

<https://doi.org/10.1038/s43856-024-00562-3>

Towards machine learning-based quantitative hyperspectral image guidance for brain tumor resection

Check for updates

David Black¹, Declan Byrne¹, Anna Walke², Sidong Liu³, Antonio Di Ieva³, Sadahiro Kaneko⁴, Walter Stummer², Tim Salcudean¹ & Eric Suero Molina^{2,3} ✉

Abstract

Background Complete resection of malignant gliomas is hampered by the difficulty in distinguishing tumor cells at the infiltration zone. Fluorescence guidance with 5-ALA assists in reaching this goal. Using hyperspectral imaging, previous work characterized five fluorophores' emission spectra in most human brain tumors.

Methods In this paper, the effectiveness of these five spectra was explored for different tumor and tissue classification tasks in 184 patients (891 hyperspectral measurements) harboring low- ($n = 30$) and high-grade gliomas ($n = 115$), non-glioma primary brain tumors ($n = 19$), radiation necrosis ($n = 2$), miscellaneous ($n = 10$) and metastases ($n = 8$). Four machine-learning models were trained to classify tumor type, grade, glioma margins, and IDH mutation.

Results Using random forests and multilayer perceptrons, the classifiers achieve average test accuracies of 84–87%, 96.1%, 86%, and 91% respectively. All five fluorophore abundances vary between tumor margin types and tumor grades ($p < 0.01$). For tissue type, at least four of the five fluorophore abundances are significantly different ($p < 0.01$) between all classes.

Conclusions These results demonstrate the fluorophores' differing abundances in different tissue classes and the value of the five fluorophores as potential optical biomarkers, opening new opportunities for intraoperative classification systems in fluorescence-guided neurosurgery.

Plain language summary

Complete surgical removal of some primary brain tumors is difficult because it can be hard to distinguish the edge of the tumor. We evaluated whether the edges of tumors and the tumor type and grade can be more accurately determined if the tumor is imaged using many different wavelengths of light. We used measurements taken from the tumors of people undergoing brain tumor surgery and developed machine-learning algorithms that could predict where the edge of the tumor was. The methods could also provide information about the type and grade of the brain tumor. These classifications could potentially be used during operations to remove brain tumors more accurately and thus improve the outcome of surgery for people with brain tumors.

Surgical resection of malignant glioma is complex, and recurrences are the rule rather than the exception, leading to patients' poor prognosis¹. This is partly due to poorly differentiated tumor tissue, especially at infiltrating margins, closely resembling healthy tissue during surgery². To address this problem, 5-aminolevulinic acid (5-ALA)-induced protoporphyrin IX (PpIX) fluorescence guidance has been established as a surgical adjunct in neurosurgery. The complete resection rate for contrast-enhancing tumors increased, in the initial approval study from 2006, from 36% operated under white light only to 65% for resections with fluorescence guidance². Improvements in neuroimaging and different surgical adjuncts, such as intraoperative ultrasound, pre- and intra-operative brain mapping, and

monitoring techniques, amongst others, currently allow for over 95% complete resection rates, whenever feasible³. Consequently, PpIX fluorescence is widely used in the resection of high-grade gliomas. Furthermore, it is the subject of research in low-grade gliomas^{4–7}, meningiomas^{8–11}, and other brain tumors¹², as well as in oral cancer^{13,14}, bladder cancer^{15,16}, and skin cancer¹⁷. Additionally, 5-ALA is used for photodynamic therapy¹⁸, to treat skin and other brain malignancies¹⁹.

Fluorescence is achieved by administering 5-ALA orally prior to surgery¹, which is metabolized to PpIX, a precursor in heme biosynthesis. The mechanisms leading to the selective accumulation of PpIX in glioma cells²⁰ are not entirely understood^{11,21}. Several explanations have been

¹Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada. ²Department of Neurosurgery, University Hospital Münster, Münster, Germany. ³Computational NeuroSurgery (CNS) Lab, Macquarie Medical School, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, NSW, Australia. ⁴Department of Neurosurgery, Hokkaido Medical Center, National Hospital Organization, Sapporo, Japan.

✉ e-mail: eric.suero@ukmuenster.de

proposed, including disruption of the blood-brain barrier, as commonly observed in high-grade gliomas (HGG), which is otherwise non-permeable to 5-ALA^{1,7,22}. Reduced activity of ferrochelatase, which would otherwise metabolize PpIX^{15,21,23}, and changes in the tumor microenvironment affecting 5-ALA uptake and PpIX efflux^{24–26} may play a part as well. However, PpIX is present in increased concentrations 7–8 hours after 5-ALA administration in HGG²⁷. PpIX fluoresces after illumination with intense 405 nm (blue) light. Absorption of such a photon lifts the molecule to an excited state, decreasing slightly through vibrational relaxation before returning to its ground state and emitting a second photon²⁸. Due to the decrease in energy before the second transition, the emitted photon has a longer wavelength of around 634 nm (red). This energy change called the Stokes shift, allows the red fluorescing areas to be differentiated from the otherwise blue-reflecting tissue²⁹.

With optical highpass filters such as the BLUE400 and BLUE400 AR systems (Carl Zeiss Meditec AG, Oberkochen, Germany) that block the relatively intense reflected blue light but transmit the red fluorescence, surgeons can differentiate 5-ALA positive tumor from surrounding tissue^{30,31}. This has led to widespread adoption. However, the proven increased resection rates are only for malignant glioma. Unfortunately, many lower-grade gliomas and even some with high-grade regions exhibit low PpIX accumulation and do not visibly fluoresce. Furthermore, autofluorescence makes it impossible to discern the diminished PpIX fluorescence even when imaging sensitivity is increased because it shares the same spectral range as the PpIX emission^{32–35}.

Hence, quantitative spectroscopic systems have been developed that measure the emission spectra and separate PpIX from autofluorescence through a priori knowledge of the fluorophores present and their individual emission spectra^{35–37}. These devices have been used primarily in research and consist of either point probes^{5,8,37–39} or wide-field hyperspectral^{12,40,41} devices. In addition to potential future intraoperative use to help distinguish tumors, spectroscopy can answer many questions about different diseases^{23,36}, how best to treat them^{27,40,41}, and how to improve imaging systems for future intra-operative integration^{12,42}.

One potential use of such a system is to distinguish between different tissue types. Several classifications are commonly applied to brain tumors, including the tumor type (e.g., glioma, metastasis, medulloblastoma, and others⁴³). Additionally, the differentiation of tumor from non-tumor or necrotic tissue (e.g., radiation necrosis, which can radiologically mimic tumor progression) is relevant in neuro-oncology and neurosurgery. Different tissue types differ substantially in behavior and prognosis, so their correct identification is paramount. Several studies have classified brain tumor type in MR images^{44,45}, but the same has not been attempted with fluorescence.

Furthermore, gliomas are classified according to their histological, genetic features, and biological behavior⁴⁶. They are categorized as grade I-IV by the World Health Organization (WHO) classification system, where traditionally, I and II are considered low grade, and III and IV are high grade⁴⁷, indicating a bad prognosis⁴⁸. Molecular parameters, i.e., isocitrate dehydrogenase (IDH)-mutation, O6-methylguanine-DNA-methyltransferase (MGMT), among others, assist in further subclassification of these tumors⁴⁹. We aim to predict tumor molecular characteristics, such as IDH mutations or genetic aberrations related to malignancy, by correlation with the spectral signature of tumors. Such knowledge would directly result in changes to surgical strategy. IDH mutations cause a shift in enzymatic activity, converting α -ketoglutarate to 2-hydroxyglutarate and inhibiting α -KG-dependent enzymes. This leads to metabolic reprogramming, hinders cell differentiation, and initiates tumorigenesis⁵⁰. Small molecule inhibitors can reverse this process, making knowledge of IDH mutation status crucial for guiding their targeted use in treatment. Knowing a tumor's grade or molecular characteristics could also help make decisions during and after surgery. The tumor types are shown in Supplementary Fig. 2. Some studies have performed automatic classification of tumor grade using convolutional neural networks (CNN) on magnetic resonance images⁵¹ and digitized histopathology slides⁵², but not through fluorescence.

Additionally, one of the primary difficulties of resection of grade II, III, and IV gliomas is the infiltrative nature of the tumors. Around the solid tumor portion, there is a region of infiltrative margin characterized by decreased tumor cellularity transitioning to healthy tissue⁵³. To delineate the margins, guidance from pre-operative MRI images is commonly used in intraoperative neuronavigation systems⁵⁴, but its accuracy suffers from factors such as brain shift⁵⁵. Differentiating more accurately between margin regions would significantly enhance the effectiveness of surgical resection of glioma and patients' outcomes^{56,57}. Leclerc et al. presented the first work applying machine learning (ML) to neurosurgical fluorescence spectroscopy, classifying different tissues based on principal component analysis (PCA) of the fluorescence spectra⁵⁸, achieving 77% accuracy on 50 samples.

