Supplementary Information for:

Hyperspectral Imaging Benchmark based on Machine Learning for Intraoperative Brain Tumour Detection

Raquel Leon[1,+], Himar Fabelo[2,1*,+], Samuel Ortega[3,1] , Ines A. Cruz-Guerrero[4], Daniel Ulises Campos-Delgado[4], Adam Szolna[5], Juan F. Piñeiro[5], Carlos Espino[5], Aruma J. O'Shanahan[5], Maria Hernandez[5], David Carrera[5], Sara Bisshopp[5], Coralia Sosa[5], Francisco J. Balea-Fernandez[6,1], Jesus Morera[5], Bernardino Clavo[7,2], and Gustavo M. Callico[1]

[1]Research Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain.
[2]Fundación Instituto de Investigación Sanitaria de Canarias (FIISC), Las Palmas de Gran Canaria, Spain.
[3]Nofima, Norwegian Institute of Food Fisheries and Aquaculture Research, Tromsø, Norway.
[4]Faculty of Sciences, Universidad Autonoma de San Luis Potosi, San Luis Potosi, Mexico.
[5]Department of Neurosurgery, University Hospital Doctor Negrin of Gran Canaria, Las Palmas de Gran Canaria, Spain.
[6]Department of Psychology, Sociology and Social Work, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain.
[7]Research Unit, University Hospital Doctor Negrin of Gran Canaria, Las Palmas de Gran Canaria, Spain.
[*]e-mail: hfabelo@iuma.ulpgc.es; [+]these authors contributed equally to this work.

List of Supplementary Figures

List of Supplementary Tables

Supplementary Figures

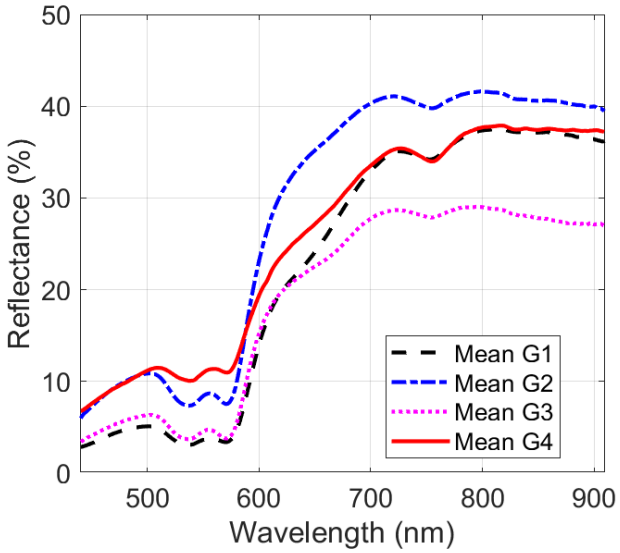

Fig. 1 Average spectral signature of the entire labelled dataset for each primary tumour grade. *G1: Grade 1; G2: Grade 2; G3: Grade 3; G4: Grade 4.*

