AI-based segmentation of intraoperative glioblastoma hyperspectral images

Marco La Salvia*[a], Emanuele Torti[a], Marco Gazzoni[a], Elisa Marenzi[a], Raquel Leon[b], Samuel Ortega[b,c], Himar Fabelo[b], Gustavo M. Callico[b], Francesco Leporati[a]

[a]Dept. of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy, IT27100; [b]Institute for Applied Microelectronics (IUMA), University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, E35017; [c]Norwegian Institute of Food Fisheries and Aquaculture Research (NOFIMA), Tromsø, Norway, 9019

ABSTRACT

Glioblastoma surgical resection is a problematic mission for neurosurgeons. Tumour complete resection improves patients healing chances and prognosis, whilst excessive resection could lead to neurological deficits. Nevertheless, surgeons' sight hardly traces the tumour's extent and boundaries. Indeed, most surgical processes result in subtotal resections. Histopathological testing might enable complete tumour elimination, though it is not feasible due to the time required for tissue investigation. Several studies reported tumour cells having unique molecular signatures and properties. Hyperspectral imaging (HSI) is an emerging, non-contact, non-ionizing, label-free and minimally invasive optical imaging technique able to extract information concerning the observed tissue at the molecular level. Here, we exploited extensive data augmentation, transfer learning, the U-Net++ and the DeepLab-V3+ architectures to perform the automatic end-to-end segmentation of intraoperative glioblastoma hyperspectral images meeting competitive processing times and segmentation results concerning the gold-standard procedure. Based on ground truths provided by the HELICoiD framework, we dramatically improved HSIs processing times, enabling the end-to-end segmentation of glioblastomas targeting the real-time processing to be employed during open craniotomy in surgery, thus improving the gold-standard ML pipeline. We measured competitive inference times concerning the standard CUDA environment offered by MatLab 2020a. The HELICoiD fastest parallel version took 1.68 s to elaborate the most prominent image of the database, whilst our methodology performs segmentation inference in 0.29 ± 0.17 s, hence being real-time compliant concerning the 21 seconds constraint imposed on processing. Furthermore, we evaluated our segmentation results qualitatively and quantitatively regarding the ground truth produced by HELICoiD.

Keywords: Glioblastoma, Deep Learning, Hyperspectral Imaging, Intraoperative Computer-Aided Assistance

1. INTRODUCTION

Brain cancer represents the most common central nervous system malignancy, leading subjects to death and morbidity. It originates either directly in the brain or the spinal cord. Physicians usually cluster nervous system tumours into primary, if cancer emerges in the brain, and secondary, or metastasis, if it forms elsewhere in the body spreading up to the brain[1,2]. Medical practitioners categorize brain tumours into various types based on their nature, origin, rate of growth and progression stage. First of all, they can be either benign or malignant lesions. The harmless cells rarely invade neighbouring healthy compartments, present distinct borders, and have a slow progression rate. The malignant cells readily attack adjacent enclosures in the brain or spinal cord, retain fuzzy contours and have a rapid advancement pace. The World Health Organization categorizes brain tumours into four grades (I, II, III and IV)[1-3]. The higher the grade, the faster the rate of growth. Physicians characterize brain tumours according to their progression stages, from 0 to 4. Stage-0 refers to abnormal cancerous biological structures that do not spread to nearby compartments. Stages-1, 2 and 3 denote cells which are cancerous and spreading rapidly. In Stage-4, cancer spreads throughout the body[2,4]. Glioblastoma (GB - grade IV) is the most deadly and typical brain tumour retaining a 5-year survival rate of 5.5%. Early and total resection of grade-II increases the overall 5-year survival rate up to 81%[1].

*marco.lasalvia01@universitadipavia.it; phone +39-0382-985678

Undoubtedly, medical practitioners could preserve numerous lives if they early noticed cancer via prompt and cost-effective diagnosis procedures. Nonetheless, it is challenging to treat cancer at higher stages[2,3,5].

A brain cancer diagnosis can be either invasive or non-invasive. The gold-standard biopsy procedure is an invasive strategy. Namely, it is the histopathological examination of a tissue sample to confirm the malignancy presence. On the other hand, non-invasive approaches comprise a body and brain scanning imaging. The scanning modalities include computed tomography (CT) and magnetic resonance imaging (MRI) of the brain. These imaging modalities help radiologists locate brain diseases, observe disease progression, and prepare surgical planning. Nonetheless, they are subject to inter-reader variability and accuracy concerning the physicians' proficiency [6–8].

