 originating from thyroid cancer and 59 from other primary cancers.

Supplementary Table 182. Performance of Slide-level Classification for Thyroid Nodule Malignancy Detection

TASK	Thyroid-BMC				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.925	0.925	0.978	0.925	0.925
DINOv2	0.666	0.663	0.725	0.666	0.666
CTransPath	0.834	0.833	0.917	0.834	0.834
ResNet-BT	0.806	0.804	0.894	0.806	0.806
UNI	0.916	0.916	0.979	0.916	0.916
Virchow	0.881	0.880	0.952	0.881	0.881

Note: A total of 849 thyroid nodule slides were included for a binary classification task: 433 malignant and 416 benign cases.

Supplementary Table 183. Performance of Slide-level Classification for Cervical Cell Abnormality Detection

TASK	Cervix-AD				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.693	0.690	0.784	0.693	0.693
DINOv2	0.537	0.516	0.651	0.531	0.531
CTransPath	0.621	0.600	0.671	0.613	0.613
ResNet-BT	0.526	0.492	0.558	0.526	0.526
UNI	0.669	0.665	0.723	0.672	0.672
Virchow	0.592	0.578	0.667	0.583	0.583

Note: A total of 468 cervical cytology slides were included for a binary classification task: 246 abnormal and 222 normal.

Supplementary Table 184. Performance of Slide-level Classification for Cervical Cell HPV Infection Detection

TASK	Cervix-HPVInf				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.731	0.702	0.802	0.703	0.703
DINOv2	0.613	0.487	0.518	0.518	0.518
CTransPath	0.664	0.610	0.669	0.612	0.612
ResNet-BT	0.633	0.492	0.586	0.537	0.537
UNI	0.696	0.650	0.784	0.646	0.646
Virchow	0.648	0.597	0.639	0.599	0.599

Note: A total of 508 cervical cytology slides were included for a binary classification task: 190 negative and 318 positive for HPV infection.

Supplementary Table 185. Performance of Slide-level Classification for Cervical Cell Candida Infection Detection

TASK	Cervix-CID				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.910	0.881	0.917	0.866	0.866
DINOv2	0.759	0.693	0.738	0.694	0.694
CTransPath	0.803	0.753	0.852	0.756	0.756
ResNet-BT	0.823	0.754	0.834	0.734	0.734
UNI	0.870	0.830	0.892	0.817	0.817
Virchow	0.850	0.799	0.892	0.792	0.792

Note: A total of 598 cervical cytology slides were included for a binary classification task: 171 Candida positive and 427 negative.

Supplementary Table 186. Performance of Slide-level Classification for Cervical Cell Bacterial Infection Detection

TASK	Cervix-BID				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.988	0.987	0.999	0.986	0.986
DINOv2	0.798	0.780	0.885	0.784	0.784
CTransPath	0.944	0.938	0.985	0.937	0.937
ResNet-BT	0.917	0.910	0.977	0.916	0.916
UNI	0.949	0.944	0.983	0.943	0.943
Virchow	0.925	0.917	0.978	0.914	0.914

Note: A total of 502 cervical cytology slides were included for a binary classification task: 180 positive and 322 negative for bacterial infection.

Supplementary Table 187. Performance of Slide-level Classification for Thyroid Nodule Malignancy Detection (Cell Block Samples)

TASK	Thyroid-BMC (Cell Block)				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.882	0.881	0.953	0.880	0.880
DINOv2	0.724	0.687	0.909	0.727	0.727
CTransPath	0.862	0.860	0.976	0.863	0.863
ResNet-BT	0.862	0.860	0.921	0.860	0.860
UNI	0.940	0.939	0.944	0.940	0.940
Virchow	0.842	0.836	0.960	0.840	0.840

Note: A total of 100 thyroid nodule cell block slides were included for a binary classification task: 50 malignant and 50 benign cases.

Supplementary Table 188. Performance of Slide-level Classification for Primary Detailed Origin of Lymph Node Metastatic Cancer

TASK	LN-Origin (Detailed)				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.900	0.567	0.874	0.579	0.942
DINOv2	0.643	0.258	0.679	0.290	0.859
CTransPath	0.803	0.389	0.827	0.467	0.910
ResNet-BT	0.832	0.380	0.817	0.400	0.914
UNI	0.875	0.480	0.875	0.496	0.942
Virchow	0.861	0.417	0.888	0.458	0.920

Note: A total of 317 lymph node metastasis slides were included for a five-class classification task: 258 thyroid cancer, 11 nasopharyngeal carcinoma, 30 lung cancer, 12 lymphoma, and 6 breast cancer.

