

Content:

Tables

- ✓ **Supplementary Table S1:** List of radiomic features derived from ADC and D maps
- ✓ **Supplementary Table S2:** Details about image pre-processing before feature extraction.
- ✓ **Supplementary Table S3.** Values of the most significant features for differentiating HPV-negative from HPV-positive patients, based on the Mann-Whitney test.
- ✓ **Supplementary Table S4** Comparison between the models in the validation set by the mid-p-value McNemar test

Figures

- ✓ **Supplementary Figure S1:** The effect of the ComBat-based feature harmonization on the validation set.
- ✓ **Supplementary Figure S2:** Comparison between accuracies obtained from different families of machine learning algorithms.

Supplementary Table S1. List of included Radiomic Features

Feature Family (Count)	Feature Name
<i>Intensity Direct</i> (39)	Energy, EnergyNorm, GlobalEntropy, GlobalMax, GlobalMean, GlobalMedian, GlobalMin, GlobalStd, GlobalUniformity, InterQuartileRange, Kurtosis, LocalEntropyMax, LocalEntropyMean, LocalEntropyMedian, LocalEntropyMin, LocalEntropyStd, LocalRangeMax, LocalRangeMean, LocalRangeMedian, LocalRangeMin, LocalRangeStd, LocalStdMax, LocalStdMean, LocalStdMedian, LocalStdMin, LocalStdStd, MeanAbsoluteDeviation, MedianAbsoluteDeviation, 5Percentile, 10Percentile, 25Percentile, 50Percentile, 75Percentile, 90Percentile, 95Percentile, Range, RootMeanSquare, Skewness, Variance
<i>Intensity Histogram</i> (13)	InterQuartileRange, Kurtosis, MeanAbsoluteDeviation, MedianAbsoluteDeviation, 5Percentile, 10Percentile, 25Percentile, 50Percentile, 75Percentile, 90Percentile, 95Percentile, Range, Skewness
<i>Grey-level co-occurrence matrix</i> (25)	JointMaximum, JointAverage, JointVariance, JointEntropy, DifferenceAverage, DifferenceVariance, DifferenceEntropy, SumAverage, SumVariance, SumEntropy, AngularSecondMoment, Contrast, Dissimilarity, InverseDifference, InverseDiffNorm, InverseDiffMoment, InverseDiffMomentNorm, InverseVariance, Correlation, AutoCorrelation, ClusterTendency, ClusterShade, ClusterProminence, InformationMeasureCor1, InformationMeasureCor2
<i>Gray Level Run Length Matrix</i> (16)	ShortRunEmphasis, LongRunEmphasis, LowGLRunEmpha, HighGLRunEmpha, ShortRunLowGLEmpha, ShortRunHighGLEmpha, LongRunLowGLEmpha, LongRunHighGLEmpha, GLNonuniformity, GLNonuniformityNorm, RLNonuniformity, RLNonuniformityNorm, RunPercentage,

	GrayLevelVariance, RunLengthVariance, RunEntropy
<i>Gray Level Distant Zone Matrix (16)</i>	SmallDistEmphasis, LargeDistEmphasis, LowGrayLevelEmpha, HighGrayLevelEmpha, SmallDistLowGLEmpha, SmallDistHighGLEmpha, LargeDistLowGLEmpha LargeDistHighGLEmpha, GLNonuniformity, GLNonuniformityNorm, ZDNonuniformity, ZDNonuniformityNorm, ZonePercentage, GrayLevelVariance, ZoneDistVariance, ZoneDistEntropy
<i>Gray Level Size Zone Matrix (16)</i>	SmallZoneEmphasis, LargeZoneEmphasis, LowGLZoneEmpha, HighGLZoneEmpha, SmallZoneLowGLEmpha SmallZoneHighGLEmpha, LargeZoneLowGLEmpha, LargeZoneHighGLEmpha, GLNonuniformity, GLNonuniformityNorm SZNonuniformity, SZNonuniformityNorm, ZonePercentage, GrayLevelVariance, SizeZoneVariance, SizeZoneEntropy
<i>Neighboring Gray Level Dependence (16)</i>	LowDepenEmphasis, HighDepenEmphasis, LowGLCountEmpha, HighGLCountEmpha LowDepenLowGLEmpha, LowDepenHighGLEmpha, HighDepenLowGLEmpha, HighDepenHighGLEmpha GLNonuniformity, GLNonuniformityNorm, DepCountNonuniformity, DepCountNonuniformityNorm GLVariance, DepCountVariance, DepCountEntropy, DepCountEnergy
<i>Morphology (16)</i>	Volume, ApproximateVolume, SurfaceArea, SurfaceToVolumeRatio, Compactness1, Compactness2 SphericalDisproportion, Sphericity, Asphericity, CentreOfMassShift, Maximum3DDiameter MajorAxisLength, MinorAxisLength, LeastAxisLength, Elongation, Flatness