Meningiomas represent the non-malignant primary tumour, whose resection can discourage further disease progression, improving the survival probabilities. Nonetheless, complete resection is not always feasible, and it might lead to neurological damage. Hence, surgeons must find a balance between tumour removal and neurological conditions[1–3]. At present, neurosurgeons employ several intraoperative guidance tools for cancer resection assistance. Namely, they extensively use Image Guided Stereotactic (IGS) neuronavigation, Magnetic Resonance Imaging (iMRI), or fluorescent tumour markers like 5-aminolevulinic acid (5-ALA)[2,3,6,7,9]. Nevertheless, these medical procedures exhibit several constraints, namely cost and time, and do not precisely outline lesions borders. Indeed, procedure duration must be reduced as much as possible, being the patient operated on with an open craniotomy. Furthermore, craniotomy and brain shift change the tumour volume in the intraoperative imaging-guided tools. Therefore, there is a present demand to research new imaging modalities that could overcome such limitations[2]. Hyperspectral imaging (HSI) is a non-invasive, non-ionizing and label-free technique, initially designed for remote-sensing and military purposes, becoming more popular in medicine for cancer detection thanks to the recent technological advances[10–12]. Hyperspectral (HS) images measure the reflected or transmitted light from the captured scene, collecting light-matter interaction values associated with several spectral bands, or wavelengths, of the electromagnetic spectrum range[11,13]. HS images comprise multiple shots aligned in neighbouring narrow wavelengths, forming a reflectance spectrum of all the pixels. Therefore, the outcome is an HS cube retaining both the spatial and the spectral information from the sample under analysis. Several studies highlighted that tumour cells present a unique molecular spectral signature and reflectance characteristics[4,8,10,11].

During the last decade, machine and deep learning (ML, DL) solutions emerged as a tool to analyse and cluster different cancer types using HSI. Indeed, HSIs are hardly visually interpretable. Hence, researchers usually carry out HS image analysis via ML approaches. Among the medical subjects concerning ML and HSIs, literature focused on brain, skin, colon and oesophageal cancer[9,10,14–17]. Concerning intraoperative glioblastoma segmentation of HS images, research mainly emerged within the European project HELICoiD (HypErspectraL Imaging Cancer Detection)[9]. Researchers gathered an in vivo human-brain HS database that developed several ML pipelines on, comprising Support Vector Machines (SVMs), KNearest Neighbours (KNN), Principal Component Analysis (PCA), and K-Means Clustering as the supporting algorithms[7]. The main challenge is retrieving a target ground truth to supervise the ML algorithms. Neurosurgeons can only partially identify the tumour and its boundaries when performing a diagnosis assisted with the traditional imaging systems. Therefore, HELICoiD-based ML studies comprised unsupervised algorithms to overcome this problem and perform the automatic segmentation of the intraoperative-captured HSIs.

Here, we research the feasibility of supervised deep learning architectures, namely U-Net++ and DeepLab-V3+, as a proof-of-concept to perform the automatic segmentation of fifteen intraoperative glioblastomas HS images retained from the HELICoiD database. We operated the ground truths coming from the HELICoiD ML-based pipeline for algorithms supervision as it currently represents the gold-standard procedure to retrieve a segmentation map of the brain cancer. Indeed, this procedure represents the only feasible way to label medical data when ground truth is unavailable via pathology-confirmed testing. Specifically, the main goal is to differentiate GB from healthy and other brain tissues, analysing the HSIs end-to-end to improve the time required to process different supervised and unsupervised algorithms in a unique ML pipeline.

2. MATERIALS AND METHODS

This section presents the HSI system used to acquire the in vivo HS brain cancer dataset, and the proposed DL approach for end-to-end HS brain image tissue classification. Here, we describe the selection and the design of the DL methodologies to segment intraoperative glioblastoma HS images and the chosen metrics to evaluate the performance.

2.1 In Vivo Human-Brain HS Dataset and Acquisition System

The HELICoiD intraoperative HS acquisition system comprised a VNIR *push-broom* camera (Hyperspec® VNIR A-Series, Headwall Photonics Inc., Fitchburg, MA, USA)[7–9] to gather data targeting the composition of an in-vivo HS human-brain database[6,7,9,18]. It captured HS images ranging from 400 to 1000 nm in 826 spectral bands and a spatial resolution of 1004 pixels. The HS camera acquires via the push-broom method providing high spectral resolution and relatively high spatial resolution. Nonetheless, the sensor only captures one spatial dimension of the scene while still capturing its entire spectral signature. Hence, the authors designed a spatial scanning to obtain the complete HS cube with a maximum image size of 1004×1787 pixels (129×230 mm). The system also included an illumination device capable of emitting cold light between 400 and 2200 nm. The engineers connected a Quartz Tungsten Halogen (QTH) lamp to the cold light emitter through a fibre optic cable to avoid brain surface exposure to high temperatures[7]. The University Hospital Doctor Negrin of Las Palmas de Gran Canaria (Spain) and the University Hospital of Southampton (UK) installed the intraoperative HS acquisition system. Both the Comité Ético de Investigación Clínica-Comité de Ética en la Investigación (CEIC/CEI) of the University Hospital Doctor Negrin and the National Research Ethics Service (NRES) Committee South Central-Oxford C for the University Hospital of Southampton approved the study and its consent procedures signed by all participating patients.