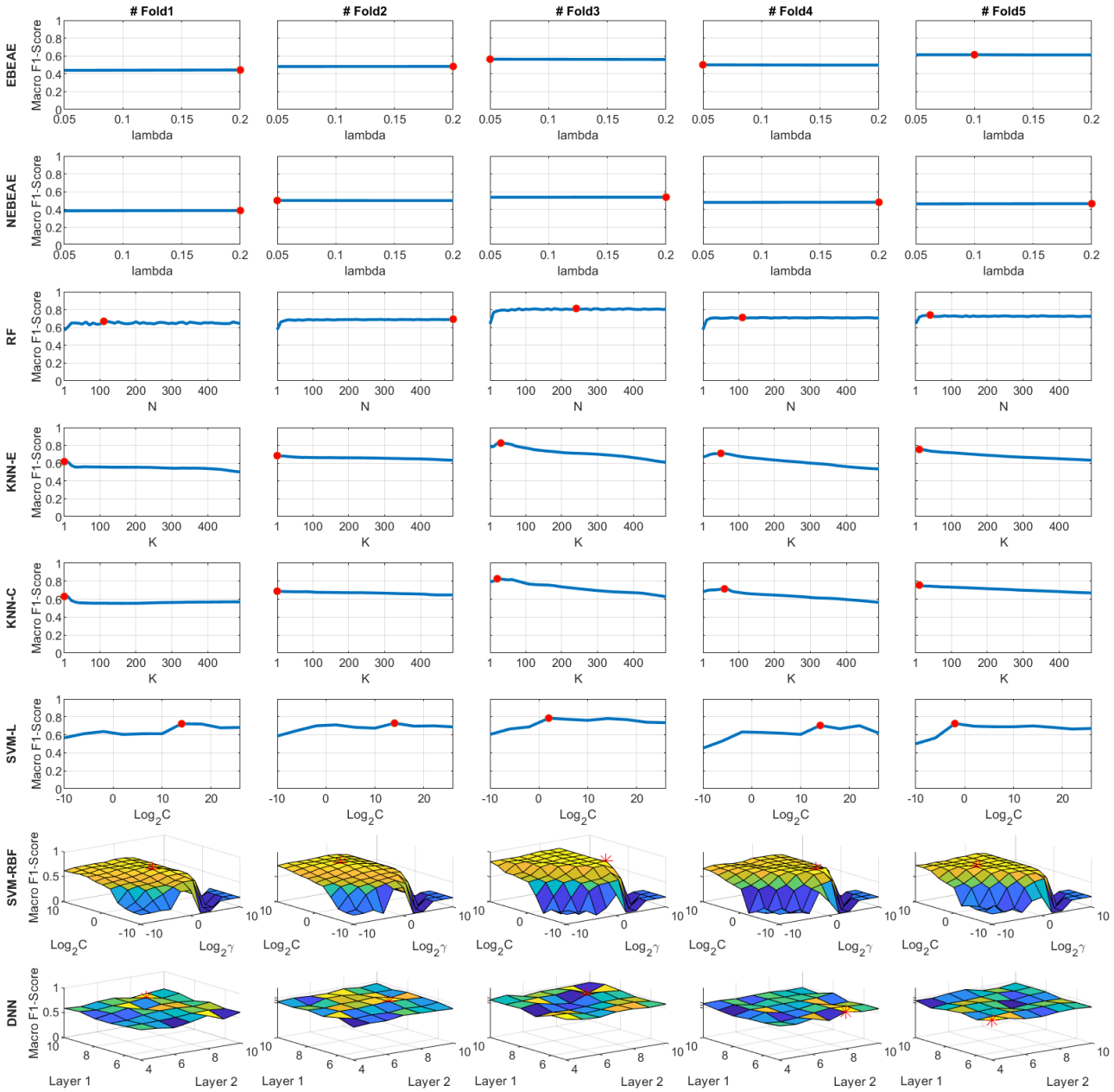

Fig. 2 Coarse search representation of the macro F1-Score results using the data reduction of 1,000 pixels per class for the hyperparameter optimization of each classification algorithm using the validation set.

