Supporting Information

Conventional Machine Learning and Deep Learning Approach for Multi-Classification of Breast Cancer Histopathology Images

TABLE S1
CLASSIFICATION ACCURACY FROM HANDCRAFTED FEATURES (HF) AND CONVENTIONAL CLASSIFIERS FOR A BALANCED AND UNBALANCED DATASET

Classifiers	Model	Magnification Factor (Accuracy in %)							
		40x		100x		200x		400x	
		Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced
Random Forest (RF)	HF+RF (400)	89.50	85.42	**91.35**	85.22	86.78	82.24	85.15	80.77
	HF+RF (500)	89.50	85.46	91.15	85.22	86.78	82.43	**86.78**	80.50
	HF+RF (1000)	**90.33**	85.42	89.88	85.32	86.78	82.37	85.15	80.98
	HF+RF (4000)	90.28	85.87	90.10	86.00	**87.43**	82.37	86.55	80.98
	HF+RF (50)	87.69	84.69	87.36	84.44	85.96	80.93	85.15	80.57
	HF+RF (200)	89.14	85.23	89.86	85.81	86.68	82.05	85.93	81.53
Support Vector Machine (SVM)	HF+SVM(L, 1)	79.15	77.51	82.77	74.17	81.88	72.08	78.90	73.64
	HF+SVM (L, 5)	82.20	80.78	87.58	79.16	86.48	76.53	82.95	77.72
	HF+SVM (R, 1)	12.15	59.62	10.41	18.58	09.07	45.61	10.00	44.46
	HF+SVM (R, 5)	38.50	63.71	40.00	32.07	44.37	52.38	37.18	52.97
	HF+SVM (S, 1)	11.28	59.49	09.16	17.70	09.07	45.61	10.00	44.46
	HF+SVM (S, 5)	14.61	60.44	15.13	27.47	18.61	46.49	12.81	47.23
Gaussian Navie Bays (NB)	HF+NB	51.80	36.73	54.58	39.82	59.58	34.12	53.90	35.13
Logistic Regression (LR)	HF+LR(L2)	72.36	74.56	76.66	66.44	77.59	69.26	73.59	69.85
Decision Tree Classifier (CART)	HF+CART	71.49	74.61	72.63	70.06	73.36	72.21	69.53	70.88
K Nearest Neighbors (kNN)	HF+kNN	79.60	81.06	81.66	79.94	82.73	79.30	79.37	80.15
Linear Discriminant Analysis (LDA)	HF+LDA	71.20	72.93	72.63	64.09	73.78	68.57	72.03	67.56

TABLE S2

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS BASELINE (BL) MODEL FOR THE 40X MAGNIFICATION FACTOR

Network Used	Model	Accuracy	Precision	Recall	F_1-score
VGG16	VGG16 as BL	15.58	0.06	0.16	0.08
VGG19	VGG19 as BL	15.58	0.02	0.10	0.03
ResNet50	ResNet50 as BL	38.96	0.42	0.39	0.34

TABLE S3

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS BASELINE (BL) MODEL FOR THE 100X MAGNIFICATION FACTOR

Network Used	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16 as BL	11.25	0.03	0.11	0.04
VGG19	VGG19 as BL	7.50	0.01	0.07	0.01
ResNet50	ResNet50 as BL	33.75	0.39	0.34	0.29

TABLE S4

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS BASELINE (BL) MODEL FOR THE 200X MAGNIFICATION FACTOR

Network Used	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16 as BL	9.21	0.01	0.09	0.02
VGG19	VGG19 as BL	9.21	0.01	0.09	0.02
ResNet50	ResNet50 as BL	39.47	0.37	0.39	0.33

TABLE S5

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS BASELINE (BL) MODEL FOR THE 400X MAGNIFICATION FACTOR

Network Used	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16 as BL	8.33	0.01	0.08	0.01
VGG19	VGG19 as BL	8.33	0.01	0.08	0.01
ResNet50	ResNet50 as BL	33.33	0.27	0.33	0.28

TABLE S6

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS FEATURE EXTRACTOR WITH CONVENTIONAL CLASSIFIERS FOR THE 40X MAGNIFICATION FACTOR