2.2 Ground-Truth Taxonomy and Labelling

Each HS image pixel received a label according to the taxonomy proposed in Figure 1. The intraoperative GB images present black rubber ring markers employed for the pathological assessment of the image labelling. The neurosurgeon placed the sterilized markers where there was either normal or tumour tissue. In the latter case, histopathological biopsy examination confirmed the cancer presence. During the classification, we assigned the markers to the *background* class[6,9,18]. Researchers recorded the tumours located in a deeper brain layer after superficial resection.
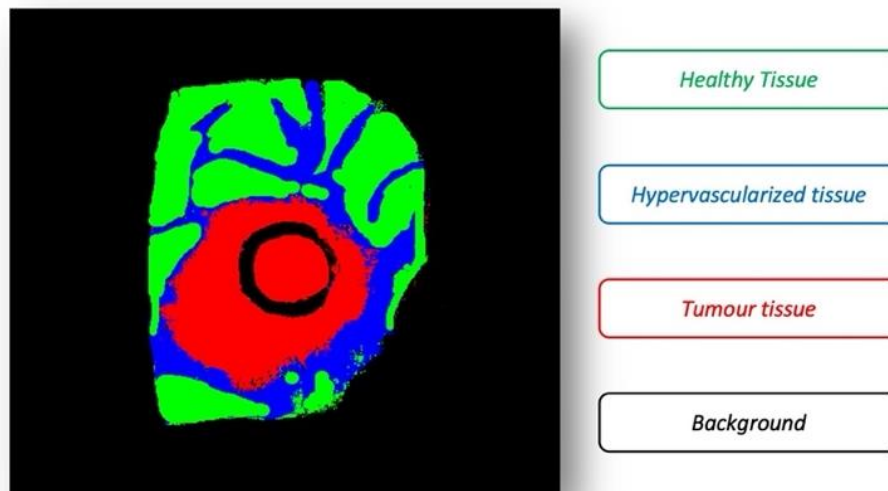

Figure 1. Intraoperative glioblastoma segmentation ground truth taxonomy.

The HELICoiD ML pipeline starts from the Spectral Angle Mapper (SAM) result[7] to segment the entire intraoperative GB hypercube (Figure 2-A). Neurosurgeons used the SAM to create the initial segmentation dataset. They selected reference pixels from healthy and tumour classes inside the circular markers. Therefore, pixels having a similar spectrum to the reference pixels were assigned with the same classes. Neurosurgeons labelled tumour pixels depending on the histopathological assessment. Likewise, physicians labelled both healthy tissue, blood vessels and background by visual inspection according to their experience.

Once the first SAM step ended, the HELICoiD ML pipeline produced the result in Figure 2-B. Since the ML pipeline produced spurious results, we cleaned the segmentation maps via adequate thresholding and filtering, obtaining the smooth mask depicted in Figure 2-C.
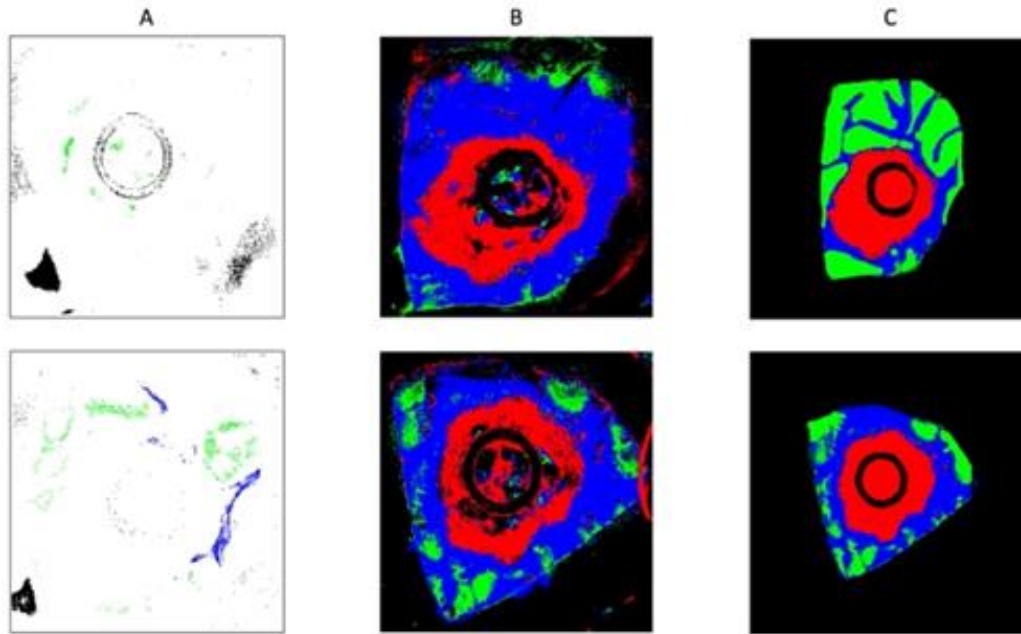

Figure 2. A is the initial ground truth derived from the Neurosurgeon with SAM labelling. B is the HELICoiD ML pipeline result, whilst C is it cleaned version.