Supplementary Table S2. Details about pre-processing before feature extraction

Feature Family	Interpolation	Re-segmentation	Intensity Discretization
<i>Intensity Direct</i>	none	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.1 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>Intensity Histogram</i>	none	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.1 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>GLCM</i>	3D ($3 \times 3 \times 3 \text{ mm}^3$)*	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.1 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>GLRLM</i>	3D ($3 \times 3 \times 3 \text{ mm}^3$) *	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.1 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>GLSZM</i>	3D ($3 \times 3 \times 3 \text{ mm}^3$)	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.35 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>GLDZM</i>	3D ($3 \times 3 \times 3 \text{ mm}^3$)	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.35 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>NGLD</i>	3D ($3 \times 3 \times 3 \text{ mm}^3$)	Range: $[0.25-3] \times 10^{-3} \text{ mm}^2/\text{s}$	FBS ($0.1 \times 10^{-3} \text{ mm}^2/\text{s}$)
<i>Morphology</i>	none	none	none

*Aggregation method: 3D average; Directions: [0 1 2 3 4 5 6 7 8 9 10 11 12]. A detailed description of radiomic features is reported in IBSI reference manual (Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. arXiv preprint arXiv:1612.07003].

Abbreviations: FBS, fixed bin size; GLCM, Grey Level Co-occurrence Matrix ; GLRLM, Gray Level Run Length Matrix; GLSZM, Gray Level Size Zone Matrix; GLDZM, Gray Level Distant Zone Matrix; NGLD, Neighboring Gray Level Dependence.

Supplementary Table S3. Values of the most significant features ($p < 0.01$) for differentiating HPV-negative from HPV-positive patients on the training cohort, based on the Mann-Whitney test.

Features (263)	HPV-negative		HPV-positive		P value
	Median	IQR*	Median	IQR	
EnergyNorm_ADCPT	2071292	921556	1371320	569555	<0.001
GlobalEntropy_ADCPT	6.14	0.59	5.94	0.38	0.002
GlobalMean_ADCPT	1409	290	1143	231	<0.001
GlobalMedian_ADCPT	1357	274	1118	233	<0.001
GlobalStd_ADCPT	330	146	267	70	0.001
GlobalUniformity_ADCPT	0.02	0.01	0.02	0.01	0.002
InterQuartileRange_ADCPT	427	153	330	105	0.002
MeanAbsoluteDeviation_ADCPT	259	112	209	61	0.001
MedianAbsoluteDeviation_ADCPT	209	81	161	51	0.001
5Percentile_ADCPT	910	179	757	235	0.003
10Percentile_ADCPT	1020	206	822	201	0.001
25Percentile_ADCPT	1150	270	975	189	<0.001
50Percentile_ADCPT	1357	274	1118	233	<0.001
75Percentile_ADCPT	1546	346	1309	262	<0.001
90Percentile_ADCPT	1782	398	1512	310	<0.001
95Percentile_ADCPT	1953	492	1628	320	<0.001
RootMeanSquare_ADCPT	1439	318	1171	238	<0.001
Skewness_ADCPT	0.17	0.61	0.61	0.50	0.001
Variance_ADCPT	109076	99425	71421	37367	0.001
InterQuartileRange_ADCPT	429	220	330	105	0.002
MeanAbsoluteDeviation_ADCPT	265	131	210	66	0.001
MedianAbsoluteDeviation_ADCPT	211	91	162	53	0.002
10Percentile_ADCPT	961	195	818	199	0.006
25Percentile_ADCPT	1150	265	964	195	<0.001
50Percentile_ADCPT	1357	277	1118	235	<0.001
75Percentile_ADCPT	1546	349	1305	246	<0.001
90Percentile_ADCPT	1774	408	1511	306	<0.001
95Percentile_ADCPT	1945	492	1628	320	<0.001
Skewness_ADCPT	-0.05	0.76	0.38	0.66	0.010
EnergyNorm_DPT	1560916	639489	941143	474717	<0.001
GlobalEntropy_DPT	6.16	0.69	5.81	0.60	0.003
GlobalMean_DPT	1203	229	934	240	<0.001
GlobalMedian_DPT	1152	248	910	242	<0.001
GlobalStd_DPT	323	135	244	104	0.006
GlobalUniformity_DPT	0.02	0.01	0.02	0.01	0.005
InterQuartileRange_DPT	404	202	296	154	0.007
MeanAbsoluteDeviation_DPT	254	107	189	88	0.006
MedianAbsoluteDeviation_DPT	204	107	145	78	0.004
10Percentile_DPT	836	226	664	263	0.001
25Percentile_DPT	988	215	769	261	<0.001
50Percentile_DPT	1152	248	910	242	<0.001
75Percentile_DPT	1413	297	1066	283	<0.001
90Percentile_DPT	1620	375	1280	365	<0.001
95Percentile_DPT	1698	419	1385	396	<0.001