Pre-trained Network	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16+RF (50)	79.22	0.83	0.81	0.80
	VGG16+RF (200)	81.82	0. 85	0.82	0.81
	VGG16+RF (400)	80.52	0.84	0.82	0.82
	VGG16+RF (500)	80.52	0.83	0.81	0.80
	VGG16+RF (1000)	80.52	0.83	0.81	0.80
	VGG16+RF (4000)	79.22	0.82	0.79	0.79
	VGG16+SVM(L, 1)	**89.61**	0.90	0.88	0.88
	VGG16+SVM (L, 5)	88.31	0.90	0.88	0.88
	VGG16+SVM (R, 1)	72.73	0.73	0.66	0.66
	VGG16+SVM (R, 5)	80.52	0.83	0.81	0.79
	VGG16+SVM (S, 1)	6.49	0.03	0.14	0.04
	VGG16+SVM (S, 5)	55.84	0.04	0.15	0.05
	VGG16+LR(L2)	**89.61**	0.91	0.90	0.90
	VGG16+CART	63.64	0.65	0.64	0.64
	VGG16+kNN	81.82	0.84	0.82	0.82
	VGG16+LDA	81.82	0.83	0.82	0.82
	VGG16+NB	63.64	0.68	0.64	0.64
VGG19	VGG19+RF (50)	79.22	0.78	0.78	0.77
	VGG19+RF (200)	79.22	0.78	0.78	0.77
	VGG19+RF (400)	79.22	0.80	0.78	0.77
	VGG19+RF (500)	79.22	0.81	0.79	0.78
	VGG19+RF (1000)	76.62	0.77	0.77	0.75
	VGG19+RF (4000)	75.32	0.76	0.75	0.74
	VGG19+SVM(L, 1)	**91.21**	0.91	0.90	0.90
	VGG19+SVM (L, 5)	**91.21**	0.91	0.90	0.90
	VGG19+SVM (R, 1)	70.13	0.59	0.60	0.58
	VGG19+SVM (R, 5)	77.92	0.72	0.73	0.72
	VGG19+SVM (S, 1)	15.58	0.03	0.16	0.05
	VGG19+SVM (S, 5)	10.39	0.03	0.16	0.05
	VGG19+LR(L2)	88.31	0.89	0.88	0.88
	VGG19+CART	54.55	0.56	0.55	0.54
	VGG19+kNN	76.62	0.81	0.81	0.80
	VGG19+LDA	77.92	0.79	0.78	0.78
	VGG19+NB	51.95	0.55	0.52	0.52
ResNet 50	ResNet 50+RF (50)	63.64	0.66	0.64	0.63
	ResNet 50+RF (200)	62.34	0.63	0.62	0.62
	ResNet 50+RF (400)	63.64	0.63	0.64	0.62
	ResNet 50+RF (500)	66.23	0.67	0.66	0.66
	ResNet 50+RF (1000)	67.53	0.67	0.66	0.66
	ResNet 50+RF (4000)	66.23	0.67	0.66	0.66
	ResNet 50+SVM(L, 1)	**67.53**	0.70	0.68	0.68
	ResNet 50+SVM (L, 5)	64.94	0.67	0.66	0.66

	ResNet 50+SVM (R, 1)	23.38	0.20	0.16	0.10
	ResNet 50+SVM (R, 5)	37.66	0.32	0.25	0.22
	ResNet 50+SVM (S, 1)	12.99	0.02	0.10	0.03
	ResNet 50+SVM (S, 5)	31.17	0.24	0.17	0.12
	ResNet 50+LR(L2)	63.64	0.66	0.64	0.63
	ResNet 50+CART	41.56	0.47	0.42	0.42
	ResNet 50+kNN	49.35	0.49	0.48	0.48
	ResNet 50+LDA	36.36	0.34	0.35	0.33
	ResNet 50+NB	29.87	0.28	0.30	0.26

TABLE S7

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS FEATURE EXTRACTOR WITH CONVENTIONAL CLASSIFIERS FOR THE 100X MAGNIFICATION FACTOR