2.3 HS Cubes Preprocessing

HS image preprocessing reduces the instrumentation noise and limits the curse of data dimensionality. We normalized the raw data for correction of the dark noise, using a dark and a white reference[7,9]. Moreover, we applied a band-selection algorithm to optimally select a subset of bands. The band reduction algorithm gathers an optimal set of bands preserving the discriminative features while still providing a good classification accuracy. Hence, we retained 128 spectral bands from the original 826[7].

2.4 Deep Learning Methodology

Three CNNs architectures were trained to perform the semantic segmentation of the HS brain lesion images. The U-Net++[19,20], and two versions of the DeepLabV3+ architecture[21] having as backbone structure a ResNet-50, but presenting alternatively 2D and 3D convolutions, were evaluated to perform semantic segmentation of the glioblastoma. Furthermore, we adopted transfer learning[22] to improve the results of the learning-based architecture by exploiting features belonging to the previous training task. The best common practice suggests using neural network architectures pre-trained on similar contexts to overcome small-sized dataset problems and poor segmentation performances. Therefore, all the listed architectures had been optimized based on the HAM10000 dataset[23].

We increased the training set statistical variability by applying data augmentation to the HS images using several methodologies including geometric, filtering, colour transformations and pixel substitution. In particular, we either performed a linear combination of random pixels of tissues belonging to the same category or directly exchanged them. Data augmentation yields promising results in computer vision, significantly reducing overfitting[24,25]. Furthermore, we introduced salt-and-pepper white noise in random image bands to enlarge the training set. The augmentation procedure was carried out iteratively and randomly. Namely, at random one of the data augmentation techniques was applied to the training set with a certain probability to be exploited. Each image at maximum received a predetermined number of augmentations. Such augmentation techniques were not applied to the test sets to prevent the results from being biased.

All architectures were modified to receive input size 384×384×128, concerning height, width, and the number of bands, and we resized the HS GB images accordingly to fit the GPUs RAM memory. We also introduced the L2 weights penalty in the loss function to additionally reduce overfitting. Cross-entropy loss function and the Adam method[26,27] were used for training. The learning step was reduced by multiplying it by the dropping factor. Namely, it steadily and linearly decreased after each predetermined number of epochs. Batch size, number of epochs, learning rate and drop factor period were set to 4, 200, 10^{-4}, and 80 respectively, for all the architectures. Drop factor and L2 penalty were set to 0.75 and $5 \cdot 10^{-3}$, respectively, for the semantic segmentation models.

The test system used to conduct our experiments was equipped with an Intel-i9-9900X CPU, working at 3.5 GHz, 128 GB of RAM, and two 2944-cores NVIDIA RTX 2080. We used MathWorks MATLAB 2020a Release - Deep Learning Toolbox for each network design and implementation.

2.5 Aggregated K-Fold Cross-Validation

Cross-validation is a resampling procedure employed to evaluate DL models on a limited data sample. It is a statistical method whose results in metrics estimations offer lower bias concerning other procedures. It has a single parameter called *k*, which refers to the number of groups in which the data sample is split. When k is as big as the data sample size, the procedure is called leave-one-out cross-validation. As such, the process is often called k-fold cross-validation. The cross-validation technique is primarily used in applied ML to estimate the performance of a model on unseen data, which was not used during the model training[27,28]. The original in-vivo human-brain HS database consisted of twenty-six images from sixteen adult patients[29]. Nine patients had a histopathologically confirmed Grade IV glioblastoma, while the remaining seven patients were either affected by other types of tumours or affected by other pathologies that required a craniotomy. However, only fifteen images offered the necessary ground truth quality obtained from the HELICoiD ML pipeline. We randomly shuffled the original HS dataset comprising the 15 HS images, and performed a leave-one-out cross-validation. Next, each unique group was selected as test data and the model was trained on the remaining groups. Hence, we applied data augmentation to the groups used for training. The model was fit on the training set and evaluated on the test set, retaining the prediction evaluated at each iteration and discarding the model. Therefore, we trained the model k times and recorded its estimate for each test set. Hence, the performance metrics for both classification and semantic segmentation were assessed on the aggregated group of predictions, namely the union of each K-fold test set, generated from each DL architecture through the procedure.