Although previous work has correlated PpIX fluorescence with WHO grade and tumor margins⁵⁹, it is not well understood how different tumor types affect the abundances of the different fluorophores⁶⁰. Through recent characterization of their basis spectra³⁵, we can now precisely analyze the abundances of the major fluorophores in hyperspectral fluorescence images. Hence, it is possible to study how the two photo-states of PpIX and the autofluorescence from flavins (i.e., flavin adenine dinucleotide), NADH, and lipofuscin are affected by the tissue type. This paper performs four classification tasks using fluorescence spectroscopy, aiming to use the abundances of the five fluorophores to differentiate between (1) tumor and tissue types (e.g., glioblastoma, meningioma, etc.), (2) WHO grades, (3) between solid tumor, infiltrating zone, and reactively altered non-tumor brain tissue, and (4) between IDH-mutant and IDH-wildtype glioma. Using random forests and multilayer perceptron, the classifiers achieve average test accuracies of 84–87%, 96.1%, 86%, and 91%, respectively, thus demonstrating the fluorophores' differing abundances in different tissue classes and the value of the five fluorophores as potential optical biomarkers. This invites further research into intraoperative classification systems in fluorescence-guided neurosurgery.

Methods

Device and Dataset

A hyperspectral imaging device was used, as previously described^{27,41,59}, and outlined below. Patients received 5-ALA orally at a dose of 20 mg/kg b.w., four hours prior to anesthesia induction for tumor resection surgery, as per routine practice in cases of suspected malignant gliomas. All Patients consented to the use of 5-ALA for fluorescence-guided resection. The tumor entity is not known prior to surgery. Thus, occasionally, patients harboring distinct tumor types also received 5-ALA. All ex vivo data collection was carried out with informed consent, complied with institutional guidelines, and was approved by the ethical committee of the University of Münster.

Tissue samples were resected during surgery and measured directly ex vivo on a petri dish. The sample was first illuminated with blue light (405 nm LED), then with no light for recording background noise, and finally with broadband white light. During each illumination phase, the emitted light from fluorescence and the reflected light were gathered in the objective lens (OPMI Pico - Carl Zeiss AG, Oberkochen, Germany) and routed to a scientific metal oxide semiconductor (sCMOS - PCOEdge, Excellitas Technologies GmbH, Wiesbaden, Germany) camera through a series of optical filters. The first filters removed the intense blue reflected light. Next, a liquid crystal tunable bandpass filter (LCTF - Meadowlark Optics, Frederick, CO) transmitted a narrow spectral band of the emitted light to the sCMOS. The LCTF was swept across the visible range (420–730 nm) in 3–5 nm increments, during each of which a grayscale image was captured by the sCMOS. In this way, a hyperspectral data cube containing all the spectral and spatial information was generated. Any pixel can be selected, and an emission spectrum extracted for that pixel from the cube. Figure 1 shows 1000 examples of typical measured spectra. During blue-light illumination, the fluorescence spectra were captured. The white-light illuminated data cube was used for dual-band normalization⁴², and the no-light spectra were used to remove the dark noise of the camera sensor from the images. Total acquisition time with 500 ms exposure time for each image is 3–4 minutes. This ensures a good signal-to-noise ratio, even with

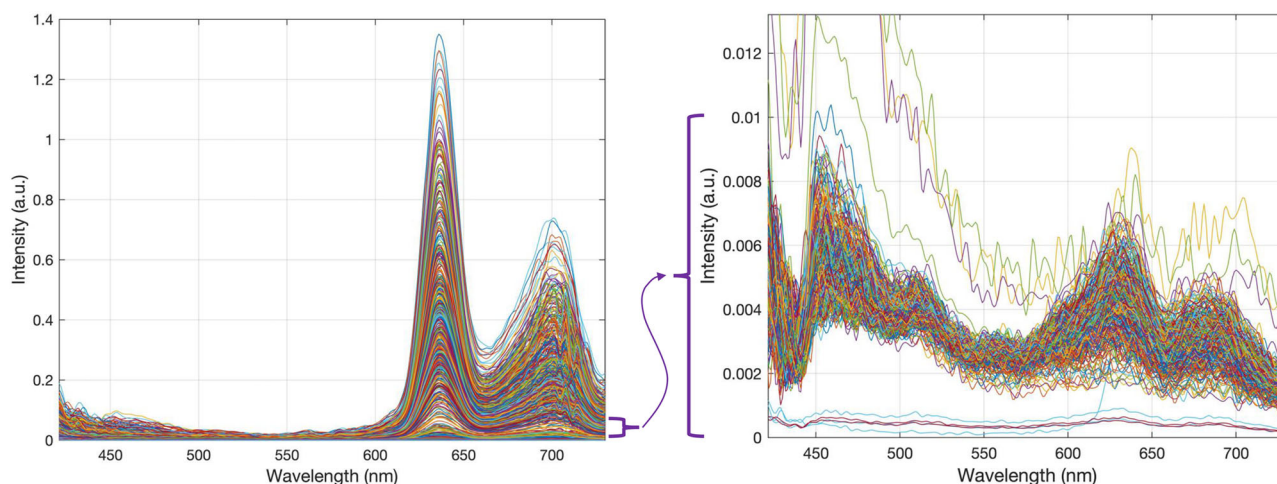


Fig. 1 | Typical Example Measured Spectra. The plot of representative 1000 spectra sampled randomly from the complete dataset (left). Furthermore, we added a zoomed-in view of some typical weak spectra, where the different autofluorescence contributions can be visualized, not just PpIX.

very little fluorescence. Each image was 2048×2048 pixels with 47.62 pixels per mm.

This device was used for ex vivo analysis of tumor biopsies at the University Hospital Münster, Germany. Typical samples had an average diameter of 288 pixels, or 6 mm, and ranged from approximately 4 to 10 mm. The resulting data set was used in this study to explore the effect of different classifications of tumors on the presence of the five main fluorophores. In particular, we analyzed tumor type, WHO grade, and whether the sample is from solid tumor, infiltrating zone, or non-tumor. The classes and number of samples are shown below and were chosen based on the availability of patient data. Many samples were measured before the updated WHO classification (CNS5, 2021)⁴⁷ and thus followed the previous 2016 WHO classification system⁴⁹.

Tissue Type Classification: ($n = 632$ biopsies)

Pilocytic astrocytoma (PA; $n = 5$), diffuse astrocytoma (DA; $n = 57$), anaplastic astrocytoma (AA; $n = 51$), glioblastoma (GB; $n = 410$), grade II oligodendroglioma (OD; $n = 24$), ganglioglioma (GG; $n = 4$), medulloblastoma (MB; $n = 6$), anaplastic ependymoma (AE; $n = 8$), anaplastic oligodendroglioma (AO; $n = 4$), meningioma (MN; $n = 37$), metastasis (MT; $n = 6$), radiation necrosis (RN; $n = 20$).

Margin Classification: ($n = 288$ biopsies)

Solid tumor (ST; $n = 131$), infiltrating zone (IZ; $n = 57$), reactively altered brain tissue (RABT; $n = 100$).

WHO Grade Classification: ($n = 571$ biopsies)

Grades I ($n = 9$), II ($n = 84$), III ($n = 57$), IV ($n = 421$).

IDH Classification: ($n = 411$ biopsies)

IDH-mutant ($n = 126$), IDH-wildtype ($n = 285$)

Note that the RABT class is the tissue outside the infiltration zone that cannot be classified as a solid tumor or infiltration zone. Intraoperatively, it is imperative to differentiate this from tumor tissue, which can be viewed as healthy tissue. It is otherwise unethical to biopsy healthy brain tissue.

The biopsies varied in size and shape, but approximately 100–1000 spectra were extracted from each. Regions of 10×10 pixels were averaged to produce one spectrum for noise reduction, and the regions were non-overlapping to ensure each data point was independent. For a given classification task, spectra were sampled randomly from the tumor portions of the biopsies. To avoid inadvertently sampling from the background glass slide, the tumor was first segmented automatically in MATLAB using the 634 nm fluorescence image. The result is shown in step 2 of Fig. 2; the resulting masks were also checked manually.

The raw fluorescence spectra were corrected, normalized, and unmixed as described previously^{12,35}. Each fluorescence spectrum is assumed to consist of a linear combination of five fluorophores whose spectra are known a priori. These basis spectra are the two photo-states of

PpIX (denoted PpIX634 and PpIX620)^{36,37}, NADH, lipofuscin, and flavins³⁵. The example measured and the basis spectra are described and shown in Fig. 2 and in previous work³⁵. The unmixing calculates the abundances of the five spectra by minimizing the squared error between the measured spectrum and a linear combination of the five spectra. If the basis spectra are combined as columns in a 310×5 matrix B , the measured spectrum is the 310×1 vector y , and the 5×1 vector of abundances is c , then we find c as

$$c = \underset{c}{\operatorname{argmin}} \left\{ \frac{1}{2} \|y - Bc\|^2 \right\}$$

using non-negative least squares. This is illustrated in steps 3–5 of Fig. 2.

Machine learning approach

With the processed data, we tried several ML classification models using Python and experimented with different hyperparameters. The various types are listed in Table 1.