RootMeanSquare _DPT	1249	258	970	243	<0.001
Skewness _DPT	0.41	0.60	0.74	0.54	0.004
Variance _DPT	104574	81226	59464	51980	0.006
MedianAbsoluteDeviation _DPT	227	130	150	93	0.008
25Percentile _DPT	958	297	737	290	0.001
50Percentile _DPT	1151	263	883	272	<0.001
75Percentile _DPT	1409	292	1051	261	<0.001
90Percentile _DPT	1590	355	1279	336	<0.001
95Percentile _DPT	1698	407	1376	381	<0.001
Skewness _DPT	-0.19	0.99	0.23	0.86	0.009
EnergyNorm _ADCLN	2046561	1635080	1330094	626770	<0.001
GlobalEntropy _ADCLN	6.32	0.69	5.93	0.42	<0.001
GlobalMax _ADCLN	2783	594	2257	552	0.001
GlobalMean _ADCLN	1384	482	1101	275	<0.001
GlobalMedian _ADCLN	1371	393	1056	284	<0.001
GlobalStd _ADCLN	372	181	284	93	<0.001
GlobalUniformity _ADCLN	0.016	0.008	0.020	0.004	<0.001
InterQuartileRange _ADCLN	467	278	337	134	<0.001
LocalEntropyMean _ADCLN	5.23	0.25	5.09	0.23	0.005
LocalEntropyMedian _ADCLN	5.35	0.28	5.13	0.25	0.001
LocalRangeMean _ADCLN	984	289	750	261	0.002
LocalRangeMedian _ADCLN	876	408	644	288	0.001
LocalStdMean _ADCLN	281	84	207	80	0.001
LocalStdMedian _ADCLN	245	105	175	85	0.001
LocalStdStd _ADCLN	149	77	132	51	0.010
MeanAbsoluteDeviation _ADCLN	293	156	219	77	<0.001
MedianAbsoluteDeviation _ADCLN	205	149	161	46	<0.001
10Percentile _ADCLN	943	270	818	243	0.008
25Percentile _ADCLN	1145	303	917	233	<0.001
50Percentile _ADCLN	1371	393	1056	284	<0.001
75Percentile _ADCLN	1559	631	1253	317	<0.001
90Percentile _ADCLN	1837	683	1497	353	<0.001
95Percentile _ADCLN	2024	622	1643	356	<0.001
Range _ADCLN	2386	666	1839	561	0.001
RootMeanSquare _ADCLN	1430	541	1153	269	<0.001
Skewness _ADCLN	0.41	0.61	0.84	0.62	0.007
Variance _ADCLN	138741	140336	80682	54523	<0.001
InterQuartileRange _ADCLN	485	260	345	153	<0.001
MeanAbsoluteDeviation _ADCLN	314	157	227	90	<0.001
MedianAbsoluteDeviation _ADCLN	214	141	164	61	<0.001
25Percentile _ADCLN	1108	308	912	244	0.009
50Percentile _ADCLN	1320	394	1037	298	<0.001
75Percentile _ADCLN	1554	657	1242	315	<0.001
90Percentile _ADCLN	1837	708	1485	352	<0.001
95Percentile _ADCLN	2004	626	1642	360	<0.001
Range _ADCLN	2783	816	2065	760	0.001
EnergyNorm _DLN	1525907	836340	925946	472561	<0.001
GlobalEntropy _DLN	6.16	0.57	5.70	0.55	<0.001
GlobalMax _DLN	2515	818	1928	564	<0.001