2.6 Performance Evaluation

We computed the occurrence of true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) values to evaluate the DL architectures' performance. Concerning the semantic segmentation task, we assessed the pixel-based occurrences. The assessment outcomes were exploited to compute the following metrics: accuracy, sensitivity, and specificity[27]. These metrics were evaluated over the aggregated prediction set of each architecture, which we conveyed through the k-fold cross-validation strategy. Furthermore, the GPU accelerated computing performance was estimated by assessing the elapsed time for each image inference to compare the researched end-to-end methodology with the HELICoiD ML pipeline processing times.

3. RESULTS AND DISCUSSION

We trained and fine-tuned the three convolutional architectures (Table 1) with a small-sized dataset. We evaluated the performance of the architectures by employing a *leave-one-out* cross-validation methodology. Furthermore, we adopted the taxonomy proposed in Figure 1 as a trade-off for a comprehensive and medically appropriate diagnosis, well-suited for DL classifiers. The structure allows physicians to treat patients according to the highest healthcare criteria whilst still retaining the best feasible segmentation performance.

Discrimination between tumour tissue and the other classes of lesions offered fair measures, meeting specificity ranging from 71% to 92% across both architectures and the considered tissue types (Table 1). We observed the DeepLab V3+ with ResNet50 backbone structure achieving the best results in tumour tissue identification with 85% Specificity. Furthermore, we measured tumour tissue accuracy above 70% for both DeepLab V3+ architectures. On the other hand, the multilabel segmentation retains sensitivity performance below 80%. Considering four distinct group classes induces each set to have fewer examples. Hence, it elicits sparse information regarding the inter-patient variability.

Table 1. U-Net++ and DeepLab V3+ (DLV3 presented in two versions with either ResNet-50 or ResNet-50 3D as backbone structures) segmentation results in terms of pixel-wise Accuracy, Specificity and Sensitivity.

		U-Net++	DLV3 RN50	DLV3 RN3D
Accuracy	Healthy Tissue	0.52	0.52	0.55
	Tumour Tissue	0.61	0.71	0.76
	Hypervascularized Tissue	0.26	0.45	0.42
	Background	0.28	0.36	0.33
Specificity	Healthy Tissue	0.74	0.74	0.73
	Tumour Tissue	0.72	0.85	0.83
	Hypervascularized Tissue	0.81	0.71	0.74
	Background	0.92	1.00	1.00
Sensitivity	Healthy Tissue	0.52	0.52	0.55
	Tumour Tissue	0.61	0.71	0.76
	Hypervascularized Tissue	0.26	0.45	0.42
	Background	0.28	0.36	0.33

The research discussed in this paper introduced a few essential matters. We designed an AI-based system to assist neurosurgeons in outlining the contours of glioblastomas during surgery, despite the limitation of the small-sized HS dataset. In particular, we explored a robust process to conceive DL algorithms and manage the small-sized dataset. We answer the demand for a deep learning end-to-end pipeline to meet the real-time constraints of the surgical procedure[7]. Indeed, CNN architectures are mainly composed of matrix computations that effectively fit the high-performance computing hardware employed in the HELICoiD project. The HELICoiD ML pipeline comprised pre-processing, band selection, SVM, KNN and K-Means clustering. It worked on HS images whose size varied from 329×377 towards 548×459 pixels at most and 128 bands. The ML pipeline yielded processing times ranging from 1.68 s to 2.68 s concerning the same test system employed in this work, both in single and multi-GPUs environments[7]. Nonetheless, we performed inference in single GPU mode on hypercubes sized $384 \times 384 \times 128$, concerning height, width, and number of selected bands, to fit the GPU RAM. We repeated measurement 100 times and yielded an average time of 0.29 [s] with 0.17 [s] of standard deviation. Hence, we proved the feasibility of our end-to-end methodology to improve the gold-standard results mainly limited by the unsupervised algorithms non-deterministic processing times.

The tool should replace the current gold-standard procedure that exploits supervised and unsupervised ML algorithms, with the latter representing the computational bottleneck, to turn the modern DL algorithms into a piece of medical equipment. Recently published articles highlighted that the Food and Drug Administration (FDA) is moving towards the approval of AI-based medical devices[30]. Concerning brain cancer, we designed a proof-of-concept whose results strongly depend on the target ground truth produced via the gold-standard HELICoiD procedure. Figure 3 depicts an example of the semantic segmentation produced by the U-Net++. Other works evaluated the segmentation metrics on the SAM result (Figure 2-A). Nonetheless, supervised CNN architectures need the entire mask to allow the gradient descent algorithm to meet convergence. Hence, not only did we evaluate our metrics over the ground truth produced by the HELICoiD pipeline and not over a few sets of pixels, targeting a future whole mask generation, but also we employed modern DL methodologies on the HELICoiD dataset. Therefore, we must stress that measuring segmentation metrics below a safety threshold, usually placed above 90% in medical contexts, is not necessarily a red flag sign. Indeed, the HELICoiD ground truth does not necessarily label every HS pixel correctly.