For a given biopsy, instead of using the abundances inferred from only one pixel, it could be more informative to use those of two or more nearby pixels. This is attempted in the *pixels per sample* hyperparameter. Further, to avoid bias, the same number of samples of each class were used to train the classifiers. Since some classes had fewer samples, we varied the number of samples per class. A class with fewer samples than the samples per class hyperparameter was excluded from the classification. Lower numbers of samples per class allowed more classes to be included in the classification. However, more significant numbers of samples per class allowed for more effective training of the classifiers, though on fewer classes. Classes with few biopsies, including anaplastic oligodendroglioma and ganglioglioma, had to be excluded from some tests.

The dataset was split 80/20 into a training and testing set. Hyperparameter and model tuning were carried out using 5-fold cross-validation on the training set to mitigate optimization bias. In total, 195 models were trained and compared for each classification task. For each of the four classification tasks, nine datasets were used with varying pixels per sample and samples per class. Finally, the best model was selected for each category and evaluated with the test set. The receiver operating characteristic curves (ROC) were determined to evaluate the multiclass classifiers, and the area under the curve (AUC) was computed. Accuracy was calculated as the number of correctly classified samples divided by the total number of samples. Confusion matrices were used to determine which classes performed better than others.

We also explored whether the calculated abundances were the most informative space to characterize the biopsies. Instead of relying on the five

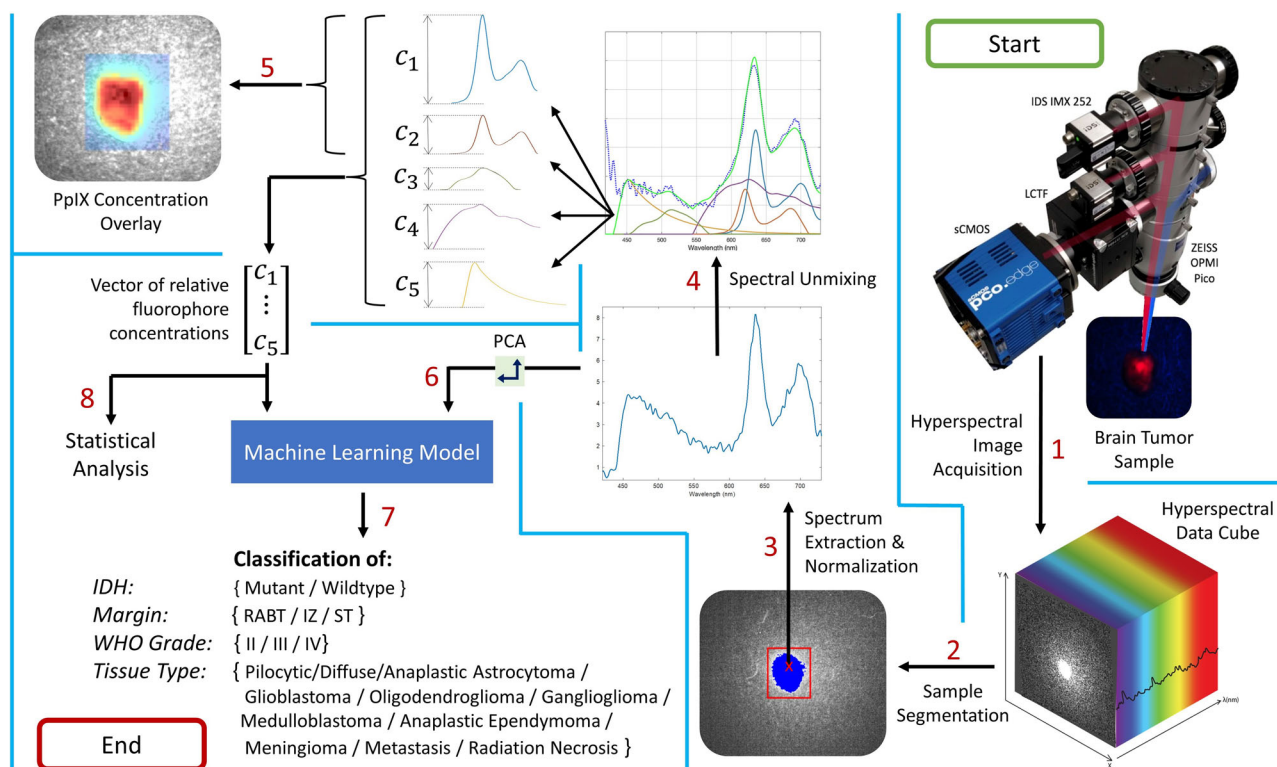


Fig. 2 | Overview of the device and method. This shows the steps of the data processing, from top right to bottom left. The imaging device first captures a hyperspectral data cube from which the ex vivo tumor tissue is segmented and the corresponding spectra are extracted, processed, and unmixed into abundance

vectors or dimensionally reduced using PCA. The abundances can be used, for example, to create a PpIX overlay plot or perform statistical analyses. In this study, they are used as inputs to machine learning models to classify tumor type, margins, WHO grades, and IDH mutation.

basis spectra, we repeated the above experiments with the mathematically optimal projection of the spectra onto a 5-dimensional hyperplane using principal component analysis (PCA). We used the same number of pixels per sample, samples per class, and all other hyperparameters as above. We also visualized the different classes using only the first two or three principal components in a scatter plot. PCA finds a set of orthogonal axes, which maximizes the variance. Depending on assumptions about the signal and noise distributions, this maximizes the mutual information between the real signal and the dimensionally-reduced output⁶¹. Nonetheless, some information was lost. To quantify how well PCA represents the original data, we utilized the variance explained parameter. Given a data matrix of fluorescence spectra, a covariance matrix whose diagonal elements (the variances) sum to the overall variability can be computed. When performing n-component PCA, only the n principal components corresponding with the n largest eigenvalues were kept. All the eigenvalues sum to the overall variability, so by choosing only the n-largest, we lost some of the information contained in the data. In particular, the ratio of the sum of the n selected eigenvalues over the total variability was called the variance explained. It effectively yielded the percent of total information the chosen principal components represented.

Statistics and reproducibility. Statistical analyses were performed to determine correlations and statistical significance. This analysis was performed in MATLAB using the two-sample Kolmogorov-Smirnov test. Python and SciKit-Learn were used for the machine learning algorithms. The entire process, from imaging to data processing and extraction to analysis, is shown in Fig. 2, and the raw results and all p values are found in Supplementary Data 1 and 2, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

As explained in detail in the Methods section, four classification tasks were explored. Tumor types belonged to the following classes: Pilocytic astrocytoma (PA), diffuse astrocytoma (DA), anaplastic astrocytoma (AA), glioblastoma (GB), grade II oligodendroglioma (OD), ganglioglioma (GG), medulloblastoma (MB), anaplastic ependymoma (AE), anaplastic oligodendroglioma (AO), meningioma (MN), metastasis (MT), and radiation necrosis (RN). Tumor margin classification was performed on three classes: solid tumor (ST), infiltrating zone (IZ), and reactively altered brain tissue (RABT). Reactively altered brain tissue can be considered as healthy tissue. IDH status (mutated and wildtype) was further assessed. Finally, WHO grades II, III, and IV were considered. After hyper-parameter tuning using cross-validation on the training dataset, the following results and metrics were calculated using the test set.

Visualization

Since the five known fluorophore abundances create a 5-dimensional space, it is impossible to visualize the data and distinguish patterns manually. Studying scatter plots of two fluorophores at a time would be feasible, but this would be a reductionistic approach drawing a very limited picture. Instead, we used PCA for dimensionality reduction, which results in some information loss. The exact amount of information lost for each classification is shown in Table 2. However, PCA fails if the data forms a non-linear manifold. One alternative that captures this more effectively is t-distributed stochastic neighbor embedding (t-SNE). This was also used to better visualize the data. Some of the more informative plots are shown in Fig. 3.

Though drawing conclusions from the visualizations in Fig. 3 is difficult, it shows that some classes can be visually distinguished from others with reasonable confidence, even in dimensionally reduced form. Furthermore, Fig. 3a and b show that the fluorophores and PCA give similar information, suggesting the fluorophore abundances’ utility as informative

Table 1 | Explored Classifier Algorithms and Hyperparameters

Pixels per Sample	[1, 2, 3]		
Samples per Class	Tumour Type	Margin	WHO Grade
	[300, 500, 800]	[1000, 3000, 5000]	[500, 1500, 3000]
Random Forests ⁶⁹	Number of trees: [50, 75, 100, 125, 150] Splitting Criteria: [Gini, Entropy, Log loss] Minimum samples to split: [2, 3, 4] Maximum features per tree: [Square root (sqrt), log2, None]		
KNNs ⁷⁰	Number of neighbors: [3, 5, 7, 9] Weights: [Uniform, Distance] p: [1, 2]		
SVM ⁷¹	Kernel: [Radial Basis Function (RBF), Linear, Polynomial, Sigmoid]		
MLP ⁷²	Number of hidden layers: [1,2,3] Number of neurons: [25, 50, 100, 150] Activation: ReLU (Rectified linear unit) Solver: [Adam, Limited-memory BFGS (LBFGS)] Nesterov Momentum: 0.9		
AdaBoost ⁷³	Number of estimators: 50 Learning rate: 1.0 Algorithm: SAMME.R		

Multilayer perceptron (MLP), support vector machine (SVM), and k-Nearest Neighbor (KNN) classifiers are described in the papers cited in the table.