Pre-trained Network	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16+RF (50)	75.00	0.74	0.74	0.73
	VGG16+RF (200)	77.50	0.83	0.79	0.78
	VGG16+RF (400)	80.00	0.85	0.80	0.80
	VGG16+RF (500)	**81.25**	0.86	0.81	0.82
	VGG16+RF (1000)	**83.75**	0.87	0.84	0.84
	VGG16+RF (4000)	**89.75**	0.89	0.87	0.87
	VGG16+SVM(L, 1)	**89.75**	0.87	0.85	0.85
	VGG16+SVM (L, 5)	87.50	0.87	0.85	0.85
	VGG16+SVM (R, 1)	60.00	0.64	0.61	0.60
	VGG16+SVM (R, 5)	77.50	0.81	0.79	0.78
	VGG16+SVM (S, 1)	16.25	0.01	0.07	0.01
	VGG16+SVM (S, 5)	15.00	0.01	0.07	0.01
	VGG16+LR(L2)	**81.25**	0.83	0.81	0.81
	VGG16+CART	56.25	0.58	0.56	0.56
	VGG16+kNN	75.00	0.80	0.78	0.77
	VGG16+LDA	78.75	0.81	0.79	0.78
	VGG16+NB	51.25	0.53	0.51	0.51
VGG19	VGG19+RF (50)	75.00	0.74	0.74	0.73
	VGG19+RF (200)	77.50	0.83	0.79	0.78
	VGG19+RF (400)	80.00	0.85	0.80	0.80
	VGG19+RF (500)	**81.25**	0.86	0.81	0.82
	VGG19+RF (1000)	**83.75**	0.87	0.84	0.84
	VGG19+RF (4000)	**89.75**	0.89	0.87	0.87
	VGG19+SVM(L, 1)	**89.75**	0.87	0.85	0.85
	VGG19+SVM (L, 5)	87.50	0.87	0.85	0.85
	VGG19+SVM (R, 1)	58.75	0.64	0.61	0.60
	VGG19+SVM (R, 5)	78.75	0.81	0.79	0.78
	VGG19+SVM (S, 1)	6.25	0.01	0.07	0.01
	VGG19+SVM (S, 5)	22.50	0.01	0.07	0.01
	VGG19+LR(L2)	**81.25**	0.83	0.81	0.81
	VGG19+CART	52.25	0.58	0.56	0.56
	VGG19+kNN	75.00	0.80	0.78	0.77
	VGG19+LDA	78.75	0.81	0.79	0.78
	VGG19+NB	51.25	0.53	0.51	0.51
ResNet 50	ResNet 50+RF (50)	52.50	0.59	0.53	0.53
	ResNet 50+RF (200)	55.00	0.59	0.54	0.53
	ResNet 50+RF (400)	55.00	0.65	0.55	0.57
	ResNet 50+RF (500)	56.25	0.65	0.56	0.58
	ResNet 50+RF (1000)	55.00	0.64	0.55	0.57
	ResNet 50+RF (4000)	55.00	0.64	0.55	0.57
	ResNet 50+SVM(L, 1)	62.50	0.70	0.65	0.66
	ResNet 50+SVM (L, 5)	62.50	0.69	0.66	0.66

	ResNet 50+SVM (R, 1)	26.25	0.31	0.23	0.19
	ResNet 50+SVM (R, 5)	41.25	0.60	0.38	0.37
	ResNet 50+SVM (S, 1)	23.75	0.37	0.20	0.16
	ResNet 50+SVM (S, 5)	31.25	0.31	0.25	0.21
	ResNet 50+LR(L2)	58.75	0.67	0.59	0.60
	ResNet 50+CART	42.50	0.49	0.42	0.43
	ResNet 50+kNN	46.25	0.54	0.53	0.53
	ResNet 50+LDA	23.75	0.31	0.26	0.26
	ResNet 50+NB	30.00	0.35	0.30	0.28

TABLE S8

PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS FEATURE EXTRACTOR WITH CONVENTIONAL CLASSIFIERS FOR THE 200X MAGNIFICATION FACTOR

