Manifold Embedding and Semantic Segmentation for Intraoperative Guidance With Hyperspectral Brain Imaging

Daniele Ravì, Himar Fabelo, Gustavo Marrero Callicò, *Member, IEEE*, and Guang-Zhong Yang, *Fellow, IEEE*

***Abstract*— Recent advances in hyperspectral imaging have made it a promising solution for intra-operative tissue characterization, with the advantages of being non-contact, non-ionizing, and non-invasive. Working with hyperspectral images *in vivo*, however, is not straightforward as the high dimensionality of the data makes real-time processing challenging. In this paper, a novel dimensionality reduction scheme and a new processing pipeline are introduced to obtain a detailed tumor classification map for intra-operative margin definition during brain surgery. However, existing approaches to dimensionality reduction based on manifold embedding can be time consuming and may not guarantee a consistent result, thus hindering final tissue classification. The proposed framework aims to overcome these problems through a process divided into two steps: dimensionality reduction based on an extension of the T-distributed stochastic neighbor approach is first performed and then a semantic segmentation technique is applied to the embedded results by using a Semantic Texton Forest for tissue classification. Detailed *in vivo* validation of the proposed method has been performed to demonstrate the potential clinical value of the system.**

***Index Terms*— Manifold embedding, hyperspectral imaging, semantic segmentation, brain cancer detection.**

I. Introduction

WHILE malignant primary brain tumours rank only 13th in the list of cancer incidence rates, their particularly poor prognosis elevates it as the fifth most common cause of cancer deaths in those under the age of 65. Among children, they are the second most common form of cancer and the most common cause of cancer death. Gliomas are the most frequent primary brain tumours and they are currently incurable. Research has shown that life expectancy increases with an extensive resection of these tumours. Gross total resection is a challenging task since gliomas infiltrate the surrounding tissue and their borders are indistinctive and difficult to identify. Currently, different techniques have been developed to help achieve this goal, but none have succeeded in reliable real-time and non-invasive tissue differentiation. For example, neuro-navigation is plagued with brain shift, while ultrasound is highly operator dependant and Magnetic Resonance Imaging (MRI) is still not accessible for real-time intra-operative use. Consequently, clinical routines are still based on subjective visual assessment by the surgeon, who decides which areas should be removed during the operation. Without accurate guidance, margin definition is poor even for experts. This is mainly due to significant visual variations between adjacent structures of the brain when they are partially obscured, and because tumour structures vary considerably across patients in terms of location, size, and extension, prohibiting the use of priors on shape and location [1]. Hyperspectral imaging, also called imaging spectroscopy, is an emerging technology that can assist surgeons to classify tumour from healthy tissue in real-time. In this paper, we propose a novel manifold embedding framework where the output generated from a hyperspectral image is semantically segmented into a tumour map. The proposed method has the main goal of delineating the exact boundaries of the brain tumours, allowing a complete resection of the malignant cells while saving as much healthy brain tissue as possible. The proposed system can also be used to improve diagnosis and treatment planning, as well as follow-up of individual patients. Hyperspectral imaging is a non-contact, non-ionizing and minimally-invasive sensing technique. Whereas a conventional camera captures images in three color channels (red, blue and green), a hyperspectral camera captures data over a large number of contiguous and narrow spectral bands [2]. Classification of the tissue under evaluation can be achieved by analysing the reflectance or fluorescence of every pixel in the hyperspectral image and the spatial structures that these pixels form. For example, it has been demonstrated that biological tissues

Manuscript received February 10, 2017; revised April 10, 2017; accepted April 11, 2017. Date of publication April 24, 2017; date of current version August 31, 2017. This work was supported in part by the European Commission through the FP7 FET Open Programme under Grant ICT-2011.9.2 and in part by the European Project HELICoiD HypErspectral Imaging Cancer Detection under Grant 618080. *(Corresponding author: Daniele Ravì.)*

D. Ravì is with The Hamlyn Centre, Imperial College London, London SW72AZ, U.K. (e-mail: d.ravi@imperial.ac.uk).

H. Fabelo and G. M. Callicò are with the Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria, E35017 Las Palmas, Spain (e-mail: hfabelo@iuma.ulpgc.es; gustavo@iuma.ulpgc.es).