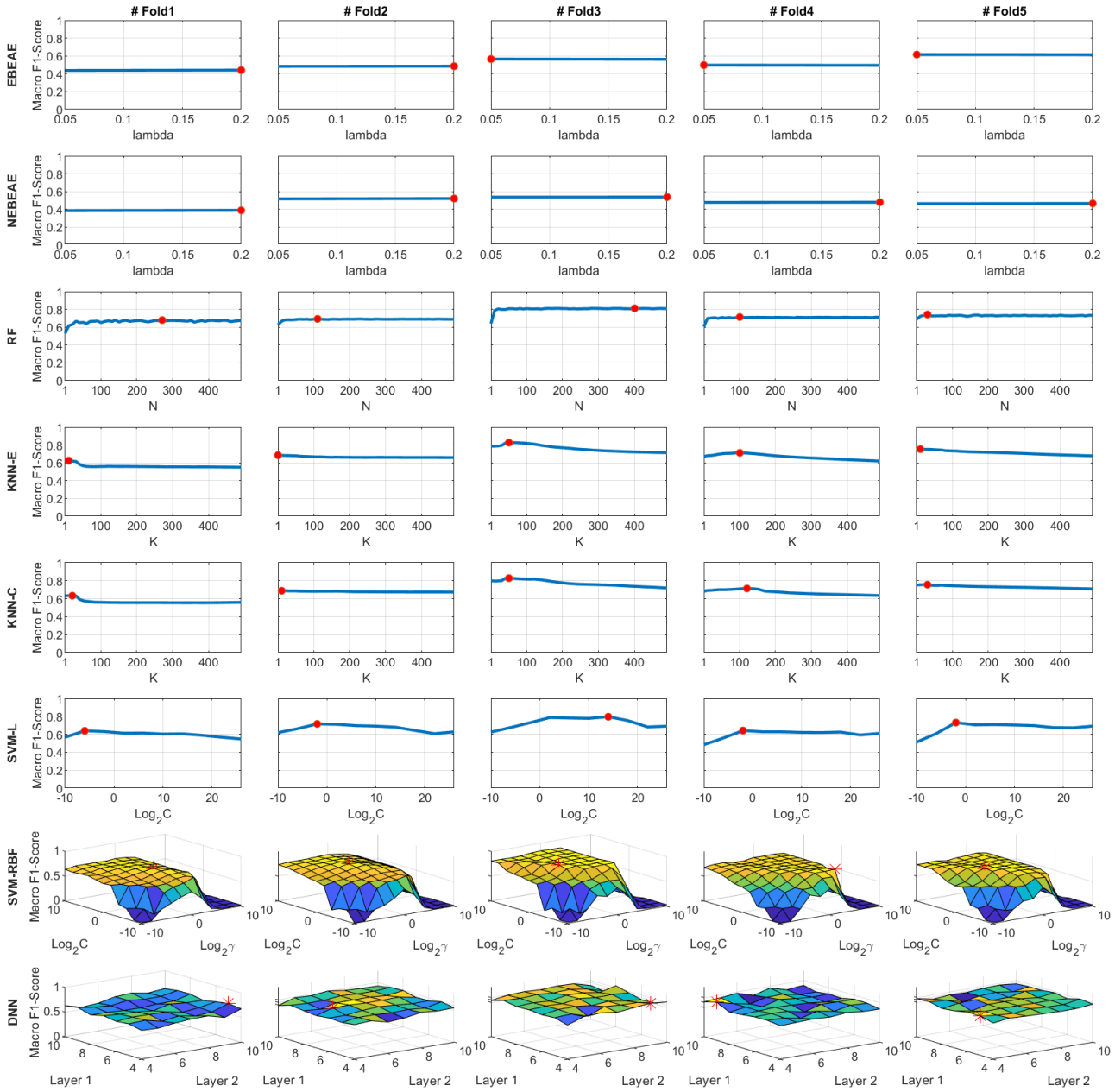

Fig. 3 Coarse search representation of the macro F1-Score results using the data reduction of 2,000 pixels per class for the hyperparameter optimization of each classification algorithm using the validation set.