Supplementary Table 189. Performance of Slide-level Classification for Lymph Node Cancer Subtype Identification

TASK	LN-CSC				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.782	0.722	0.948	0.718	0.877
DINOv2	0.473	0.263	0.575	0.341	0.677
CTransPath	0.729	0.709	0.850	0.692	0.843
ResNet-BT	0.594	0.550	0.791	0.563	0.774
UNI	0.796	0.725	0.908	0.735	0.882
Virchow	0.608	0.567	0.781	0.567	0.781

Note: A total of 136 lymph node cancer slides were included for a three-class classification task: 43 squamous cell carcinomas, 67 adenocarcinomas, and 26 small cell carcinomas.

Supplementary Table 190. Performance of Slide-level Classification for Cervical Cell HPV Subtype Infection Detection

TASK	Cervix-HPVInf (Detailed)				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.415	0.334	0.612	0.357	0.800
DINOv2	0.328	0.206	0.526	0.257	0.755
CTransPath	0.316	0.245	0.582	0.278	0.763
ResNet-BT	0.328	0.242	0.536	0.274	0.762
UNI	0.367	0.301	0.604	0.311	0.780
Virchow	0.355	0.296	0.562	0.309	0.778

Note: A total of 508 cervical cytology slides were included for a four-class classification task: 190 negative, 108 low-risk HPV positive, 141 high-risk HPV positive, and 69 positive for both low- and high-risk HPV.

Supplementary Table 191. Performance of Slide-level Classification for TIS System Classification in Effusion Cytology Cell Blocks

TASK	TIS-CLS (Cell Block)				
Metrics	Accuracy	F1	AUC	Sensitivity	Specificity
CROWN	0.493	0.428	0.752	0.496	0.828
DINOv2	0.493	0.428	0.752	0.496	0.828
CTransPath	0.396	0.341	0.715	0.428	0.814
ResNet-BT	0.396	0.302	0.705	0.327	0.790
UNI	0.622	0.567	0.802	0.603	0.870
Virchow	0.513	0.428	0.753	0.428	0.830

Note: A total of 96 effusion cytology slides were included for a four-class classification task: 44 TIS-II, 16 TIS-III, 9 TIS-IV, and 27 TIS-V cases.

XIII. Performance of Slide-level Few-shot Classification

In each diagnostic task, few (K)-shot classification was performed by selecting K slides from each class as the support set. The feature representations of these support slides were averaged to construct the corresponding class prototypes. For all remaining query slides, we computed their distances to the class prototypes, and each slide was assigned to the class with the nearest prototype. Classification accuracy was then calculated based on these assignments. This approach directly evaluates the discriminative capacity of slide-level few-shot prototypes across diverse cytology domains. In our experiments, K was evaluated up to 256-shot when sufficient slides were available. For classes with limited sample size, the maximum K was restricted by the total number of available slides.

Supplementary Table 192. Performance of Slide-level Few-shot Classification for Lymph Node Malignancy Detection

Task	LN-BMC							
	1	4	8	16	32	64	128	256
CROWN	0.525	0.648	0.716	0.741	0.764	0.762	0.780	0.779
DINOv2	0.511	0.584	0.577	0.581	0.589	0.623	0.613	0.617
CTransPath	0.556	0.616	0.710	0.699	0.722	0.721	0.720	0.728
ResNet-BT	0.593	0.582	0.688	0.707	0.713	0.712	0.724	0.734
UNI	0.529	0.661	0.686	0.742	0.762	0.762	0.765	0.748
Virchow	0.574	0.609	0.712	0.714	0.724	0.728	0.734	0.738

Note: A total of 791 lymph node slides were included for a binary classification task: 396 with malignancy and 395 without malignancy.

Supplementary Table 193. Performance of Slide-level Few-shot Classification for Primary Origin of Lymph Node Metastatic Cancer

Task	LN-Origin				
	1	4	8	16	32
CROWN	0.485	0.664	0.773	0.795	0.826
DINOv2	0.565	0.613	0.611	0.666	0.660
CTransPath	0.641	0.563	0.659	0.678	0.652
ResNet-BT	0.476	0.646	0.660	0.690	0.723
UNI	0.548	0.646	0.665	0.754	0.784
Virchow	0.569	0.530	0.608	0.649	0.702

Note: A total of 317 lymph node metastasis slides were included for a binary classification task: 258 originating from thyroid cancer and 59 from other primary cancers.