G.-Z. Yang is with The Hamlyn Centre, Imperial College London, London SW72AZ, U.K. (e-mail: g.z.yang@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2017.2695523

exhibit fluorescent properties when excited with ultra-violet light, and significant differences in these fluorescent properties occur between malignant and healthy tissues [3]. Previous works demonstrate that hyperspectral imaging can be used for certain cancer detection in animals [4], [5]. Thus far, limited work has been performed by using the technique for detecting cancer in vivo [6]–[8] and none has been used for brain tissue. The work described in this paper presents results obtained from our collaborative EU project HELICoiD [9]–[11] where four universities, three industrial partners and two hospitals were involved. To the best of our knowledge, we are the first to explore hyperspectal imaging for the identification of cancer tissue during in-vivo brain surgery. The rest of the paper is organised as follows. We first briefly review, in Section II, the current state-of-the-art of brain tumour segmentation techniques developed in other domains. We then describe the mechanism behind the optical processes involved in the acquisition of a hyperspectral image and how tissue characterisation can be achieved in Section II-A. Theory behind the dimensionality reduction is provided in Section II-B. The preprocessing pipeline, the proposed method used to evaluate the quality of a dimensionality reduction output, and the developed manifold embedding framework are presented in Section III. The protocol used for acquiring the hyperspectral images in the operating theatre and the obtained database are described in Section IV. Finally, we report and discuss the results in Sections V, and Section VI concludes the paper.

II. Related Work

Existing approaches to brain tumour visualisation are commonly based on the use of Computed Tomography (CT) or MRI. Current methods try to automatically discover glial tumour [12], [13], meningioma [14]–[18] or glioma subtypes [19] that can occur in the brain tissues. In general, medical imaging approaches for brain tissue characterization can be categorized into two types: i) the generative probabilistic methods that segment the brain by exploiting detailed prior information about the appearance and spatial distribution of the different tissue types [20]–[26] and ii) discriminative approaches that learn from labelled images about different appearance of tissues, analysing local features that are relevant to tumour segmentation task [27]–[33]. These features usually represent local intensity differences, intensity distributions, texture and spatial regularity of tissue labels. Once these features are extracted, they are fed into a classifier to obtain a semantic segmentation of the image. The semantic segmentation can be thought of as an extension of the popular scene classification problem where the entity to classify is no longer the whole image, but a single group of pixels. The output of the semantic segmentation classifier, therefore, highlights the tumour classification map. Generative models represent the state-of-the-art for brain tissue segmentation. However, the main problem is that they require a significant effort for transforming an arbitrary semantic interpretation of the image into appropriate probabilistic models [1]. Hyperspectral imaging is a relatively new area of research that can be also used in this context. In the literature, few techniques based on this technology have been proposed for tissue characterization and tumour analysis [4]–[8]. Due to the high-dimensionality of the hyperspectral images, most of the existing approaches classify the tissues by exploiting just spectral information without taking into account the spatial correlations, essential for describing the underlying brain structures. Therefore, they classify the spectral signature of each sample independently using standard classifiers such as the Support Vector Machine (SVM) [8]. We propose to extend this idea and reduce the problem of segmenting a hyperspectral image by analysing an embedded version of it so that the spatial information can be considered in the low-dimensional space. This makes our method capable of processing a hyperspectral image in real-time.