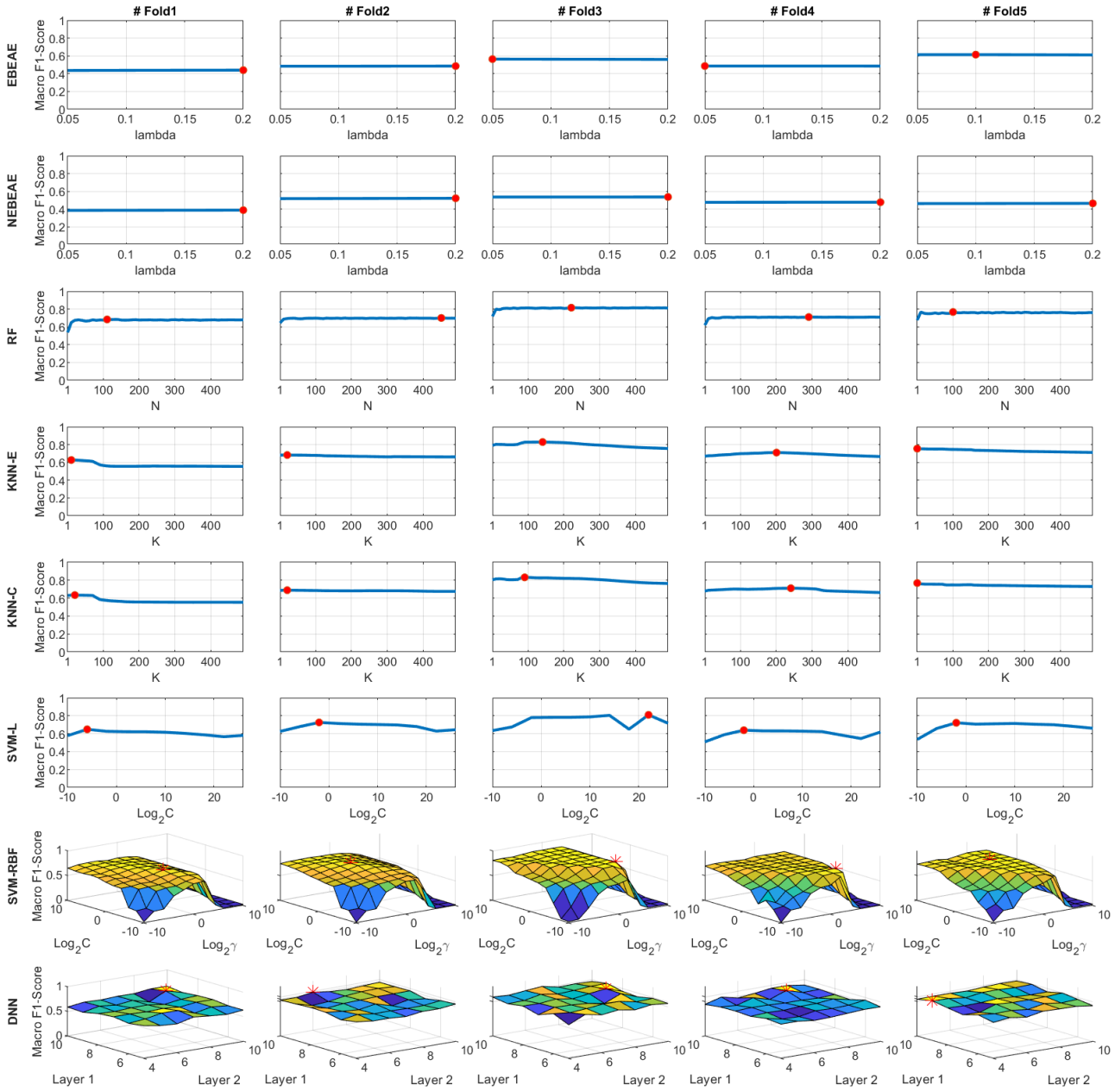

Fig. 4 Coarse search representation of the macro F1-Score results using the data reduction of 4,000 pixels per class for the hyperparameter optimization of each classification algorithm using the validation set.

Supplementary Tables

Table 1 Summary of the in-vivo HS human brain image database, including number of pixels labelled of each class, spatial and spectral dimension, and the diagnosis. *NT: Normal; TT: Tumour; BV: Blood vessel; BG: Background; S: Secondary; P: Primary; G1: Grade 1; G2: Grade 2; G3: Grade 3; G4: Grade 4; DL: Deep layer tumour.*