Nevertheless, the analysis reveals boundaries. The major one, which yields the others, concerns the dataset dimensions. Indeed, HS imaging is a powerful instrument compared to classical RGB pictures. Several studies highlighted that tumour cells present a unique molecular spectral signature and reflectance characteristics [4,8,10,11]. They allow the classification of pixels of tissues into different aetiologies. HS imaging systems gather brain reflected and transmitted light into several wavelength ranges of the electromagnetic spectrum, enabling potential glioblastoma lesion tracing through DL algorithms. Indeed, each pixel contains meaningful information concerning the captured tissue properties. However, not only are some brain portions transitioning from healthy to malignant tissue, but cancer, healthy and hypervascularized signatures might slightly differ from each other. Each patient possesses a unique tissue signature which causes the test images to be very

different from the training ones, increasing inter-patient variability. Figure 4 represents the spectral signature means and standard deviations of the investigated tissue classes concerning the labels assigned by the HELICoiD ML pipeline. Therefore, a broader dataset should cope with this issue and let CNNs concentrate more on the significant parts of the wavelengths, enhancing the segmentation outcomes accomplished in this work.
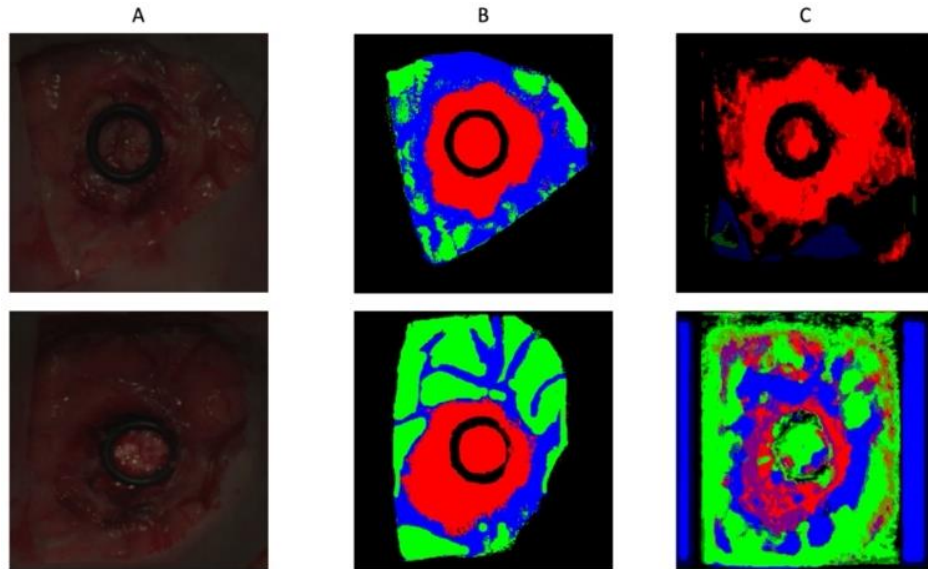

Figure 3. U-Net++ examples of semantic segmentation predictions compared to the cleaned HELICoiD ground truth. A) Synthetic RGB image. B) Cleaned HELICoiD ground truth. C) U-Net++ result.
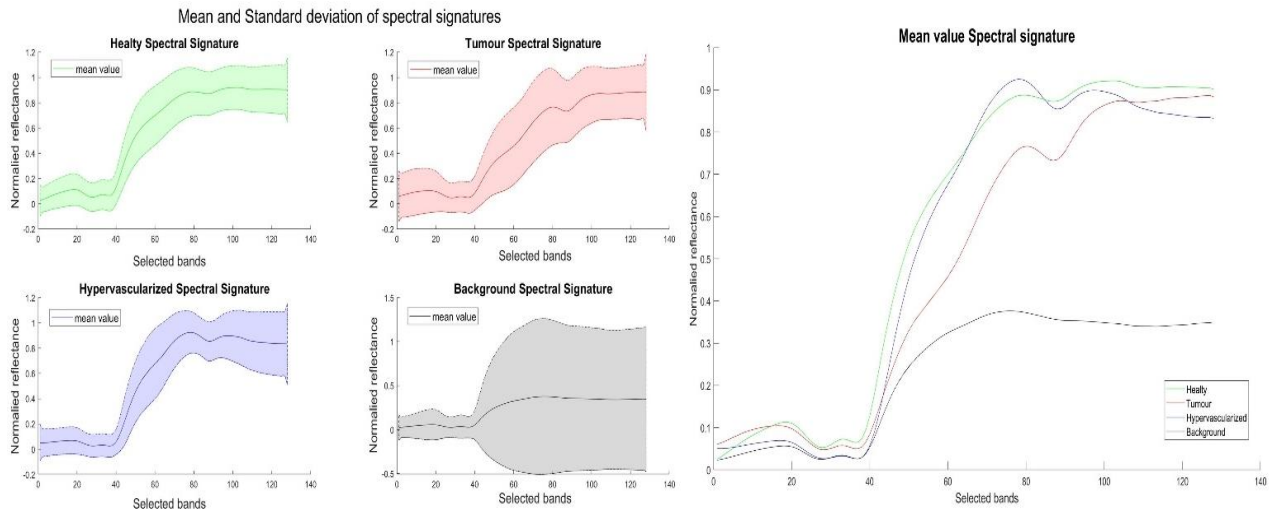

Figure 4. Average spectral signature and standard deviation of all brain HS images tissues and background.