Supplementary Table 194. Performance of Slide-level Few-shot Classification for Thyroid Nodule Malignancy Detection

Task	Thyroid-BMC							
Shot	1	4	8	16	32	64	128	256
CROWN	0.535	0.608	0.642	0.666	0.695	0.721	0.734	0.749
DINOv2	0.556	0.601	0.608	0.639	0.629	0.632	0.636	0.631
CTransPath	0.524	0.553	0.563	0.598	0.615	0.622	0.664	0.637
ResNet-BT	0.539	0.540	0.569	0.596	0.618	0.639	0.640	0.672
UNI	0.554	0.570	0.616	0.639	0.646	0.678	0.706	0.693
Virchow	0.520	0.528	0.573	0.584	0.597	0.613	0.632	0.644

Note: A total of 849 thyroid nodule slides were included for a binary classification task: 433 malignant and 416 benign cases.

Supplementary Table 195. Performance of Slide-level Few-shot Classification for Cervical Cell Abnormality Detection

Task	Cervix-AD						
Shot	1	4	8	16	32	64	128
CROWN	0.514	0.524	0.548	0.578	0.579	0.620	0.609
DINOv2	0.514	0.510	0.550	0.516	0.520	0.533	0.509
CTransPath	0.507	0.514	0.527	0.539	0.544	0.531	0.561
ResNet-BT	0.495	0.520	0.505	0.501	0.516	0.514	0.545
UNI	0.497	0.510	0.538	0.533	0.527	0.567	0.566
Virchow	0.505	0.522	0.544	0.530	0.566	0.580	0.571

Note: A total of 468 cervical cytology slides were included for a binary classification task: 246 abnormal and 222 normal.

Supplementary Table 196. Performance of Slide-level Few-shot Classification for Cervical Cell HPV Infection Detection

Task	Cervix-HPVInf						
Shot	1	4	8	16	32	64	128
CROWN	0.459	0.515	0.525	0.559	0.604	0.614	0.622
DINOv2	0.501	0.497	0.503	0.517	0.508	0.528	0.536
CTransPath	0.532	0.524	0.529	0.515	0.532	0.534	0.546
ResNet-BT	0.492	0.488	0.527	0.517	0.538	0.532	0.555
UNI	0.518	0.511	0.523	0.528	0.549	0.557	0.571
Virchow	0.535	0.486	0.535	0.535	0.577	0.570	0.575

Note: A total of 508 cervical cytology slides were included for a binary classification task: 190 negative and 318 positive for HPV infection.

Supplementary Table 197. Performance of Slide-level Few-shot Classification for Cervical Cell Candida Infection Detection

Task	Cervix-CID						
Shot	1	4	8	16	32	64	128
CROWN	0.591	0.671	0.698	0.761	0.780	0.790	0.834
DINOv2	0.474	0.506	0.477	0.517	0.496	0.493	0.485
CTransPath	0.496	0.538	0.546	0.542	0.601	0.621	0.624
ResNet-BT	0.486	0.576	0.595	0.603	0.619	0.642	0.651
UNI	0.495	0.551	0.570	0.627	0.647	0.684	0.711
Virchow	0.521	0.583	0.556	0.613	0.595	0.627	0.602

Note: A total of 598 cervical cytology slides were included for a binary classification task: 171 Candida positive and 427 negative.

Supplementary Table 198. Performance of Slide-level Few-shot Classification for Cervical Cell Bacterial Infection Detection

Task	Cervix-BID						
Shot	1	4	8	16	32	64	128
CROWN	0.748	0.897	0.935	0.941	0.941	0.952	0.963
DINOv2	0.477	0.496	0.535	0.561	0.569	0.561	0.537
CTransPath	0.514	0.606	0.611	0.682	0.728	0.724	0.747
ResNet-BT	0.539	0.720	0.744	0.757	0.779	0.777	0.752
UNI	0.559	0.627	0.687	0.713	0.779	0.750	0.785
Virchow	0.540	0.604	0.682	0.703	0.708	0.715	0.700

Note: A total of 502 cervical cytology slides were included for a binary classification task: 180 positive and 322 negative for bacterial infection.

Supplementary Table 199. Performance of Slide-level Few-shot Classification for Thyroid Nodule Malignancy Detection (Cell Block Samples)

Task	Thyroid-BMC (cell block)				
Shot	1	4	8	16	32
CROWN	0.523	0.546	0.583	0.600	0.661
DINOv2	0.560	0.537	0.546	0.666	0.656
CTransPath	0.542	0.565	0.630	0.656	0.694
ResNet-BT	0.522	0.522	0.513	0.566	0.608
UNI	0.551	0.577	0.625	0.631	0.669
Virchow	0.528	0.571	0.552	0.587	0.642

Note: A total of 100 thyroid nodule cell block slides were included for a binary classification task: 50 malignant and 50 benign.