GlobalMean_DLN	1169	308	944	247	<0.001
GlobalMedian_DLN	1190	334	917	265	<0.001
GlobalStd_DLN	359	190	233	116	<0.001
GlobalUniformity_DLN	0.017	0.007	0.023	0.009	<0.001
InterQuartileRange_DLN	454	233	295	156	<0.001
LocalEntropyMax_DLN	5.88	0.24	5.75	0.20	0.007
LocalEntropyMean_DLN	5.12	0.28	4.86	0.33	<0.001
LocalEntropyMedian_DLN	5.17	0.33	4.91	0.28	<0.001
LocalRangeMax_DLN	2597	911	1996	588	0.008
LocalRangeMean_DLN	841	506	615	347	0.001
LocalRangeMedian_DLN	782	480	532	384	<0.001
LocalRangeStd_DLN	398	253	331	130	0.008
LocalStdMax_DLN	822	470	607	272	0.005
LocalStdMean_DLN	247	142	172	99	0.001
LocalStdMedian_DLN	217	128	145	90	<0.001
LocalStdStd_DLN	123	81	97	50	0.005
MeanAbsoluteDeviation_DLN	272	153	180	91	<0.001
MedianAbsoluteDeviation_DLN	219	121	140	70	<0.001
25Percentile_DLN	951	282	772	221	0.002
50Percentile_DLN	1190	334	917	265	<0.001
75Percentile_DLN	1352	458	1069	285	<0.001
90Percentile_DLN	1635	581	1278	308	<0.001
95Percentile_DLN	1841	840	1407	303	<0.001
Range_DLN	2149	770	1543	555	<0.001
RootMeanSquare_DLN	1235	333	962	250	<0.001
Variance_DLN	129395	144447	54114	60061	<0.001
InterQuartileRange_DLN	514	297	318	181	<0.001
MeanAbsoluteDeviation_DLN	328	136	203	133	<0.001
MedianAbsoluteDeviation_DLN	250	147	152	85	<0.001
50Percentile_DLN	1174	353	894	237	0.001
75Percentile_DLN	1342	453	1055	288	<0.001
90Percentile_DLN	1633	614	1269	279	<0.001
95Percentile_DLN	1812	847	1405	305	<0.001
Range_DLN	2394	871	1843	483	0.001
JointAverage_ADCPT	8.84	3.21	6.30	2.64	0.002
JointVariance_ADCPT	6.21	7.22	5.24	2.93	0.002
DifferenceAverage_ADCPT	2.15	1.20	1.87	0.63	0.007
SumAverage_ADCPT	18	6	13	5	0.002
SumVariance_ADCPT	19	20	14	9	0.002
Contrast_ADCPT	7.85	10.13	6.16	3.77	0.008
Dissimilarity_ADCPT	2.15	1.20	1.87	0.63	0.007
InverseDifference_ADCPT	0.44	0.11	0.48	0.08	0.004
InverseDiffMoment_ADCPT	0.37	0.13	0.42	0.10	0.005
InverseVariance_ADCPT	0.36	0.11	0.39	0.05	0.004
AutoCorrelation_ADCPT	81	56	41	36	0.002
ClusterTendendcy_ADCPT	19	20	14	9	0.002
ClusterProminence_ADCPT	1364	2565	752	768	0.004
InformationMeasureCor2_ADCPT	0.81	0.16	0.74	0.16	0.009
GLNonuniformityNorm_ADCPT	0.08	0.04	0.10	0.03	0.007