A. Tissue Characterization Through Hyperspectral Image

Due to significant improvements and miniaturization of hyperspectral cameras in recent years, hyperspectral imaging is becoming a well-established technique for disease diagnosis in a wide range of medical applications. In this section, we briefly describe the mechanisms behind the optical processes steps involved during a hyperspectral image acquisition. Specifically, when the light hits a biological tissue, it can deflect in three main components called absorbed, reflected and scattered light [34]. While reflection appears on surfaces built from a non-absorbing powder, or from fibers and polycrystalline material, scattering occurs where there is a spatial variation in the refractive index of the substances, generally caused by inhomogeneous structures; finally, absorption happens when the photons' energy matches the energy gap of the molecules of the tissues and is usually significantly high in haemoglobin, melanin, and water [35]. The penetration depth of the light depends on the molecular composition of the tissue. Consequently, the absorption, reflection and the scattering characteristics change across tissues, showing differences also during the progression of a disease [36]. Measurement of these tissue characteristics can provide quantitative diagnostic information about pathology. For example, absorption spectra characterize the concentration and oxygen saturation of haemoglobin, which can reveal if an angiogenesis or hypermetabolism is present [37]. In other words, when the light is absorbed by the tissue, it is either converted to heat or radiated in the form of luminescence (fluorescence and phosphorescence). Therefore, cells in different disease states may have different intrinsic characteristics, resulting in different fluorescence emission spectra. Hyperspectral images have the capability of measuring this fluorescence, making a possible real-time investigation of the tissues for diagnosis purposes. These changes of the optical properties make this technology an ideal non-invasive probe for tissue analysis [38]. The output of a hyperspectral camera is a three dimensional matrix that contains samples arranged in columns (x dimension), lines (y dimension) and bands (wavelength dimension). The third dimension represents the spectral information and it usually consists of several hundreds of spectral bands that cover a contiguous portion of the light spectrum, such as the visible and near-infrared range (VNIR), and the shortwave infrared range (SWIR). This matrix is known as a hyperspectral cube. To classify samples in the hyperspectral cube and identify the various

pathological conditions, our method tries to exploit both the spatial information and spectral signatures at the same time.

B. Dimensionality Reduction

Hyperspectral images tend to have a high dimensionality, making real-time processing difficult. In the context of this paper, the high dimensionality is due to the large number of wavelength bands that create the hyperspectral cube (for example, an image that has 400×400 pixels captured at 200 wavelengths will be represented by 160000 vectors lying in the space $\mathbb{R}^{200}$). In order to handle a hyperspectral image adequately for real-time applications, its dimensionality needs to be reduced through the projection of the hyperspectral cube to a space with only a few dimensions.

Dimensionality reduction transforms high-dimensional data into a reduced dimensional representation that is still capable of describing the initial data. The intrinsic dimensionality is, therefore, the minimum number of parameters required to accurately describe all the observed properties of the data. Since in our application we want to use dimensionality reduction just to project the data into a lower dimensional space and make the subsequent tissue categorization as precise as possible, the constraint of preserving all these initial properties is not strictly necessary. Consequently, in this case, the intrinsic dimension can be made small, as long as the new representation allows the correct classification of all the tissue types.

The problem of dimensionality reduction can be mathematically defined as follows. Assume that we have a dataset represented by an $n \times D$ matrix X consisting of n data vectors x_i $(i \in 1, 2, \ldots, n)$ with dimensionality D and intrinsic dimensionality of d (where $d \ll D$). In mathematical terms, the points in dataset X are lying on a manifold with dimensionality d that is embedded in the D-dimensional space. Dimensionality reduction is the process that transforms the dataset X into a new dataset Y with dimensionality d, while the observed properties of the data are retained as much as possible:

$$X \in \mathbb{R}^D \rightarrow Y \in \mathbb{R}^d \tag{1}$$

Since the geometry of the data and the intrinsic dimensionality d are in general unknown a priori, dimensionality reduction is, therefore, an ill-posed problem that can only be solved by making some assumptions. Many algorithms for dimensionality reduction have been developed in the past. Principal Components Analysis (PCA) [39] is one of the most popular linear techniques for dimensionality reduction. However, its effectiveness is limited by its global linearity: its purpose is to capture the new dimension so that the original variance in the data is preserved. Preserving local features such as continuity, conformity or manifold structure sometimes can be more important than preserving global properties like PCA does. Moreover, previous studies have shown that nonlinear techniques outperform linear techniques on complex artificial tasks [40].