Supplementary Table 200. Performance of Slide-level Few-shot Classification for Primary Detailed Origin of Lymph Node Metastatic Cancer

Task	LN-Origin (detailed)	
Shot	1	4
CROWN	0.253	0.356
DINOv2	0.269	0.341
CTransPath	0.220	0.307
ResNet-BT	0.230	0.324
UNI	0.265	0.371
Virchow	0.247	0.292

Note: A total of 317 lymph node metastasis slides were included for a five-class classification task: 258 thyroid cancer, 11 nasopharyngeal carcinoma, 30 lung cancer, 12 lymphoma, and 6 breast cancer.

Supplementary Table 201. Performance of Slide-level Few-shot Classification for Lymph Node Cancer Subtype Identification

Task	LN-CSC			
Shot	1	4	8	16
CROWN	0.492	0.487	0.603	0.638
DINOv2	0.357	0.373	0.389	0.382
CTransPath	0.416	0.452	0.470	0.482
ResNet-BT	0.402	0.449	0.431	0.528
UNI	0.441	0.494	0.548	0.578
Virchow	0.417	0.468	0.471	0.461

Note: A total of 136 lymph node cancer slides were included for a three-class classification task: 43 squamous cell carcinomas, 67 adenocarcinomas, and 26 small cell carcinomas.

Supplementary Table 202. Performance of Slide-level Few-shot Classification for Cervical Cell HPV Subtype Infection Detection

Task	Cervix-HPVInf (detailed)					
Shot	1	4	8	16	32	64
CROWN	0.315	0.298	0.324	0.339	0.360	0.422
DINOv2	0.260	0.247	0.253	0.248	0.252	0.204
CTransPath	0.263	0.278	0.277	0.291	0.304	0.316
ResNet-BT	0.265	0.271	0.270	0.277	0.288	0.297
UNI	0.244	0.272	0.282	0.301	0.302	0.352
Virchow	0.302	0.293	0.287	0.320	0.341	0.348

Note: A total of 508 cervical cytology slides were included for a four-class classification task: 190 negative, 108 low-risk HPV positive, 141 high-risk HPV positive, and 69 positive for both low- and high-risk HPV.

Supplementary Table 203. Performance of Slide-level Few-shot Classification for TIS System Classification in Effusion Cytology Cell Blocks

Task	TIS-CLS (cell block)		
Shot	1	4	8
CROWN	0.314	0.341	0.328
DINOv2	0.279	0.325	0.339
CTransPath	0.300	0.409	0.384
ResNet-BT	0.326	0.359	0.378
UNI	0.298	0.343	0.400
Virchow	0.291	0.321	0.344

Note: A total of 96 effusion cytology slides were included for a four-class classification task: 44 TIS-II, 16 TIS-III, 9 TIS-IV, and 27 TIS-V cases.

Supplementary Table 204. Details of Training Dataset

Cohort		SMH	ZPH
Total slide number		23,556	2,893
Site	Cervix	12,844	0
	Lymph node	1,956	69
	Head and Neck	8,756	838
	Effusion	0	1,387
	Thorax	0	433
	Urinary Tract	0	113
	Abdomen	0	53
Sex	Man	2,712	1,218
	Woman	20,844	1,675
Mean Age (years)		41.39	44.65
Proportion of abnormal/malignant cases	Cervix	5%	NA
	Lymph node	41%	36%
	Head and Neck	68%	32%
	Effusion	NA	12%
	Thorax	NA	29%
	Urinary Tract	NA	44%
	Abdomen	NA	19%

Note: NA, not applicable.

Supplementary Table 205. Expert evaluation of attention heatmaps

Feature	Cytopathologist1	Cytopathologist2	Kappa
Nuclear atypia	94%	97%	0.653
Chromatin irregularity	40%	29%	0.716
Multinucleation	8%	3%	0.525
Nuclear hyperchromasia	44%	26%	0.618
Increased nuclear-to-cytoplasmic ratio	41%	53%	0.763
Cytoplasmic vacuolation	55%	59%	0.715
Scattered cells	4%	5%	0.884
Papillary structures	6%	12%	0.638
Glandular structures	28%	22%	0.735
Clustered cell arrangement	59%	62%	0.728