Pre-trained Network	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16+RF (50)	59.21	0.64	0.61	0.60
	VGG16+RF (200)	67.11	0.69	0.68	0.67
	VGG16+RF (400)	69.74	0.71	0.70	0.69
	VGG16+RF (500)	69.74	0.71	0.70	0.69
	VGG16+RF (1000)	67.11	0.69	0.67	0.66
	VGG16+RF (4000)	68.42	0.71	0.68	0.67
	VGG16+SVM(L, 1)	**78.95**	0.81	0.82	0.81
	VGG16+SVM (L, 5)	**78.95**	0.81	0.82	0.81
	VGG16+SVM (R, 1)	56.58	0.57	0.55	54
	VGG16+SVM (R, 5)	**71.05**	0.69	0.68	0.67
	VGG16+SVM (S, 1)	15.79	0.01	0.09	0.02
	VGG16+SVM (S, 5)	25.00	0.01	0.09	0.02
	VGG16+LR(L2)	**88.58**	0.89	0.89	0.89
	VGG16+CART	40.79	0.43	0.41	0.41
	VGG16+kNN	**72.37**	0.78	0.74	0.74
	VGG16+LDA	**87.26**	0.88	0.86	0.86
	VGG16+NB	50.00	0.53	0.50	0.50
VGG19	VGG19+RF (50)	67.11	0.68	0.67	0.67
	VGG19+RF (200)	**71.05**	0.73	0.71	0.71
	VGG19+RF (400)	**71.05**	0.72	0.71	0.71
	VGG19+RF (500)	**69.74**	0.74	0.72	0.73
	VGG19+RF (1000)	**69.74**	0.71	0.70	0.70
	VGG19+RF (4000)	**69.74**	0.69	0.68	0.68
	VGG19+SVM(L, 1)	68.42	0.69	0.68	0.68
	VGG19+SVM (L, 5)	69.74	0.69	0.68	0.68
	VGG19+SVM (R, 1)	56.58	0.58	0.54	0.53
	VGG19+SVM (R, 5)	65.79	0.65	0.63	0.63
	VGG19+SVM (S, 1)	15.79	0.01	0.09	0.02
	VGG19+SVM (S, 5)	18.42	0.01	0.09	0.02
	VGG19+LR(L2)	69.74	0.71	0.70	0.70
	VGG19+CART	36.84	0.39	0.37	0.37
	VGG19+kNN	**72.37**	0.78	0.74	0.74
	VGG19+LDA	69.74	0.72	0.70	0.70
	VGG19+NB	53.95	0.52	0.54	0.51
ResNet 50	ResNet 50+RF (50)	40.79	0.44	0.42	0.41
	ResNet 50+RF (200)	43.42	0.43	0.43	0.42
	ResNet 50+RF (400)	47.37	0.53	0.47	0.47
	ResNet 50+RF (500)	46.05	0.57	0.46	0.45
	ResNet 50+RF (1000)	44.74	0.58	0.45	0.43
	ResNet 50+RF (4000)	43.42	0.49	0.43	0.42
	ResNet 50+SVM(L, 1)	53.95	0.61	0.54	0.53
	ResNet 50+SVM (L, 5)	52.63	0.55	0.53	0.51

	ResNet 50+SVM (R, 1)	22.37	0.13	0.22	0.14
	ResNet 50+SVM (R, 5)	21.05	0.16	0.22	0.16
	ResNet 50+SVM (S, 1)	23.68	0.11	0.20	0.11
	ResNet 50+SVM (S, 5)	21.05	0.14	0.22	0.14
	ResNet 50+LR(L2)	53.95	0.56	0.54	0.52
	ResNet 50+CART	26.32	0.26	0.26	0.26
	ResNet 50+kNN	36.84	0.45	0.45	0.43
	ResNet 50+LDA	22.37	0.13	0.22	0.14
	ResNet 50+NB	23.68	0.20	0.24	0.19

TABLE S9

C PERFORMANCE METRICS OF PRE-TRAINED NETWORK AS FEATURE EXTRACTOR WITH CONVENTIONAL CLASSIFIERS FOR THE 200X MAGNIFICATION FACTOR