Table 2 | Variance Explained from 2 to 5-Dimensional PCA

PCA Dimension	Tumor Type	Margin	WHO Grade
2	79%	79%	85%
3	87%	89%	93%
4	94%	95%	95%
5	97%	99%	96%

A value of 100% means that all variance is accounted for. With five values, almost no information is lost, indicating that the 5 primary fluorophores most likely cause the measured emissions.

input features to ML classifiers. Table 2 additionally demonstrates that almost all the spectral information is stored in five dimensions, which lends further credence to the five a priori basis spectra³⁵.

Tissue type classification

Since the tissue type data is categorical, it is not possible to determine linear correlations. However, pairwise Kolmogorov-Smirnov tests were used to test the null hypothesis for each fluorophore: abundances from each pair of classes come from the same distribution. All available spectra were used in the significance tests for each class, which ranged from 200 to over 5000, depending upon the class. The chosen test (kstest2.m in MATLAB) does not require equal sample sizes. All class pairs were significantly different ($p < 0.01$) across all the fluorophores except those listed in Supplementary Table 1. Thus, even if a pair of classes did not differ significantly in one fluorophore, it did in the remaining four. All p values are found in Supplementary Data 2.

Both multilayer perceptrons (MLPs) and random forests performed well for classifying the tumor type. None of the other classifiers were effective. Performance was better with more pixels per class, partly because the models were trained on more data and partly because they had to distinguish between fewer classes. Therefore, we present the best results overall and the best results from training with all the classes.

The best-performing classifier, including most classes, was a random forest model (150 trees, 500 samples per class, 3 pixels per sample, square root of the number of features $\sqrt{n_{\text{features}}}$ features per tree, cross-entropy splitting criterion). The average accuracy was 83.6%, with an AUC of 0.98 in classifying between all classes. Figure 4 shows this classifier's confusion matrix and multiclass ROC.

This shows outstanding performance for some classes and relatively poor for others. In particular, GB, DA, and AA have low classification

accuracy, and GB is frequently mistaken for AE. However, AE is rare and generally easy to distinguish visually from healthy tissue, so fluorescence is not commonly used. Hence, despite its good classification accuracy, AE was removed. Additionally, this data used the 2016 WHO classification because it was partly collected before 2021. After the 2021 WHO classification, however, IDH-wildtype anaplastic astrocytomas are considered GB. Thus, these samples ($n = 33$) were re-labeled as GB, and the models were re-trained.

With these changes, the best-performing model used the same algorithm and hyperparameters as above but gave 87.3% accuracy and an average AUC of 0.98, as shown in Fig. 5.

We also considered classifying between higher-level groupings of tumor types (i.e., higher in the hierarchy of Supplementary Fig. 2). The classes were glioma, meningioma, medulloblastoma, and radiation necrosis; ependymoma and metastasis were excluded for small sample size. The best-performing model for this task was an MLP (150 neurons per layer, 3 hidden layers, 500 samples per class, 2 pixels per sample, Adam solver). The model achieved 89.8% accuracy and an AUC of 0.97, as shown in Supplementary Fig. 3. The glioma class is very broad, likely leading to slightly lower accuracy. Generally, however, the performance is good. The accuracy is boosted to 90.67% by using 800 samples per class, which necessitates leaving out radiation necrosis.

Therefore, tumor type does indeed have a strong effect on the five fluorophores: strong enough that it is possible to distinguish between eleven types of tumors and other tissue types with a relatively high degree of accuracy. Some gliomas, specifically GB, DA, and AA, were difficult to discern. Meningioma was also sometimes misclassified as medulloblastoma. On the other hand, many tissue types, including ganglioglioma, radiation necrosis, and, to a slightly lesser degree, medulloblastoma, could be classified with almost perfect accuracy.

Using PCA instead of fluorophore abundances to perform this analysis did not improve results. This is likely due to the fact that PCA does not seem to provide more information than the fluorophores.

Margin classification in gliomas

The margin classification would be practically useful in an intraoperative setting. It could delineate solid tumor from infiltrating cells and non-tumor tissue. However, the distinction is relatively subjective (see Discussion section), and as shown in Fig. 3e, there may be substantial overlap between classes. Again, however, all fluorophore abundances significantly differ between every category ($p < 0.01$).

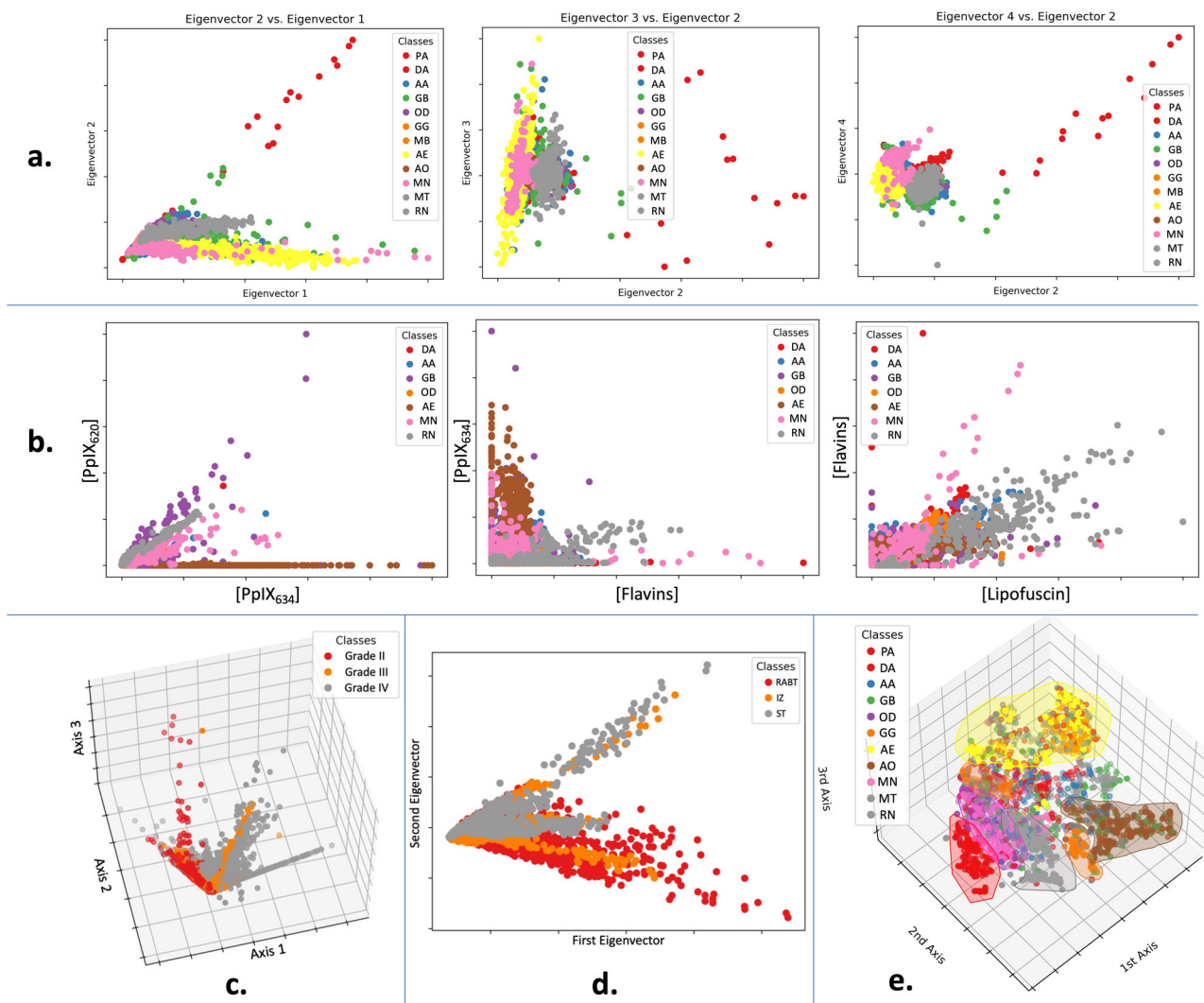


Fig. 3 | Visualizations of tumor data for different classifications. Several classes are visually distinguishable or nearly so despite being very low-dimensional representations, which is promising for ML classification. Panels **a** and **b** compare the effectiveness of PCA versus fluorophore abundances for visualizing the classes, with similar results. This suggests that the fluorophore abundances are already close to the most informative features for representing the data. Subplots **c** and **d** show PCA on WHO grade and tumor margin respectively, while the others show tumor type. Finally, panel **e** shows a 3-component T-SNE on tumor type. Here, we see distinctions between some groups, though others, including glioblastoma and astrocytoma, remain very mixed. All axes are in arbitrary units. WHO grade I and some tumor

types are excluded due to relatively small sample size and unbalanced datasets. The classification abbreviations are as follows: Pilocytic astrocytoma (PA), diffuse astrocytoma (DA), anaplastic astrocytoma (AA), glioblastoma (GB), grade II oligodendroglioma (OD), ganglioglioma (GG), medulloblastoma (MB), anaplastic ependymoma (AE), anaplastic oligodendroglioma (AO), meningioma (MN), metastasis (MT), radiation necrosis (RN), solid tumor (ST), infiltrating zone (IZ), reactively altered brain tissue (RABT). Several of these tumors were operated on before 2021 and are therefore classified according to the older WHO classification system, which is no longer in use.