Campaign	Image ID	Size (height×width×bands)	#Labeled Pixels				Diagnosis
			NT	TT	BV	BG	
First Data Campaign	Op04C2	389×345×826	5,007	0	965	1,992	Normal Brain
	Op05C1	483×488×826	6,061	0	1,727	20,483	Renal Carcinoma (S)
	Op07C1	582×400×826	7,714	0	1,089	0	Normal Brain
	Op08C1	460×549×826	2,295	1,221	1,331	630	G4 Glioblastoma (P)
	Op08C2	480×553×826	2,187	138	1,000	7,444	G4 Glioblastoma (P)
	Op10C3	371×461×826	10,626	0	2,332	3,972	G4 Glioblastoma (P)
	Op12C1	443×497×826	4,516	855	8,697	1,685	G4 Glioblastoma (P)
	Op12C2	445×498×826	6,553	3,139	6,041	8,731	G4 Glioblastoma (P)
	Op13C1	298×253×826	1,827	0	129	589	Lung Carcinoma (S)
	Op14C1	317×244×826	0	30	64	1,866	G4 Glioblastoma (P)
	Op15C1	376×494×826	1,251	2,046	4,089	696	G4 Glioblastoma (P)
	Op16C1	335×323×826	3,970	0	246	12,002	Normal Brain
	Op16C2	335×326×826	349	0	0	2,767	Normal Brain
	Op16C3	315×321×826	603	0	234	1,696	Normal Brain
	Op16C4	383×297×826	1,178	0	1,064	956	G4 Glioblastoma (P)
	Op16C5	414×292×826	2,643	0	452	5,125	G4 Glioblastoma (P)
	Op17C1	441×399×826	1,328	0	68	3,069	G4 Glioblastoma (P)
	Op18C1	479×462×826	13,450	0	488	9,773	G1 Ganglioglioma (P)
	Op18C2	510×434×826	4,813	0	958	5,895	G1 Ganglioglioma (P)
	Op19C1	601×535×826	6,499	0	1,350	1,933	G1 Meningioma (P)
	Op20C1	378×330×826	1,842	3,655	1,513	2,625	G4 Glioblastoma (P)
	Op21C1	452×334×826	3,405	167	793	5,330	Breast Carcinoma (S)
	Op21C2	448×324×826	2,353	31	555	2,137	Breast Carcinoma (S)
	Op21C5	433×340×826	969	0	1,637	1,393	Breast Carcinoma (S)
	Op22C1	597×527×826	2,806	0	1,064	3,677	G3 Anaplastic Oligodendroglioma (P)
	Op22C2	611×527×826	8,174	0	680	0	G3 Anaplastic Oligodendroglioma (P)
	Op22C3	592×471×826	0	96	0	0	G3 Anaplastic Oligodendroglioma (P)
Second Data Campaign	Op34C1	319×356×826	0	0	0	15,609	G3 Anaplastic Astrocytoma (P)
	Op34C2	300×342×826	512	145	0	12,979	G3 Anaplastic Astrocytoma (P)
	Op34C3	290×301×826	0	360	0	10,533	G3 Anaplastic Astrocytoma (P)
	Op35C1	431×503×826	9,025	0	7,287	485	G2 Oligodendroglioma (P) [DL]
	Op35C2	312×535×826	0	1,338	629	1,353	G2 Oligodendroglioma (P)
	Op36C1	412×324×826	11,665	0	4,461	4,621	G4 Glioblastoma (P) [DL]
	Op36C2	432×322×826	2,940	888	2,980	5,853	G4 Glioblastoma (P)
	Op37C1	434×453×826	12,719	0	2,524	11,161	G4 Glioblastoma (P) [DL]
	Op37C2	315×526×826	2,997	0	375	4,166	G4 Glioblastoma (P)
	Op37C3	290×422×826	3,201	407	0	0	G4 Glioblastoma (P)
	Op37C4	280×444×826	0	330	0	0	G4 Glioblastoma (P)
	Op38C1	497×490×826	18,511	2,295	4,229	3,669	G1 Meningioma (P)
	Op39C1	415×446×826	4,003	244	489	9,829	G4 Glioblastoma (P)
	Op39C2	399×439×826	7,705	1,629	822	9,867	G4 Glioblastoma (P)
	Op40C1	303×374×826	2,728	394	1,151	3,492	G1 Meningioma (P)
	Op40C2	294×344×826	817	700	2,130	902	G1 Meningioma (P)
	Op41C1	449×486×826	2,359	69	1,047	5,030	G1 Ganglioglioma (P)
	Op41C2	437×488×826	4,874	158	2,150	4,888	G1 Ganglioglioma (P)
	Op42C1	629×646×826	20,565	0	10,956	8,991	G2 Astrocytoma (glioma) (P)
	Op42C2	623×584×826	19,435	428	5,110	3,983	G2 Astrocytoma (glioma) (P)
	Op42C3	650×582×826	2,385	401	979	716	G2 Astrocytoma (glioma) (P)
	Op43C1	575×543×826	28,285	1,177	4,012	5,995	G4 Glioblastoma (P)
	Op43C2	554×446×826	17,236	475	2,103	1,964	G4 Glioblastoma (P)
	Op43C3	522×533x826	11,798	589	2,727	1,587	G4 Glioblastoma (P)
	Op43C4	538×525×826	14,160	0	4,749	687	G4 Glioblastoma (P)
Third Data Campaign	Op50C1	565×533×826	2,116	1,091	620	5,502	G1 Meningioma (P)
	Op51C1	635×617×826	1,164	0	424	31,247	G4 Glioblastoma (P)
	Op53C1	546×446×826	361	5,549	0	33,606	Breast Carcinoma (S)
	Op54C1	515×504×826	2,697	0	3,506	9,535	G4 Glioblastoma (P)
	Op55C1	397×435×826	3,128	0	901	8,278	G3 Astrocytoma (glioma) (P)
	Op55C2	500×349×826	0	1,046	545	9,415	G3 Astrocytoma (glioma) (P)
	Op56C1	446×598×826	1,346	4,081	2,200	28,370	G2 Astrocytoma (glioma) (P)
	Op56C2	467×566×826	1,326	372	1,116	7,702	G2 Astrocytoma (glioma) (P)
	Op57C1	440×535×826	1,773	771	1,263	23,415	Breast Carcinoma (S)
	Op58C2	721×752×826	6,589	1,629	4,565	43,565	G2 Meningioma (P)
Total number of labelled pixels			**305,449**	**37,944**	**101,305**	**405,401**	