Pre-trained Network	Model	Accuracy (in %)	Precision	Recall	F_1-score
VGG16	VGG16+RF (50)	62.50	0.65	0.64	0.64
	VGG16+RF (200)	68.06	0.66	0.67	0.66
	VGG16+RF (400)	**69.44**	0.70	0.69	0.69
	VGG16+RF (500)	**69.44**	0.70	0.69	0.69
	VGG16+RF (1000)	68.06	0.69	0.68	0.68
	VGG16+RF (4000)	68.06	0.69	68	68
	VGG16+SVM(L, 1)	**80.00**	0.80	0.79	0.79
	VGG16+SVM (L, 5)	**80.00**	0.80	0.79	0.79
	VGG16+SVM (R, 1)	51.39	0.52	0.51	0.51
	VGG16+SVM (R, 5)	62.50	0.62	0.61	0.61
	VGG16+SVM (S, 1)	26.39	0.01	0.08	0.01
	VGG16+SVM (S, 5)	38.89	0.01	0.08	0.01
	VGG16+LR(L2)	**69.44**	0.71	0.69	0.69
	VGG16+CART	41.67	0.46	0.42	0.42
	VGG16+kNN	**68.06**	0.75	0.72	0.72
	VGG16+LDA	62.50	0.64	62	0.62
	VGG16+NB	40.28	0.43	0.40	0.40
VGG19	VGG19+RF (50)	62.50	0.60	0.60	0.60
	VGG19+RF (200)	63.89	0.65	0.62	0.63
	VGG19+RF (400)	62.50	0.64	0.62	0.63
	VGG19+RF (500)	63.89	0.65	0.64	0.64
	VGG19+RF (1000)	**63.89**	0.65	0.64	0.64
	VGG19+RF (4000)	**63.89**	0.65	0.64	0.64
	VGG19+SVM(L, 1)	**69.44**	0.72	0.71	0.71
	VGG19+SVM (L, 5)	**66.67**	0.72	0.71	0.71
	VGG19+SVM (R, 1)	47.22	0.48	0.43	0.43
	VGG19+SVM (R, 5)	61.11	0.53	0.53	0.53
	VGG19+SVM (S, 1)	29.17	0.01	0.08	0.01
	VGG19+SVM (S, 5)	26.39	0.01	0.08	0.01
	VGG19+LR(L2)	**68.06**	0.69	0.68	0.68
	VGG19+CART	34.72	0.35	0.35	0.33
	VGG19+kNN	**63.89**	0.66	0.62	0.63
	VGG19+LDA	58.33	0.62	0.58	0.60
	VGG19+NB	34.72	0.36	0.35	0.35
ResNet 50	ResNet 50+RF (50)	36.11	0.36	0.35	0.33
	ResNet 50+RF (200)	31.94	0.36	0.33	0.32
	ResNet 50+RF (400)	33.33	0.36	0.33	0.32
	ResNet 50+RF (500)	33.33	0.38	0.33	0.33
	ResNet 50+RF (1000)	31.94	0.34	0.32	0.31
	ResNet 50+RF (4000)	34.72	0.35	0.35	0.33
	ResNet 50+SVM(L, 1)	34.72	0.34	0.38	0.35
	ResNet 50+SVM (L, 5)	34.72	0.35	0.38	0.35

	ResNet 50+SVM (R, 1)	16.67	0.03	0.12	0.05
	ResNet 50+SVM (R, 5)	26.39	0.11	0.21	0.14
	ResNet 50+SVM (S, 1)	15.28	0.07	0.12	0.07
	ResNet 50+SVM (S, 5)	22.22	0.16	0.21	0.14
	ResNet 50+LR(L2)	33.33	0.34	0.33	0.32
	ResNet 50+CART	30.56	0.32	0.31	0.30
	ResNet 50+kNN	25.00	0.25	0.24	0.24
	ResNet 50+LDA	16.67	0.15	0.14	0.14
	ResNet 50+NB	16.67	0.23	0.17	0.16

Fig. S1. ROC curve analysis for VGG16 (Baseline Model) at (a) 40X, (b) 100X, (c) 200X, and (d) 400X magnification factor.