4. CONCLUSIONS

Here, we researched three DL architectures targeting the semantic segmentation of fifteen HS images belonging to the HELICoiD dataset. Modern DL methodologies allow the end-to-end segmentation of the HS images targeting the real-time processing to be employed during open craniotomy in surgery, thus improving the gold-standard ML pipeline. We measured competitive inference times measured with the standard CUDA environment offered by MatLab 2020a, hence

without a custom implementation, concerning the HELICoiD processing times. HELICoiD's fastest parallel version took 1.68 s to elaborate the largest image of the database, whilst our methodology performs segmentation inference in 0.29 ± 0.17 s thoroughly satisfying the real-time constraint, classifying the images in less than 21 seconds. Furthermore, we evaluated our segmentation results both qualitatively and quantitatively in comparison with the ground truth produced by the HELICoiD project.

REFERENCES

[1]   Miller, K. D., Ostrom, Q. T., Kruchko, C., Patil, N., Tihan, T., Cioffi, G., Fuchs, H. E., Waite, K. A., Jemal, A., Siegel, R. L. and Barnholtz-Sloan, J. S., "Brain and other central nervous system tumor statistics, 2021," CA Cancer J Clin **71**(5), 381–406 (2021).

[2]   Tandel, G. S., Biswas, M., Kakde, O. G., Tiwari, A., Suri, H. S., Turk, M., Laird, J. R., Asare, C. K., Ankrah, A. A., Khanna, N. N., Madhusudhan, B. K., Saba, L. and Suri, J. S., "A Review on a Deep Learning Perspective in Brain Cancer Classification," Cancers (Basel) **11**(1) (2019).

[3]   Alifieris, C. and Trafalis, D. T., "Glioblastoma multiforme: Pathogenesis and treatment," Pharmacol Ther **152**, 63–82 (2015).

[4]   Barberio, M., Benedicenti, S., Pizzicannella, M., Felli, E., Collins, T., Jansen-Winkeln, B., Marescaux, J., Viola, M. G. and Diana, M., "Intraoperative Guidance Using Hyperspectral Imaging: A Review for Surgeons," Diagnostics (Basel) **11**(11) (2021).

[5]   Manni, F., van der Sommen, F., Fabelo, H., Zinger, S., Shan, C., Edström, E., Elmi-Terander, A., Ortega, S., Callicó, G. M. and de With, P. H. N., "Hyperspectral Imaging for Glioblastoma Surgery: Improving Tumor Identification Using a Deep Spectral-Spatial Approach," Sensors (Basel) **20**(23), 1–20 (2020).

[6]   Fabelo, H., Ortega, S., Lazcano, R., Madroñal, D., Callicó, G. M., Juárez, E., Salvador, R., Bulters, D., Bulstrode, H., Szolna, A., Piñeiro, J. F., Sosa, C., O, A. J., Bisshopp, S., Hernández, M., Morera, J., Ravi, D., Ravi Kiran, B., Vega, A., et al., "An Intraoperative Visualization System Using Hyperspectral Imaging to Aid in Brain Tumor Delineation," Sensors **18**, 430 (2018).

[7]   Florimbi, G., Fabelo, H., Torti, E., Ortega, S., Marrero-Martin, M., Callico, G. M., Danese, G. and Leporati, F., "Towards Real-Time Computing of Intraoperative Hyperspectral Imaging for Brain Cancer Detection Using Multi-GPU Platforms," IEEE Access **8**, 8485–8501 (2020).

[8]   Leon, R., Fabelo, H., Ortega, S., Piñeiro, J. F., Szolna, A., Hernandez, M., Espino, C., O'Shanahan, A. J., Carrera, D., Bisshopp, S., Sosa, C., Marquez, M., Morera, J., Clavo, B. and Callico, G. M., "VNIR–NIR hyperspectral imaging fusion targeting intraoperative brain cancer detection," Scientific Reports 2021 11:1 **11**(1), 1–12 (2021).

[9]   Fabelo, H., Ortega, S., Kabwama, S., Callico, G. M., Bulters, D., Szolna, A., Pineiro, J. F. and Sarmiento, R., "HELICoiD project: a new use of hyperspectral imaging for brain cancer detection in real-time during neurosurgical operations," https://doi.org/10.1117/12.2223075 **9860**, 986002 (2016).