Table 2 Coarse search to optimize the hyperparameters of the processing algorithms to classify the data based on the spectral information using the validation and training set with 1000 pixels per class. *HP: Hyperparameter; I: Initial value; S: Step value; F: Final value; γ: Lambda; C: Cost; T: Number of trees; K: Number of nearest neighbours; L: Size of the output layers; EBEAE: Extended blind end-member and abundance extraction; NEBEAE: Nonlinear extended blind end-member and abundance extraction; RF: Random forest; KNN-E: K-nearest neighbour with Euclidean distance; KNN-C: K-nearest neighbour with cosine distance; SVM-L: Support vector machine with linear kernel; SVM-RBF: Support vector machine with radial basis function kernel; DNN: deep neural network.*

Classifier	HP	# Fold	Coarse Search		Macro F1-Score (%)
			I/S/F	Optimal	
EBEAE	γ	1	0.05/0.05/0.2	0.20	0.443
		2	0.05/0.05/0.2	0.20	0.482
		3	0.05/0.05/0.2	0.05	0.564
		4	0.05/0.05/0.2	0.05	0.501
		5	0.05/0.05/0.2	0.10	0.614
NEBEAE	γ	1	0.05/0.05/0.2	0.20	0.389
		2	0.05/0.05/0.2	0.05	0.501
		3	0.05/0.05/0.2	0.20	0.538
		4	0.05/0.05/0.2	0.20	0.482
		5	0.05/0.05/0.2	0.20	0.466
RF	T	1	1/10/500	50	0.668
		2	1/10/500	50	0.692
		3	1/10/500	50	0.813
		4	1/10/500	50	0.711
		5	1/10/500	50	0.740
KNN-E	N	1	1/10/500	1	0.617
		2	1/10/500	1	0.684
		3	1/10/500	31	0.825
		4	1/10/500	51	0.710
		5	1/10/500	11	0.754
KNN-C	N	1	1/10/500	1	0.630
		2	1/10/500	1	0.688
		3	1/10/500	21	0.828
		4	1/10/500	61	0.714
		5	1/10/500	11	0.756
SVM-L	C	1	$2^{-10}/2^{4}/2^{26}$	2^{14}	0.724
		2	$2^{-10}/2^{4}/2^{26}$	2^{14}	0.730
		3	$2^{-10}/2^{4}/2^{26}$	2^{2}	0.786
		4	$2^{-10}/2^{4}/2^{26}$	2^{14}	0.705
		5	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.726
SVM-RBF	C γ	1	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{0} 2^{0}	0.711
		2	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{4} 2^{-2}	0.787
		3	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{-2} 2^{4}	0.852
		4	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{-4} 2^{2}	0.762
		5	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{2} 2^{-4}	0.784
DNN	L 1 NL 2	1	4/1/10 4/1/10	9 8	0.670
		2	4/1/10 4/1/10	7 8	0.776
		3	4/1/10 4/1/10	8 8	0.830
		4	4/1/10 4/1/10	5 8	0.747
		5	4/1/10 4/1/10	5 4	0.807