Consequently, a multilayer perceptron (5000 samples per class, 3 pixels per sample, 100 neurons per hidden layer, 5 hidden layers, Adam solver) distinguishes with 85.7% accuracy and a mean AUC of 0.95, as shown in Fig. 6.

It is reasonable that the infiltrating zone is the most commonly mistaken for solid tumor and reactively altered brain tumor, whereas the latter two are infrequently confused.

WHO grade classification

Furthermore, knowing what WHO-grade certain tumors or tissue regions have during an operation would be valuable and further reinforce the malignant character of the tissue. In practice, many low-grade tumors have an anaplastic focus, which fluoresces, while the surrounding tumor does not. All differences in fluorophore abundances between WHO grades were found to be statistically significant ($p < 0.01$). For this task, a multilayer perceptron classifier performed best (3000 samples per class, 3 pixels per

sample, 150 neurons per layer, 5 hidden layers, Adam solver), with 96.1% accuracy and a testing AUC of 0.99, as shown in Fig. 7. While WHO grades II and III were classified with good accuracy, grade III tumors were more frequently confused for grade IV.

IDH mutation

Finally, another relevant classification task is the IDH mutation. The best-performing classifier for IDH mutation was a random forest model (150 trees, all available samples per class, 2 pixels per sample, $\sqrt{n_features}$ features per tree, Shannon entropy splitting criterion). The average accuracy was 93% in classifying between IDH-mutant and IDH-wildtype tumors. Figure 8 shows this classifier’s confusion matrix and multiclass ROC.

Discussion

In this paper, four ML-based classifiers were developed for tissue types, IDH mutation, tumor margins, and glioma WHO grades. Each classifier achieved

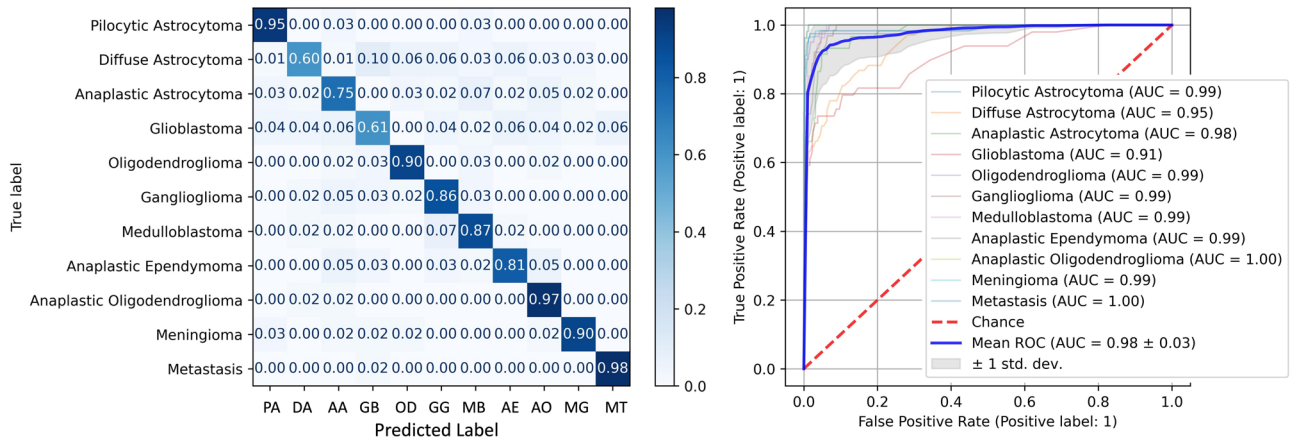


Fig. 4 | Tumor type classifier performance. Confusion matrix and ROC for the best-performing random forest classifier.

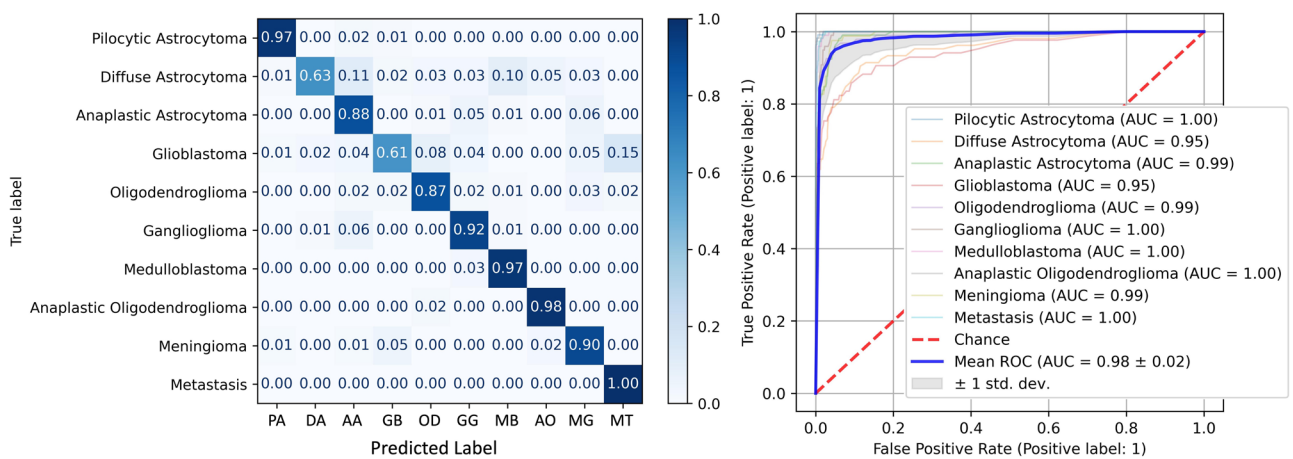


Fig. 5 | Tumor type classifier performance (No AE). Confusion matrix and ROC are used for the best-performing random forest classifier of tumor types with AE excluded, and IDH-wildtype AA relabeled as GB.

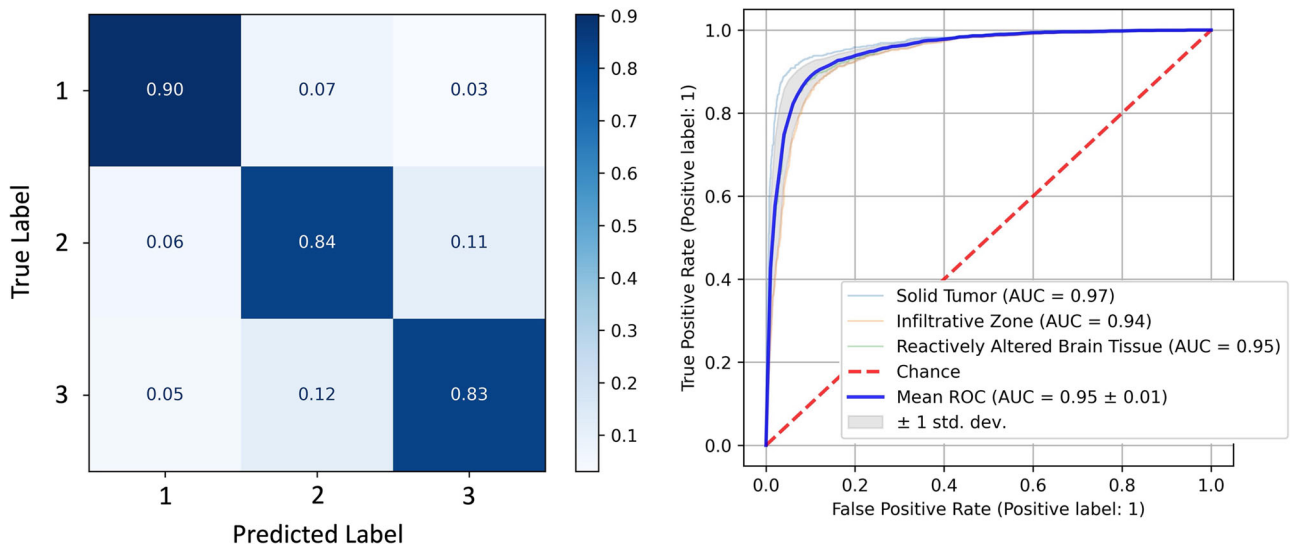


Fig. 6 | Tumor Margin Classifier Performance. Confusion matrix and ROC for the best performing MLP classifier of tumor margins. The classes are [1: Solid tumor, 2: Infiltrating zone, 3: Reactively altered brain tissue].