GrayLevelVariance _ADCPT	15.69	7.68	11.21	4.66	0.001
ShortRunEmphasis _ADCPT	0.93	0.05	0.90	0.05	0.010
HighGLRunEmpha _ADCPT	91	68	52	41	0.003
ShortRunHighGLEmpha _ADCPT	85	57	47	37	0.002
LongRunHighGLEmpha _ADCPT	125	84	78	45	0.007
GrayLevelVariance _ADCPT	8.43	7.15	6.49	3.89	0.004
GLNonuniformityNorm _ADCPT	0.24	0.09	0.31	0.11	0.001
GrayLevelVariance _ADCPT	2.15	1.38	1.47	0.89	0.005
HighDepenEmphasis _ADCPT	16	19	27	24	0.009
HighGLCountEmpha _ADCPT	89	65	48	40	0.002
HighDepenLowGLEmpha _ADCPT	0.51	0.62	1.09	1.81	0.001
GLVariance _ADCPT	7.64	7.47	6.24	3.65	0.004
DepCountVariance _ADCPT	4.00	6.70	7.49	6.41	0.004
JointAverage _DPT	8.11	3.29	5.79	1.85	<0.001
SumAverage _DPT	16	7	12	4	<0.001
AutoCorrelation _DPT	68	52	37	22	<0.001
LargeDistHighGLEmpha _DPT	114	73	75	38	0.007
GrayLevelVariance _DPT	13	8	10	5	0.009
HighGLRunEmpha _DPT	73	50	46	26	0.001
ShortRunHighGLEmpha _DPT	66	50	40	25	0.001
LongRunLowGLEmpha _DPT	0.051	0.050	0.090	0.068	0.001
LongRunHighGLEmpha _DPT	101	69	66	40	0.002
GLNonuniformityNorm _DPT	0.27	0.10	0.32	0.14	0.004
GrayLevelVariance _DPT	1.80	1.00	1.28	0.93	0.003
HighGLCountEmpha _DPT	71	53	43	27	0.001
LowDepenHighGLEmpha _DPT	18	24	11	12	0.008
HighDepenLowGLEmpha _DPT	0.53	1.27	1.42	1.72	<0.001
DepCountVariance _DPT	4.19	6.60	9.12	9.96	0.008
JointMaximum _ADCLN	0.04	0.04	0.07	0.03	0.002
JointAverage _ADCLN	8.49	5.71	6.38	2.71	<0.001
JointVariance _ADCLN	9.42	10.70	6.13	4.74	<0.001
JointEntropy _ADCLN	6.25	1.23	5.63	0.92	0.001
DifferenceAverage _ADCLN	2.65	1.47	1.93	0.88	<0.001
DifferenceVariance _ADCLN	5.06	6.17	2.89	2.14	<0.001
DifferenceEntropy _ADCLN	2.82	0.73	2.46	0.40	<0.001
SumAverage _ADCLN	17	11	13	5	<0.001
SumVariance _ADCLN	28	30	17	13	0.002
SumEntropy _ADCLN	4.17	0.68	3.79	0.53	0.001
AngularSecondMoment _ADCLN	0.017	0.016	0.030	0.014	<0.001
Contrast _ADCLN	11.9	15.5	6.8	5.7	<0.001
Dissimilarity _ADCLN	2.65	1.47	1.93	0.88	<0.001
InverseDifference _ADCLN	0.39	0.06	0.48	0.10	<0.001
InverseDiffMoment _ADCLN	0.30	0.07	0.41	0.12	<0.001
InverseVariance _ADCLN	0.31	0.08	0.39	0.09	<0.001
AutoCorrelation _ADCLN	80	131	44	34	<0.001
ClusterTendendcy _ADCLN	28	30	17	13	0.002
ClusterProminence _ADCLN	2307	5508	1195	1781	0.004
HighGrayLevelEmpha _ADCLN	115	125	71	46	0.001
SmallDistHighGLEmpha _ADCLN	102	98	66	46	0.002