Table 3 Coarse search to optimize the hyperparameters of the processing algorithms to classify the data based on the spectral information using the validation and training set with 2000 pixels per class. *HP: Hyperparameter; I: Initial value; S: Step value; F: Final value; HP: Hyperparameter; I: Initial value; S: Step value; F: Final value; γ: Lambda; C: Cost; T: Number of trees; K: Number of nearest neighbours; L: Size of the output layers; EBEAE: Extended blind end-member and abundance extraction; NEBEAE: Nonlinear extended blind end-member and abundance extraction; RF: Random forest; KNN-E: K-nearest neighbour with Euclidean distance; KNN-C: K-nearest neighbour with cosine distance; SVM-L: Support vector machine with linear kernel; SVM-RBF: Support vector machine with radial basis function kernel; DNN: deep neural network.*

Classifier	HP	# Fold	Coarse Search		Macro F1-Score (%)
			I/S/F	Optimal	
EBEAE	γ	1	0.05/0.05/0.2	0.20	0.441
		2	0.05/0.05/0.2	0.20	0.484
		3	0.05/0.05/0.2	0.05	0.565
		4	0.05/0.05/0.2	0.05	0.497
		5	0.05/0.05/0.2	0.05	0.615
NEBEAE	γ	1	0.05/0.05/0.2	0.20	0.389
		2	0.05/0.05/0.2	0.20	0.521
		3	0.05/0.05/0.2	0.20	0.538
		4	0.05/0.05/0.2	0.20	0.479
		5	0.05/0.05/0.2	0.20	0.466
RF	T	1	1/10/500	50	0.680
		2	1/10/500	50	0.693
		3	1/10/500	50	0.811
		4	1/10/500	50	0.713
		5	1/10/500	50	0.742
KNN-E	N	1	1/10/500	11	0.624
		2	1/10/500	1	0.685
		3	1/10/500	51	0.826
		4	1/10/500	101	0.711
		5	1/10/500	11	0.753
KNN-C	N	1	1/10/500	21	0.633
		2	1/10/500	11	0.688
		3	1/10/500	51	0.828
		4	1/10/500	121	0.713
		5	1/10/500	31	0.756
SVM-L	C	1	$2^{-10}/2^{4}/2^{26}$	2^{-6}	0.639
		2	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.715
		3	$2^{-10}/2^{4}/2^{26}$	2^{14}	0.795
		4	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.641
		5	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.730
SVM-RBF	C γ	1	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{2} 2^{-2}	0.937
		2	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{6} 2^{-6}	0.820
		3	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{4} 2^{-2}	0.896
		4	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{0} 2^{0}	0.897
		5	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{0} 2^{0}	0.899
DNN	L 1 L 2	1	4/1/10 4/1/10	5 10	0.707
		2	4/1/10 4/1/10	7 5	0.759
		3	4/1/10 4/1/10	4 9	0.844
		4	4/1/10 4/1/10	9 4	0.779
		5	4/1/10 4/1/10	5 4	0.805