The idea behind most of the nonlinear methods is to decompose the nonlinear structures into linear subspaces in the same way that some curves can be approximated by locally linear functions, and use solutions similar to PCA on each of these subspaces. For example, Isomap [41], Locally Linear Embedding (LLE) [42] and Hessian Locally Linear Embedding (HLLE) [43] are convex techniques that attempt to find the global minimum of some objective function. These approaches are based on k-NN graphs and the main algorithms share a similar flow chart. First, K nearest neighbours for each point are generated, then some properties of the manifold are computed by looking at neighbourhoods previously found. Finally, an embedding that preserves these local properties is proposed. In contrast to Isomap, which preserves geodesic distances by computing the shortest path between pairs of points on the high-dimensional space, LLE tries to analyze local properties of these points and describes them as a linear combination of their k-nearest neighbours. The advantage of LLE is that it does not suffer from common issues that occur when, for example, the geodesic distances are computed on graph with short-circuits. For this reason, LLE is in general more robust. HLLE is instead a variant of LLE that assumes a local isometric representation of the data and tries to minimize the curviness of the high-dimensional manifold. This minimization is obtained by an eigen-analysis of the local Hessian matrix. The advantage of using the local Hessian matrix is that it is invariant to differences in the positions of the manifold datapoints. Another well-known approach is the Kernel PCA (KPCA) [44] that can be seen as the nonlinear counterpart of the traditional PCA. The main difference is that KPCA does not apply the PCA directly to the high-dimensional points, but instead these points are first projected into a nonlinear feature space by exploiting a kernel function. Unfortunately, it is not always clear how to select this kernel function and to this end, approaches such as the Maximum Variance Unfolding (MVU) [45], are proposed to learn a proper kernel matrix automatically. Laplacian Eigenmaps (LE) [46] is another dimensionality reduction approach based on the Laplace-Beltrami operator. Here, a heat kernel is used to weight a Laplacian of adjacency graph extracted from the manifold. Ziemann and Messinger [47] have recently proposed an adaptive variant of the LE approach that is based on an Adaptive Nearest Neighbours (ANN) graph to separate target data from the background data.

All the approaches described so far attempt to obtain the optimal solution with respect to the objective function considered. This is usually satisfied thanks to the convexity of the eigen-decomposition analysis used to find the solution. However, there are alternative techniques, although not optimal, which are worth considering. For example, the Sammon mapping [48] that, rather to explicitly represent the transformation function, simply provides a measure of how well the results of the transformation reflect the structures present in the original dataset. Autoencoder and Stuck Autoencoder [49] are techniques based on training a deep neural network (DNN). These deep learning based methods [50] use unsupervised training to recreate the input vector samples and project them into a lower dimensional space. Their network configurations are relatively straightforward. The number of nodes in each hidden layer is first gradually

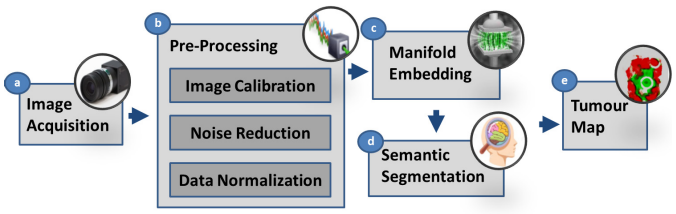


Fig. 1. Processing pipeline of the proposed approach for the semantic segmentation of hyperspectral images.

reduced to encode the data into a lower dimensional space and then gradually increased to decode back to the original high-dimensional space. Many variations of the Autoencoder have been proposed to increase the stability against small variations of the input pattern. These methods include the Sparse Autoencoder [51], the Denoising Autoencoder [52] and the Contractive Autoencoder [53]. Although deep learning approaches can obtain good results since they are data driven, the main disadvantage is that they usually require big datasets to avoid overfitting.