LargeDistHighGLEmpa _ADCLN	137	297	92	84	0.004
GLNonuniformityNorm _ADCLN	0.07	0.02	0.09	0.04	<0.001
ZonePercentage _ADCLN	0.34	0.17	0.24	0.23	0.004
GrayLevelVariance _ADCLN	18	14	12	7	0.001
ZoneDistEntropy _ADCLN	4.3	0.9	3.9	0.7	0.004
ShortRunEmphasis _ADCLN	0.9	0.0	0.9	0.1	0.005
HighGLRunEmpha _ADCLN	92	134	52	40	<0.001
ShortRunHighGLEmpa _ADCLN	86	125	47	38	<0.001
LongRunLowGLEmpa _ADCLN	0.05	0.06	0.08	0.13	0.004
LongRunHighGLEmpa _ADCLN	125	184	69	48	0.001
GLNonuniformityNorm _ADCLN	0.09	0.04	0.12	0.04	<0.001
RLNonuniformityNorm _ADCLN	0.83	0.07	0.78	0.10	0.005
RunPercentage _ADCLN	0.91	0.04	0.88	0.07	0.008
GrayLevelVariance _ADCLN	12	12	7	5	<0.001
RunEntropy _ADCLN	4.13	0.60	3.88	0.40	0.004
LargeZoneLowGLEmpa _ADCLN	88	382	590	1100	0.009
GLNonuniformityNorm _ADCLN	0.23	0.09	0.29	0.13	0.001
GrayLevelVariance _ADCLN	2.72	1.53	1.41	1.34	<0.001
LowDepenEmphasis _ADCLN	0.28	0.14	0.21	0.14	0.002
HighGLCountEmpha _ADCLN	89	134	50	40	<0.001
LowDepenHighGLEmpa _ADCLN	31	37	15	19	<0.001
HighDepenLowGLEmpa _ADCLN	0.43	0.42	1.01	2.02	0.001
GLNonuniformityNorm _ADCLN	0.10	0.04	0.13	0.05	<0.001
GLVariance _ADCLN	11	12	7	5	<0.001
JointMaximum _DLN	0.05	0.03	0.08	0.06	<0.001
JointAverage _DLN	7.61	3.52	5.81	1.96	0.001
JointVariance _DLN	8.39	9.68	4.15	4.38	<0.001
JointEntropy _DLN	6.08	1.14	5.19	1.08	<0.001
DifferenceAverage _DLN	2.57	1.00	1.66	0.79	<0.001
DifferenceVariance _DLN	3.61	3.81	2.16	1.54	<0.001
DifferenceEntropy _DLN	2.68	0.42	2.27	0.46	<0.001
SumAverage _DLN	15	7	12	4	0.001
SumVariance _DLN	24	32	12	15	<0.001
SumEntropy _DLN	4.08	0.78	3.54	0.52	<0.001
AngularSecondMoment _DLN	0.020	0.017	0.035	0.031	<0.001
Contrast _DLN	10.40	8.83	4.95	4.49	<0.001
Dissimilarity _DLN	2.57	1.00	1.66	0.79	<0.001
InverseDifference _DLN	0.40	0.08	0.50	0.10	<0.001
InverseDiffMoment _DLN	0.32	0.09	0.44	0.13	<0.001
InverseVariance _DLN	0.33	0.09	0.41	0.09	<0.001
AutoCorrelation _DLN	66	52	36	26	0.001
ClusterTendendcy _DLN	24	32	11	15	<0.001
ClusterProminence _DLN	1802	10840	504	1511	0.001
InformationMeasureCor2 _DLN	0.89	0.19	0.76	0.21	0.008
HighGrayLevelEmpha _DLN	82	80	54	29	0.004
SmallDistHighGLEmpa _DLN	69	73	46	28	0.006
GLNonuniformityNorm _DLN	0.08	0.03	0.11	0.04	<0.001
ZonePercentage _DLN	0.32	0.19	0.18	0.17	0.006
GrayLevelVariance _DLN	14	13	10	6	<0.001