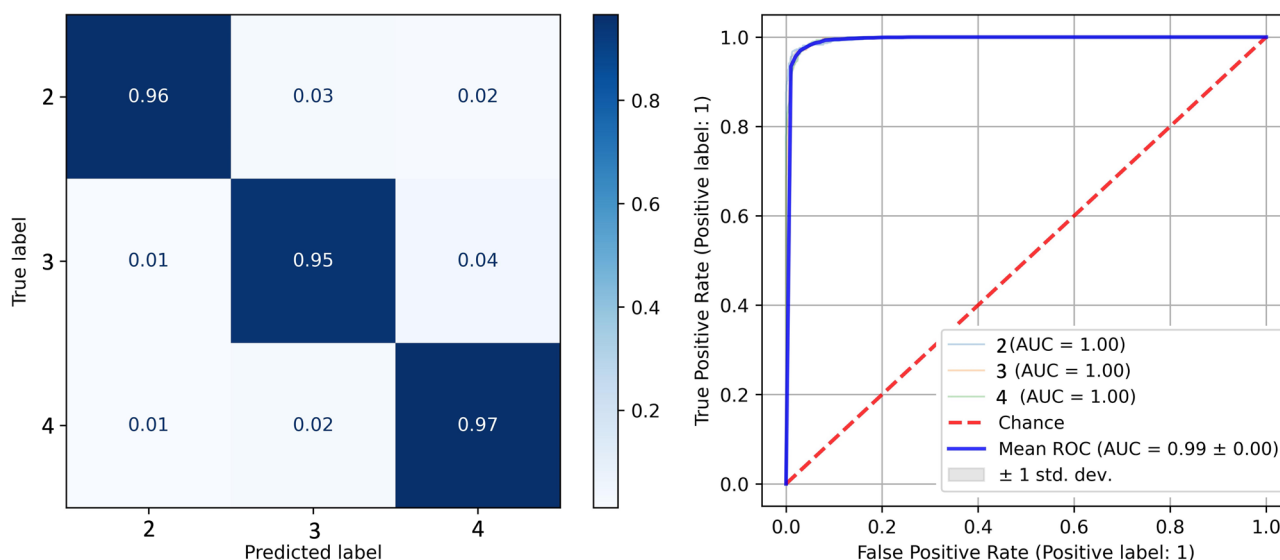


Fig. 7 | WHO Grade Classifier Performance. Confusion matrix and ROC for the best performing random forest classifier of WHO grades. The classes are [2: Grade II, 3: Grade III, 4: Grade IV].

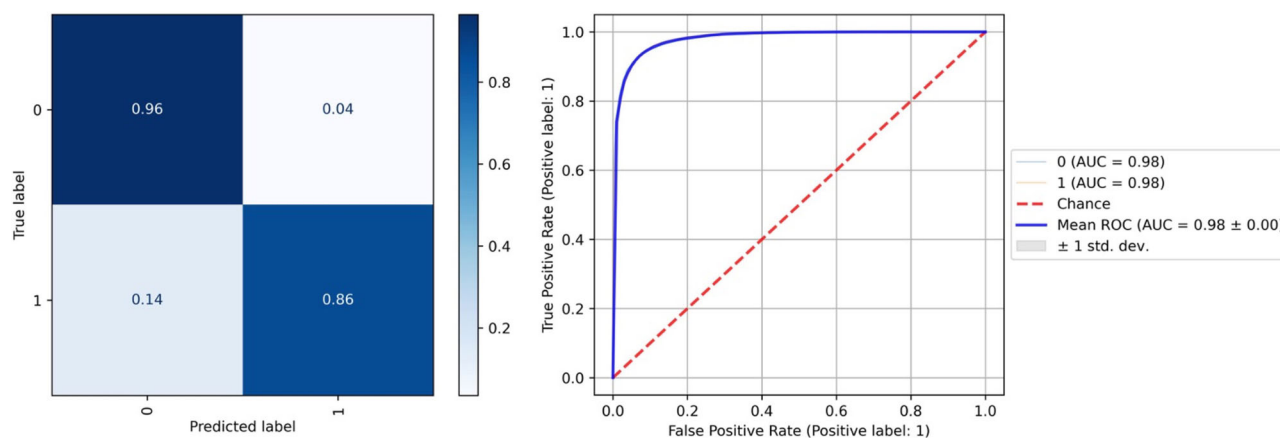


Fig. 8 | IDH Mutation Classifier Performance. Confusion matrix and ROC for the best performing random forest classifier of IDH mutation (1 = mutant, 0 = wildtype). The classifier could predict IDH mutation with an accuracy of 93%.

high accuracy. The margin classification accuracy of 85.7% outperforms that of Leclerc et al.⁵⁸ (77% on 50 samples), likely due to the use of much more data. This represents the accuracy of the model of differentiating tumor (solid tumor and infiltration zone) from healthy non-tumorous tissue (reactive brain altered tissue). Our classifier also used fluorophore abundances instead of PCA. Data from 891 hyperspectral measurements of 184 patients was utilized, corresponding to up to 15000 spectra for a given test. Furthermore, the tissue type and WHO grade classifications performed well. The classifiers achieved an average test accuracy of 87.3% when classifying between tissue types and 89.5% when classifying between higher-level groupings of tissue type.

A possible limitation of the accuracies presented in this study is that some of the training data might be noisy if the automatic tumor segmentation did not work well. This is especially relevant for very small, oddly shaped biopsies or ones with a large splash of fluid next to the tumor, which tended to show up very brightly in the fluorescence images. The masks were checked manually, but some background inclusion was inevitable. Current work explores using convolutional neural networks (CNNs) to improve the labeling process. Another source of noise in the labels is described in the Methods and Results sections, as inaccurate or uncertain labeling could decrease training effectiveness. For instance, the delineation between the margin classes is not entirely straightforward, and some tumor types are no

longer distinguished since the 2021 WHO classification. This was seen between Figs. 5 and 6. Future work should be carried out to replicate the results in different classes, as defined in the 2021 WHO classification system.

The results could likely be improved with more data on the less frequent tumor types (e.g., APXA, ganglioglioma). For the MLP and random forest models, there was a strong trend of increasing accuracy as we increased the amount of training data. This could be due to the reduced number of labels to classify, resulting in a lower probability of error, or because the training process improved with an increased amount of data. It is likely a combination of both, but it is promising that with continued data collection, we will likely be able to improve classification accuracy.

Furthermore, looking at individual spectra or a few randomly selected ones per biopsy loses spatial information, which might be valuable for the presented classification tasks and the initial calculation of fluorophore abundances. Hence, these tasks could benefit from a convolutional approach that considers spatial information and translational invariance. Future work will explore inputting the hyperspectral data cubes directly into a CNN⁶². In general, we have only considered relatively simple, off-the-shelf models, so accuracy may be improved by considering other learning algorithms. In our tests, random forest models outperformed MLP models for every task. However, with more data or more careful design of the neural network architecture and hyperparameters, a neural network would be

expected to outperform the random forests. This is left for future work as this paper focuses on establishing the feasibility of a fluorescence-based approach to tumor classification.

IDH mutations are oncogenic drivers in gliomas and act as a relevant prognostic marker for these tumors⁶³. Thus, having this information during surgery could be valuable. In the context of eloquence, a surgeon may adopt a less aggressive approach in younger patients upon identifying an IDH-mutant oligodendroglioma. In the case of an IDH-wildtype tumor, the survival chances would increase relevantly by maximizing resection, and thus an aggressive surgery would be warranted. Our model can differentiate mutations in IDH with an accuracy of 93%. This is promising for future intraoperative applications.

The labeling of biopsies assumes that the entire biopsy belongs to a particular class. This is a reasonable assumption as the biopsies are small (typically 4–10 mm diameter), carefully removed, and thus relatively homogeneous. In general, however, the tissue is very heterogeneous, and the distinction between classes is blurred. This could introduce noise into the margin classification training data. For example, a pixel from a biopsy labeled solid tumor might be in a small region of infiltrating tumor and thus be incorrectly labeled. This could partially explain why the classification accuracy is only 85.7%. The only way to improve this would be to register the fluorescence image spatially with the histopathology to obtain a pixel-for-pixel map of class labels. However, the histopathological assessment is made from thin biopsy slices in unknown orientations and planes, so spatial registration is infeasible with the current setup and constitutes future work.

In addition, heterogeneity is an issue not only for an ML classifier but also for pathologists. The categorization of tumor margins can be subjective, and there is an intra-observer variability among different pathologists and centers, which affects the testing dataset^{64,65}. Thus, it is likely that the classification model presented in this paper would perform better if pathologists applied a standardized, quantitative measure to distinguish between different tissue regions. With this in mind, the classification accuracy achieved at 85.7% is very good. More rigorous labeling of margins could be performed in the future by having several pathologists from different centers label the data.

The performance of the classifiers may also be affected by the accuracy of the spectral unmixing, particularly the method and the basis spectra used. This paper used the spectra described previously with non-negative least squares unmixing³⁵. However, other spectra have also been proposed for brain tumor fluorescence⁶⁶. The effectiveness of both sets of spectra is evaluated in a recent paper⁶⁷, which also explores the unmixing algorithm, including considerations of sparsity and the underlying probabilistic models. Therefore, future work should extend the analysis presented in this paper to see if new unmixing approaches can improve the accuracy of the classifiers.