Table 4 Coarse search to optimize the hyperparameters of the processing algorithms to classify the data based on the spectral information using the validation and training set with 4000 pixels per class. *HP: Hyperparameter; I: Initial value; S: Step value; F: Final value; HP: Hyperparameter; I: Initial value; S: Step value; F: Final value; γ: Lambda; C: Cost; T: Number of trees; K: Number of nearest neighbours; L: Size of the output layers; EBEAE: Extended blind end-member and abundance extraction; NEBEAE: Nonlinear extended blind end-member and abundance extraction; RF: Random forest; KNN-E: K-nearest neighbour with Euclidean distance; KNN-C: K-nearest neighbour with cosine distance; SVM-L: Support vector machine with linear kernel; SVM-RBF: Support vector machine with radial basis function kernel; DNN: deep neural network.*

Classifier	HP	# Fold	Coarse Search		Macro F1-Score (%)
			I/S/F	Optimal	
EBEAE	γ	1	0.05/0.05/0.2	0.20	0.440
		2	0.05/0.05/0.2	0.20	0.485
		3	0.05/0.05/0.2	0.05	0.564
		4	0.05/0.05/0.2	0.05	0.486
		5	0.05/0.05/0.2	0.10	0.614
NEBEAE	γ	1	0.05/0.05/0.2	0.20	0.389
		2	0.05/0.05/0.2	0.20	0.523
		3	0.05/0.05/0.2	0.20	0.537
		4	0.05/0.05/0.2	0.20	0.478
		5	0.05/0.05/0.2	0.20	0.465
RF	T	1	1/10/500	50	0.684
		2	1/10/500	50	0.700
		3	1/10/500	50	0.816
		4	1/10/500	50	0.711
		5	1/10/500	50	0.767
KNN-E	N	1	1/10/500	11	0.627
		2	1/10/500	21	0.683
		3	1/10/500	141	0.829
		4	1/10/500	201	0.711
		5	1/10/500	1	0.755
KNN-C	N	1	1/10/500	21	0.635
		2	1/10/500	21	0.689
		3	1/10/500	91	0.832
		4	1/10/500	241	0.711
		5	1/10/500	1	0.769
SVM-L	C	1	$2^{-10}/2^{4}/2^{26}$	2^{-6}	0.648
		2	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.726
		3	$2^{-10}/2^{4}/2^{26}$	2^{22}	0.810
		4	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.639
		5	$2^{-10}/2^{4}/2^{26}$	2^{-2}	0.722
SVM-RBF	C γ	1	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{-2} 2^{0}	0.715
		2	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{2} 2^{-2}	0.795
		3	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{-4} 2^{4}	0.854
		4	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{-6} 2^{4}	0.767
		5	$2^{-10}/2^{2}/2^{10}$ $2^{-10}/2^{2}/2^{10}$	2^{4} 2^{0}	0.792
DNN	L 1 L 2	1	4/1/10 4/1/10	9 9	0.717
		2	4/1/10 4/1/10	10 6	0.775
		3	4/1/10 4/1/10	8 9	0.851
		4	4/1/10 4/1/10	9 8	0.784
		5	4/1/10 4/1/10	9 4	0.807

Table 5 Data partition detail of the five folds. ** indicates patients without tumour samples labelled*.

#Fold	Training Patients (#Total Images)	Validation Patients (#Total Images)	Test Patients (#Total Images)
1	Remaining 23 patients (34)	Op4*, Op8, Op22, Op39, Op41 (10)	Op10*, Op16*, Op21, Op37, Op42, Op54* (17)
2	Remaining 22 patients (37)	Op37, O41, Op42, Op50, Op58 (11)	Op4*, Op12, Op17*, Op22, Op38, Op43, Op55 (13)
3	Remaining 22 patients (41)	Op4*, Op36, Op43, Op53, Op57 (8)	Op5*, Op13*, Op18*, Op34, Op39, Op50, Op56 (12)
4	Remaining 22 patients (45)	Op10, Op18, Op36, Op50, Op54 (7)	Op7*, Op14, Op19*, Op35, Op40, Op51*, Op57 (9)
5	Remaining 22 patients (41)	Op10*, Op35, Op37, Op 40, Op50 (10)	Op8, Op15, Op20, Op36, Op41, Op53, Op58 (10)