[10] Torti, E., Leon, R., Salvia, M. la, Florimbi, G., Martinez-Vega, B., Fabelo, H., Ortega, S., Callicó, G. M. and Leporati, F., "Parallel Classification Pipelines for Skin Cancer Detection Exploiting Hyperspectral Imaging on Hybrid Systems," Electronics 2020, Vol. 9, Page 1503 **9**(9), 1503 (2020).

[11] Lu, G. and Fei, B., "Medical hyperspectral imaging: a review," J Biomed Opt **19**(1), 010901 (2014).

[12] Kumar, D. and Kumar, D., "Hyperspectral Image Classification Using Deep Learning Models: A Review," J Phys Conf Ser **1950**(1), 012087 (2021).

[13] Rey-Barroso, L., Peña-Gutiérrez, S., Yáñez, C., Burgos-Fernández, F. J., Vilaseca, M. and Royo, S., "Optical Technologies for the Improvement of Skin Cancer Diagnosis: A Review," Sensors (Basel) **21**(1), 1–31 (2021).

[14] Jansen-Winkeln, B. T., Barberio, M. L., Gockel, I. and Maktabi, M. L., "Feedforward Artificial Neural Network-Based Colorectal Cancer Detection Using Hyperspectral Imaging: a Step Toward Automatic Optical Biopsy," J Am Coll Surg **233**(5), S57 (2021).

[15] Collins, T., Maktabi, M., Barberio, M., Bencteux, V., Jansen-Winkeln, B., Chalopin, C., Marescaux, J., Hostettler, A., Diana, M. and Gockel, I., "Automatic Recognition of Colon and Esophagogastric Cancer with Machine Learning and Hyperspectral Imaging," Diagnostics (Basel) **11**(10) (2021).

[16] Khan, U., Paheding, S., Elkin, C. P. and Devabhaktuni, V. K., "Trends in Deep Learning for Medical Hyperspectral Image Analysis," IEEE Access **9**, 79534–79548 (2021).

[17] Ortega, S., Halicek, M., Fabelo, H., Camacho, R., Plaza De La Luz, M., Godtliebsen, F., Callicó, G. M. and Fei, B., "Hyperspectral Imaging for the Detection of Glioblastoma Tumor Cells in H&E Slides Using Convolutional Neural Networks," Sensors (Basel) **20**(7) (2020).

[18] Fabelo, H., Ortega, S., Ravi, D., Kiran, B. R., Sosa, C., Bulters, D., Callicó, G. M., Bulstrode, H., Szolna, A., Piñeiro, J. F., Kabwama, S., Madroñal, D., Lazcano, R., JO'Shanahan, A., Bisshopp, S., Hernández, M., Báez, A., Yang, G. Z., Stanciulescu, B., et al., "Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations," PLoS One **13**(3), e0193721 (2018).

[19] Ronneberger, O., Fischer, P. and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **9351**, 234–241 (2015).

[20] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. and Liang, J., "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation," IEEE Trans Med Imaging **39**(6), 1856 (2020).

[21] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., "Encoder-decoder with atrous separable convolution for semantic image segmentation," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **11211 LNCS**, 833–851 (2018).

[22] Yosinski, J., Clune, J., Bengio, Y. and Lipson, H., "How transferable are features in deep neural networks?," Adv Neural Inf Process Syst **27** (2014).

[23] Tschandl, P., Rosendahl, C. and Kittler, H., "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Scientific Data 2018 5:1 **5**(1), 1–9 (2018).

[24] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B. and Sánchez, C. I., "A survey on deep learning in medical image analysis," Med Image Anal **42**, 60–88 (2017).

[25] Shorten, C. and Khoshgoftaar, T. M., "A survey on Image Data Augmentation for Deep Learning," J Big Data **6**(1), 1–48 (2019).

[26] Kingma, D. P. and Ba, J. L., "Adam: A Method for Stochastic Optimization," 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings (2014).

[27] Hastie, T., Tibshirani, R. and Friedman, J., "The Elements of Statistical Learning" (2009).

[28] James, G., Witten, D., Hastie, T. and Tibshirani, R., "Bias-Variance Trade-Off for k-Fold Cross-Validation," An Introduction to Statistical Learning - with Applications in R **7**(10), 24 (2013).

[29] Fabelo, H., Ortega, S., Szolna, A., Bulters, D., Pineiro, J. F., Kabwama, S., J-O'Shanahan, A., Bulstrode, H., Bisshopp, S., Kiran, B. R., Ravi, D., Lazcano, R., Madronal, D., Sosa, C., Espino, C., Marquez, M., de La Luz Plaza, M., Camacho, R., Carrera, D., et al., "In-Vivo Hyperspectral Human Brain Image Database for Brain Cancer Detection," IEEE Access **7**, 39098–39116 (2019).

[30] Yu, K. H., Beam, A. L. and Kohane, I. S., "Artificial intelligence in healthcare," Nature Biomedical Engineering 2018 2:10 **2**(10), 719–731 (2018).