Finally, though PCA was used in this paper primarily for visualization, analysis of information loss, and comparison to previous work⁵⁸, it should not be used as an alternative for spectral unmixing. PCA is non-unique, so the factors can be rotated to produce a set of axes that is equivalently optimal from the PCA perspective but which provides an entirely different spectral unmixing result. Instead, independent component analysis (ICA) may be more suitable, assuming the different fluorophores are statistically independent⁶⁸. As a result, the PCA components do not match the shape of the basis spectra, except for the first one, which resembles PpIX₆₃₄ (see Supplementary Fig. 1). Additionally, the fact that the classification results with PCA were worse than those with spectral unmixing could be due to several factors. The scaling of the fluorophore abundances varied greatly, which PCA does not capture well. Furthermore, PCA requires zero-mean data and thus subtracts the mean from the data as one of the first processing steps. However, spectral data is inherently non-negative, leading to unphysical and inaccurate representations.

This paper has explored the effect of different neurosurgically relevant categorizations of brain tumors and tissue on the five previously-characterised³⁵ fluorescence emission spectra. At least four of five

fluorophore abundances were found to vary statistically significantly ($p < 0.01$) among tumor margins, WHO grades, and tissue types. To test the predictive value of the unmixing, we introduced four different ML-based classifiers for tissue type, tumor margins, WHO grades, and IDH status. Each classifier achieved high accuracy, thus promising practical utility for similar systems in the future and demonstrating the differential expression of the fluorophores in different tissue classes and tumors. Together with the fact that the five mathematically optimal PCA-derived components matched very closely to the physically justified fluorophores, this also shows the value and accuracy of the five fluorophores as biomarkers. Moreover, it shows potential for automatic intraoperative classification systems in fluorescence-guided neurosurgery in the future.

Data availability

Patient data cannot be made publicly available for privacy reasons, but spectral data may be shared with individuals at reasonable request. The raw results from the ML models from which Figs. 4–8 were generated are included in Supplementary Data 1.

Code availability

The code is not made publicly available because it primarily handles our specific data. The algorithms used (random forests, multi-layer perceptrons, etc.) were implemented in SciKit-Learn and are available open source at scikit-learn.org or any similar library. However, the code may be made available upon reasonable request.

Received: 19 July 2023; Accepted: 25 June 2024;

Published online: 04 July 2024

References

1. Stepp, H. & Stummer, W. 5-ALA in the management of malignant glioma. *Lasers Surg. Med* **50**, 399–419 (2018).
2. Stummer, W. et al. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *Lancet Oncol.* **7**, 392–401 (2006).
3. Schucht, P. et al. Gross total resection rates in contemporary glioblastoma surgery: results of an institutional protocol combining 5-aminolevulinic acid intraoperative fluorescence imaging and brain mapping. *Neurosurgery* **71**, 927–935 (2012).
4. Widhalm, G. et al. The value of visible 5-ALA fluorescence and quantitative protoporphyrin IX analysis for improved surgery of suspected low-grade gliomas. *J. Neurosurg.* **133**, 79–88 (2019).
5. Valdés, P. A. et al. Quantitative fluorescence using 5-aminolevulinic acid-induced protoporphyrin IX biomarker as a surgical adjunct in low-grade glioma surgery. *J. Neurosurg.* **123**, 771–780 (2015).
6. Stummer, W. et al. Fluorescence-guided resection of glioblastoma multiforme by using 5-aminolevulinic acid-induced porphyrins: A prospective study in 52 consecutive patients. *J. Neurosurg.* **93**, 1003–1013 (2000).
7. Roberts, D. W. et al. Coregistered fluorescence-enhanced tumor resection of malignant glioma: relationships between δ -aminolevulinic acid-induced protoporphyrin IX fluorescence, magnetic resonance imaging enhancement, and neuropathological parameters. *Clin. Artic. J. Neurosurg.* **114**, 595–603 (2011).
8. Valdes, P. A., Millesi, M., Widhalm, G. & Roberts, D. W. 5-aminolevulinic acid induced protoporphyrin IX (ALA-PpIX) fluorescence guidance in meningioma surgery. *J. Neurooncol.* **141**, 555–565 (2019).
9. Kajimoto, Y. et al. Use of 5-aminolevulinic acid in fluorescence-guided resection of meningioma with high risk of recurrence: Case report. *J. Neurosurg.* **106**, 1070–1074 (2007).
10. Motekallemi, A. et al. The current status of 5-ALA fluorescence-guided resection of intracranial meningiomas—a critical review. *Neurosurg. Rev.* **38**, 619–628 (2015).

11. Valdes, P. A. et al. 5-Aminolevulinic acid-induced protoporphyrin IX fluorescence in meningioma: Qualitative and quantitative measurements in Vivo. *Neurosurgery* **10**, 74–82 (1982).
12. Suero Molina, E., Kaneko, S., Black, D. & Stummer, W. 5-Aminolevulinic acid-induced porphyrin contents in various brain tumors: implications regarding imaging device design and their validation. *Neurosurgery* **89**, 1132–1140 (2021).
13. Leunig, A. et al. Detection of squamous cell carcinoma of the oral cavity by imaging 5-Aminolevulinic acid-induced Protoporphyrin IX fluorescence. *Laryngoscope* **110**, 78–83 (2000).
14. Betz, C. S. et al. A comparative study of normal inspection, autofluorescence and 5-ALA-induced PPIX fluorescence for oral cancer diagnosis. *Int. J. Cancer* **97**, 245–252 (2002).
15. Harada, Y., Murayama, Y., Takamatsu, T., Otsuji, E. & Tanaka, H. 5-Aminolevulinic acid-induced Protoporphyrin IX fluorescence imaging for tumor detection: recent advances and challenges. *Int. J. Mol. Sci.* **23**, 6478 (2022).
16. Koenig, F. et al. Diagnosis of bladder carcinoma using protoporphyrin IX fluorescence induced by 5-aminolevulinic acid. *BJU Int.* **83**, 129–135 (1999).
17. Van Der Beek, N., De Leeuw, J., Demmendaal, C., Bjerring, P. & Neumann, H. A. M. PpIX fluorescence combined with autofluorescence is more accurate than PpIX fluorescence alone in fluorescence detection of non-melanoma skin cancer: An intra-patient direct comparison study. *Lasers Surg. Med.* **44**, 271–276 (2012).
18. Kennedy, J. C. & Pottier, R. H. New trends in photobiology: Endogenous protoporphyrin IX, a clinically useful photosensitizer for photodynamic therapy. *J. Photochem. Photobiol. B* **14**, 275–292 (1992).
19. Schipmann, S. et al. Combination of ALA-induced fluorescence-guided resection and intraoperative open photodynamic therapy for recurrent glioblastoma: case series on a promising dual strategy for local tumor control. *J. Neurosurg.* **134**, 426–436 (2020).
20. Sachar, M., Anderson, K. E. & Ma, X. Protoporphyrin IX: the Good, the Bad, and the Ugly. *J. Pharmacol. Exp. Ther.* **356**, 267–275 (2016).
21. McNicholas, K., MacGregor, M. N. & Gleadle, J. M. In order for the light to shine so brightly, the darkness must be present—why do cancers fluoresce with 5-aminolevulinic acid? *Br. J. Cancer* **121**, 631–639 (2019).
22. Olivo, M. & Wilson, B. C. Mapping ALA-induced PPIX fluorescence in normal brain and brain tumour using confocal fluorescence microscopy. *Int. J. Oncol.* **25**, 37–45 (2004).
23. Valdés, P. A. et al. Deferoxamine iron chelation increases delta-aminolevulinic acid induced protoporphyrin IX in xenograft glioma model. *Photochem Photobiol.* **86**, 471–475 (2010).
24. Mazurek, M., Szczepanek, D., Orzyłowska, A. & Rola, R. Analysis of factors affecting 5-ALA fluorescence intensity in visualizing glial tumor cells—literature review. *Int. J. Mol. Sci.* **23**, 926 (2022).
25. Harmatys, K. M., Musso, A. J., Clear, K. J. & Smith, B. D. Small molecule additive enhances cell uptake of 5-aminolevulinic acid and conversion to protoporphyrin IX. *Photochem Photobiol. Sci.* **15**, 1408 (2016).
26. Hagiya, Y. et al. Pivotal roles of peptide transporter PEPT1 and ATP-binding cassette (ABC) transporter ABCG2 in 5-aminolevulinic acid (ALA)-based photocytotoxicity of gastric cancer cells in vitro. *Photodiagn. Photodyn. Ther.* **9**, 204–214 (2012).
27. Kaneko, S. et al. Fluorescence real-time kinetics of protoporphyrin IX after 5-ALA administration in low-grade glioma. *J. Neurosurg.* **1**, 1–7 (2021).
28. Lichtman, J. W. & Conchello, J. A. Fluorescence microscopy. *Nat. Methods* **2**, 910–919 (2005).
29. Vollmer, F., Rettig, W. & Birkner, E. Photochemical mechanisms producing large fluorescence Stokes shifts. *J. Fluoresc.* **4**, 65–69 (1994).
30. Suero Molina, E., Stögbauer, L., Jeibmann, A., Warneke, N. & Stummer, W. Validating a new generation filter system for visualizing 5-ALA-induced PpIX fluorescence in malignant glioma surgery: a proof of principle study. *Acta Neurochir.* **162**, 785–793 (2020).
31. Stummer, W. et al. Technical principles for protoporphyrin-IX-fluorescence guided microsurgical resection of malignant glioma tissue. *Acta Neurochir.* **140**, 995–1000 (1998).
32. Bottiroli, G. et al. Brain tissue autofluorescence: an aid for intraoperative delineation of tumor resection margins. *Cancer Detect. Prev.* **22**, 330–339 (1998).
33. Mochizuki, Y., Park, M. K., Mori, T. & Kawashima, S. The difference in autofluorescence features of lipofuscin between brain and adrenal. **12**, 283–288 <https://doi.org/10.2108/zsj.12.283> (1995).
34. Lifante, J. et al. The near-infrared autofluorescence fingerprint of the brain. *J. Biophotonics* **13**, e202000154 (2020).
35. Black, D. et al. Characterization of autofluorescence and quantitative protoporphyrin IX biomarkers for optical spectroscopy-guided glioma surgery. *Sci. Rep.* **11**, 1–12 (2021).
36. Alston, L. et al. Spectral complexity of 5-ALA induced PpIX fluorescence in guided surgery: a clinical study towards the discrimination of healthy tissue and margin boundaries in high and low grade gliomas. *Biomed. Opt. Express* **10**, 2478 (2019).
37. Montcel, B., Mahieu-Williams, L., Armoiry, X., Meyronet, D. & Guyotat, J. Two-peaked 5-ALA-induced PpIX fluorescence emission spectrum distinguishes glioblastomas from low grade gliomas and infiltrative component of glioblastomas. *Biomed. Opt. Express* **4**, 548 (2013).
38. Li, Y., Rey-Dios, R., Roberts, D. W., Valdés, P. A. & Cohen-Gadol, A. A. Intraoperative fluorescence-guided resection of high-grade gliomas: a comparison of the present techniques and evolution of future strategies. *World Neurosurg.* **82**, 175–185 <https://doi.org/10.1016/j.wneu.2013.06.014> (2014).
39. Alston, L., Rousseau, D., Hebert, M. & Mahieu-Williams, L. Nonlinear relation between concentration and fluorescence emission of protoporphyrin IX in calibrated phantoms. *J. Biomed. Opt.* **23**, 1 (2018).
40. Kaneko, S., Suero Molina, E., Ewelt, C., Warneke, N. & Stummer, W. Fluorescence-based measurement of real-time kinetics of Protoporphyrin IX after 5-Aminolevulinic acid administration in human in situ malignant gliomas. *Clin. Neurosurg.* **85**, E739–E746 (2019).
41. Molina, E. S., Black, D., Kaneko, S., Mütter, M. & Stummer, W. Double dose of 5-aminolevulinic acid and its effect on protoporphyrin IX accumulation in low-grade glioma. *J. Neurosurg.* **137**, 943–952 (2022).
42. Valdés, P. A. et al. A spectrally constrained dual-band normalization technique for protoporphyrin IX quantification in fluorescence-guided surgery. *Opt. Lett.* **37**, 1817 (2012).
43. Bondy, M. L. et al. Brain tumor epidemiology: consensus from the Brain Tumor Epidemiology Consortium (BTEC). *Cancer* **113**, 1953 (2008).
44. Zacharaki, E. I. et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn. Reson. Med.* **62**, 1609–1618 (2009).
45. Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A. & Mengko, T. R. Brain tumor classification using convolutional neural network. *IFMBE Proc.* **68**, 183–189 (2019).
46. Omuro, A. & DeAngelis, L. M. Glioblastoma and other malignant gliomas: a clinical review. *JAMA* **310**, 1842–1850 (2013).
47. Torp, S. H., Solheim, O. & Skjulsvik, A. J. The WHO 2021 Classification of Central Nervous System tumours: a practical update on what neurosurgeons need to know—a minireview. *Acta Neurochir.* **164**, 2453–2464 (2022).
48. Weller, M. et al. Glioma. *Nat. Rev. Dis. Prim.* **1**, 1–18 (2015).
49. Louis, D. N. et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica.* **131**, 803–820 <https://doi.org/10.1007/s00401-016-1545-1> (2016).