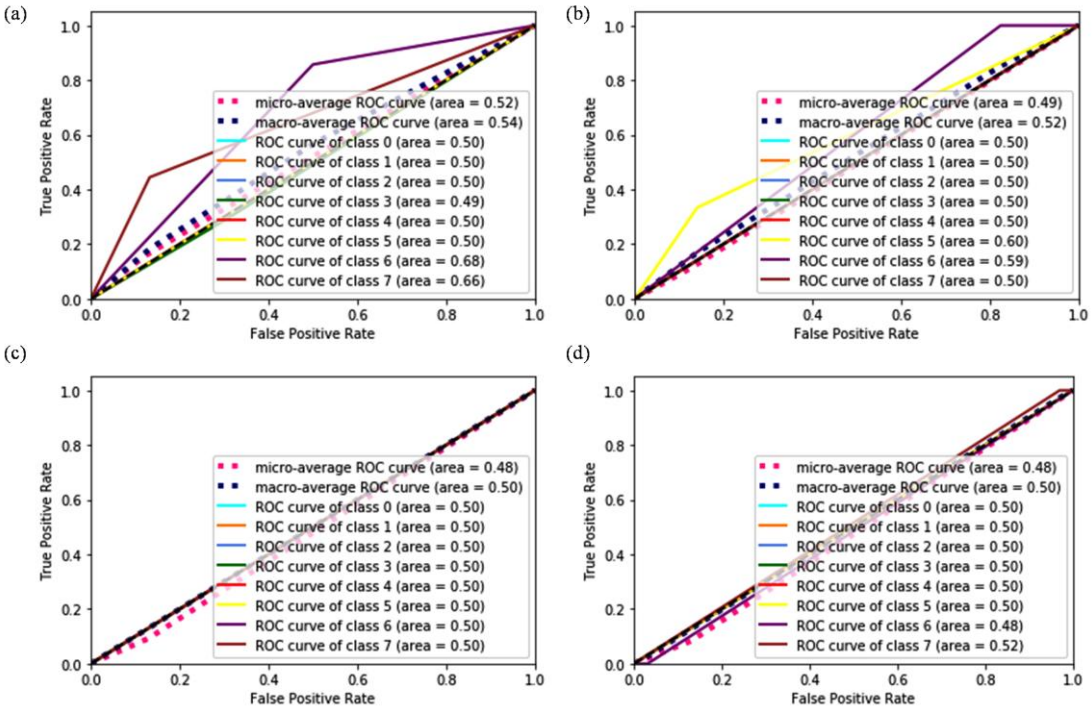
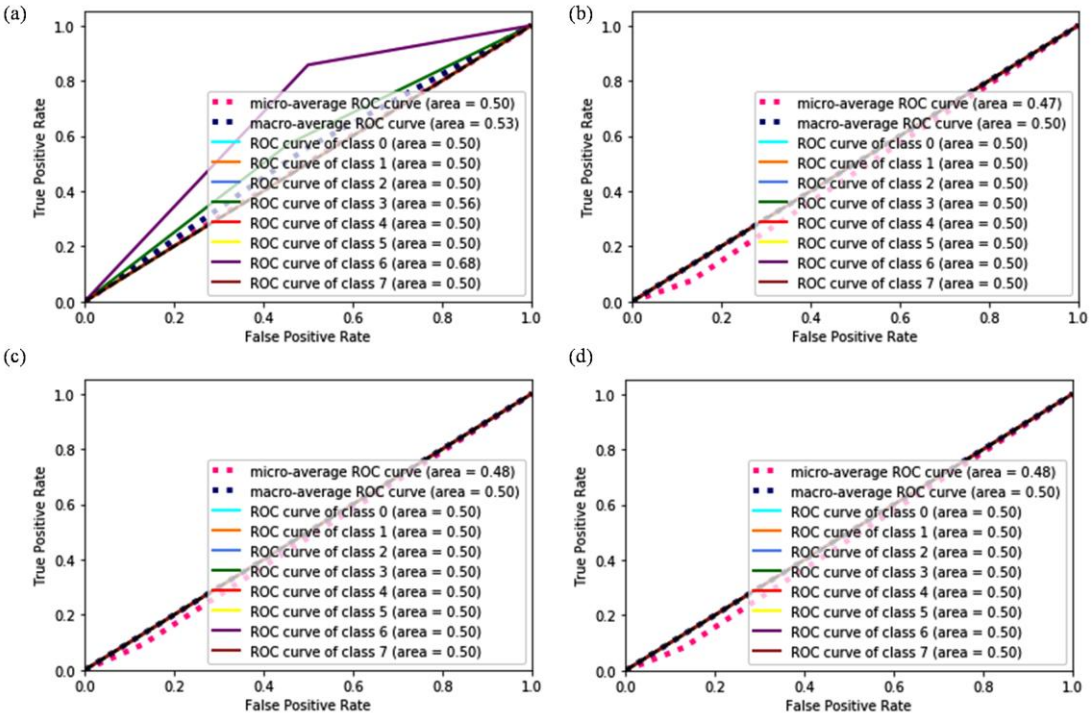
Fig. S2. ROC curve analysis for VGG19 (Baseline Model) at (a) 40X, (b) 100X, (c) 200X, and (d) 400X magnification factor