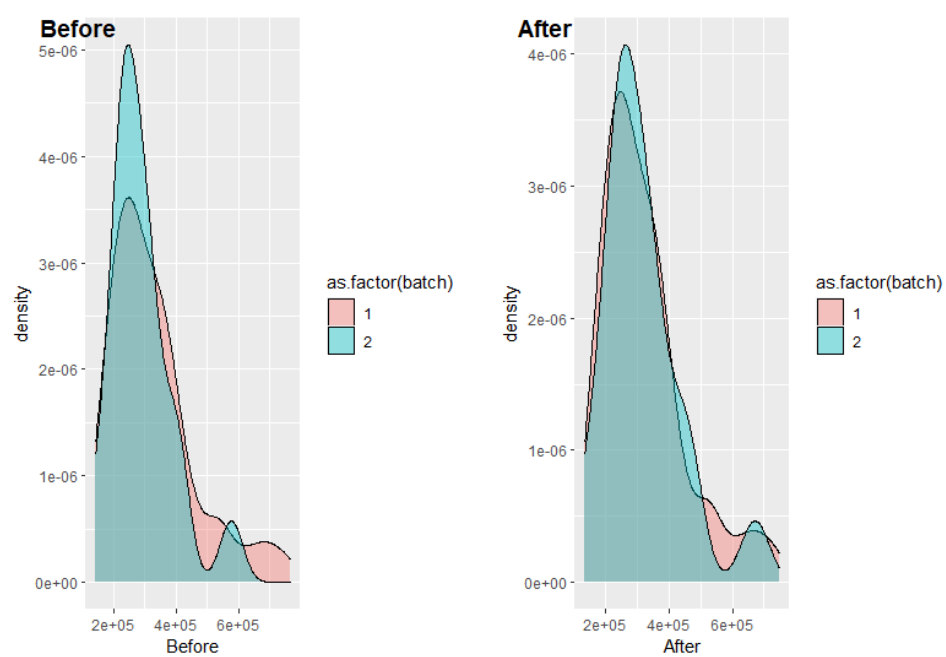
ZoneDistEntropy _DLN	4.15	1.01	3.57	0.96	0.004
ShortRunEmphasis _DLN	0.93	0.04	0.89	0.06	0.002
LongRunEmphasis _DLN	1.32	0.23	1.58	0.45	0.003
HighGLRunEmpha _DLN	77	45	40	28	<0.001
ShortRunHighGLEmpha _DLN	72	43	36	27	<0.001
LongRunHighGLEmpha _DLN	99	64	59	37	0.007
GLNonuniformityNorm _DLN	0.10	0.04	0.14	0.05	<0.001
RLNonuniformityNorm _DLN	0.84	0.10	0.75	0.13	0.003
RunPercentage _DLN	0.91	0.06	0.85	0.08	0.003
GrayLevelVariance _DLN	9.4	10.6	5.0	5.4	<0.001
RunLengthVariance _DLN	0.11	0.09	0.20	0.17	0.004
RunEntropy _DLN	4.04	0.55	3.63	0.48	0.002
HighGLZoneEmpha _DLN	10.1	7.2	7.7	3.9	0.007
GLNonuniformityNorm _DLN	0.25	0.09	0.33	0.13	<0.001
GrayLevelVariance _DLN	2.24	1.46	1.22	1.05	0.002
LowDepenEmphasis _DLN	0.26	0.15	0.17	0.11	0.003
HighDepenEmphasis _DLN	15	12	29	29	0.005
HighGLCountEmpha _DLN	74	44	39	30	<0.001
LowDepenHighGLEmpha _DLN	20	21	8	10	<0.001
HighDepenLowGLEmpha _DLN	0.54	1.08	1.50	2.63	0.001
GLNonuniformityNorm _DLN	0.10	0.04	0.15	0.07	<0.001
DepCountNonuniformityNorm _DLN	0.17	0.09	0.12	0.08	0.006
GLVariance _DLN	9.2	10.4	4.5	5.4	<0.001