50. Guo, J. et al. Biological roles and therapeutic applications of IDH2 mutations in human cancer. *Front. Oncol.* **11**, 644857 (2021).
51. Ahammed Muneer, K. V., Rajendran, V. R. & Paul Joseph, K. Glioma tumor grade identification using artificial intelligent techniques. *J. Med Syst.* **43**, 1–12 (2019).
52. Jose, L. et al. Artificial intelligence-assisted classification of gliomas using whole-slide images. *Arch. Pathol. Lab. Med.* <https://doi.org/10.5858/ARPA.2021-0518-OA> (2022).
53. Ferrer, V. P., Moura Neto, V. & Mentlein, R. Glioma infiltration and extracellular matrix: Key players and modulators. *Glia* **66**, 1542–1565 (2018).
54. Kubben, P. L. et al. Intraoperative MRI-guided resection of glioblastoma multiforme: A systematic review. *Lancet Oncol.* **12**, 1062–1070 (2011).
55. Nabavi, A. et al. Serial intraoperative magnetic resonance imaging of brain shift. *Neurosurgery* **48**, 787–798 (2001).
56. Laws, E. R. et al. Survival following surgery and prognostic factors for recently diagnosed malignant glioma: data from the Glioma Outcomes Project. *J. Neurosurg.* **99**, 467–473 (2003).
57. Li, Y. M., Suki, D., Hess, K. & Sawaya, R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *J. Neurosurg.* **124**, 977–988 (2016).
58. Leclerc, P. et al. Machine learning-based prediction of glioma margin from 5-ALA induced PpIX fluorescence spectroscopy. *Sci. Rep.* **10**, 1–9 (2020).
59. Walke, A. et al. Challenges in, and recommendations for, hyperspectral imaging in ex vivo malignant glioma biopsy measurements. *Sci. Rep.* **13**, 3829 (2023).
60. Liberti, M. V. & Locasale, J. W. The Warburg effect: how does it benefit cancer cells? *Trends Biochem Sci.* **41**, 211–218 (2016).
61. Wold, S., Esbensen, K. & Geladi, P. Principal component analysis. *Chemometr. Intell. Lab. Syst.* **2**, 37–52 (1987).
62. Black, D. et al. Deep learning-based correction and unmixing of hyperspectral images for brain tumor surgery. Preprint at <https://doi.org/10.48550/arXiv.2402.03761> (2024).
63. Alshiekh Nasany, R. & de la Fuente, M. I. Therapies for IDH-Mutant Gliomas. *Curr. Neurol. Neurosci. Rep.* **23**, 225–233 (2023).
64. Van Den Bent, M. J. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician’s perspective. *Acta Neuropathol.* **120**, 297–304 (2010).
65. Di Ieva, A. Angioarchitectural morphometrics of brain tumors: are there any potential histopathological biomarkers? *Microvasc. Res.* **80**, 522–533 (2010).
66. Fürtjes, G. et al. Intraoperative microscopic autofluorescence detection and characterization in brain tumors using stimulated Raman histology and two-photon fluorescence. *Front Oncol.* **13**, 1146031 (2023).
67. Black, D. et al. A spectral library and method for sparse unmixing of hyperspectral images in fluorescence guided resection of brain tumors. *Biomedical Optics Express* (2024) (In Press).
68. Stone, J. V. Independent component analysis: an introduction. *Trends Cogn. Sci.* **6**, 59–64 (2002).
69. Biau, G. & Scornet, E. A random forest guided tour. *Test* **25**, 197–227 (2016).
70. Larose, D. T. & Larose, C. D. k -nearest neighbor algorithm. *Discov. Knowl. Data* 149–164 <https://doi.org/10.1002/9781118874059.CH7> (2014).
71. Noble, W. S. What is a support vector machine? *Nat. Biotechnol.* **24**, 1565–1567 (2006).
72. Murtagh, F. Multilayer perceptrons for classification and regression. *Neurocomputing* **2**, 183–197 (1991).
73. Freund, Y. & Schapire, R. E. A Short Introduction to Boosting. *J. Jpn. Soc. Artif. Intell.* **14**, 771–780 (1999).

Acknowledgements

We thank Carl Zeiss Meditec AG (Oberkochen, Germany) for providing us with the required BLUE400 filter and the OPMI PICO system.

Author contributions

Conception and design: D.B., E.S.M. Acquisition of data: S.Ka., E.S.M. Analysis and Interpretation: D.B., D.By., T.S., A.W., A.D., S.L., W.S., E.S.M. Drafting the article: D.B., E.S.M. Critically revising the article: all authors. Statistical analysis: D.B., E.S.M. Approved the final version of the manuscript on behalf of all authors: E.S.M. Study supervision: E.S.M.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare the following competing interests: Eric Suero Molina received research support from Carl Zeiss Meditec AG. Walter Stummer has received speaker and consultant fees from SBI ALA Pharma, medac, Carl Zeiss Meditec, and NXDC, as well as research support from Zeiss. All the other authors do not have any competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43856-024-00562-3>.

Correspondence and requests for materials should be addressed to Eric Suero Molina.

Peer review information *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024