*Diffusion coefficients, ADC and D, are expressed in 10^{-6} mm²/s; Abbreviation: IQR, interquartile range

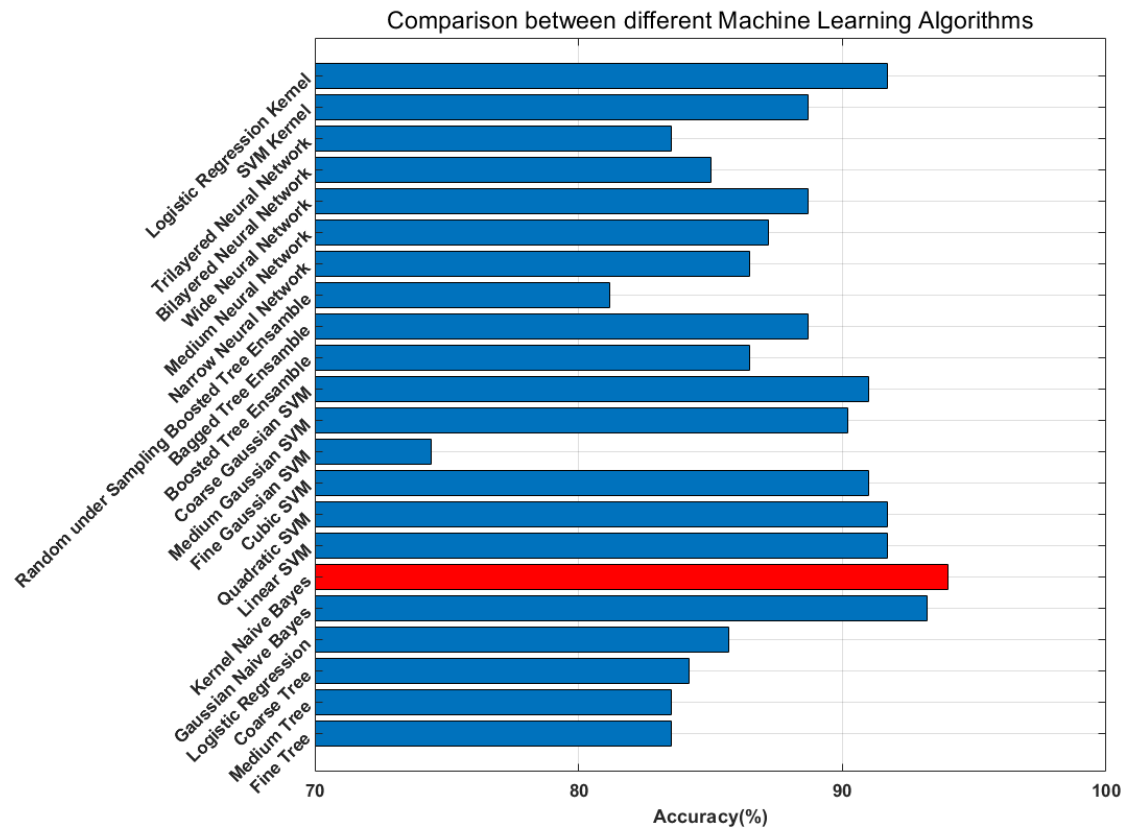
Supplementary Table S4. Comparison between the performances of the proposed models in the validation set

Model	1	2	3	4	5	6
1	1.00	1.00	0.69	0.34	1.00	0.14
2	1.00	1.00	0.50	0.22	1.00	0.09
3	0.69	0.50	1.00	0.38	0.73	0.15
4	0.34	0.22	0.38	1.00	0.29	0.39
5	1.00	1.00	0.73	0.29	1.00	0.17
6	0.14	0.09	0.15	0.39	0.17	1.00

P-values refer to the mid-p-value McNemar test.



Supplementary Figure S1. The effect of the ComBat-based feature harmonization on the validation set, which were in part examined on the 1.5T system ($n = 30$) and in part on the 3T system ($n = 19$).



Supplementary Figure S2. Comparison between accuracies obtained from different families of machine learning algorithms.