In recent years, there are extensive interests in developing manifold embedding techniques explicitly for hyperspectral imaging. For example, in [54] a manifold learning approach is proposed for hyperspectral ocean water data analysis and in [55] a manifold coordinate representation is used to represent hyperspectral images for bathymetry retrieval. Olson and Doster [56] proposed an unsupervised anomaly detection framework that uses nonlinear techniques to learn a model for the non-anomalous data collected from an 8-band multispectral sensor as well as panchromatic infrared images. Cheng et al. [57] proposed a Biologically Inspired Feature (BIF) embedded in a high-dimensional space for identifying Peripapillary Atrophy (PPA). Finally, in [58] a Sparse Transfer Manifold Embedding (STME) approach is presented for the training of hyperspectral remote sensing dataset that has unbalanced or small-size classes.

III. Methods

The pipeline of the proposed approach is summarised in Fig. 1. It consists of a hyperspectral image acquisition block followed by a pre-processing step, then the proposed manifold embedding framework and finally a semantic segmentation classifier that generates the tumour map. Each of these blocks is described in more detail as follows.

A. Image Acquisition

To capture the hyperspectral images, Headwall's Hyperspec VNIR A-Series and Headwall's Hyperspec NIR X-Series cameras have been used. The VNIR camera captures 826 spectral bands, covering the spectral range of 400-1000 nm with a spectral resolution of 2-3 nm, and each pixel has a dimension of 128.7 μm × 128.7 μm, while the NIR camera captures 172 spectral bands, covering the spectral range of 900-1700 nm and each pixel has a dimension of 0.48 μm × 0.48 μm. The light source used in our system is a wide spectrum light that provides a uniform illumination on the subject. The distance between the lens and the subject

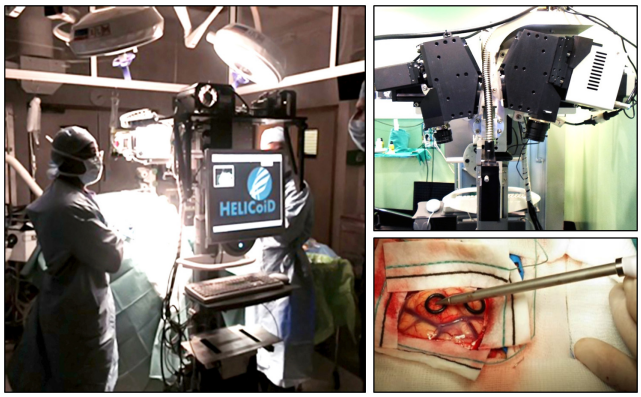


Fig. 2. Setup of our system during the in vivo tumour brain resection. Two cameras (VNIR and NIR) are used for the image acquisition while two sterilised fiducial markers are placed on the brain surface for labelling purposes.

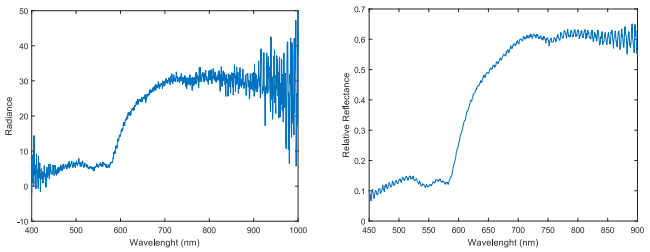
is kept constant at 40 cm and manual image focus is used. The setup of our system during the in vivo tumour brain resection is depicted in Fig. 2.

B. Pre-Processing

Although during the acquisition of a hyperspectral image we have used a wide spectrum light that provides a uniform illumination, the light uniformity along all areas of the brain cannot always be guaranteed due to the 3D cortical folding of the brain that can introduce occlusions and shadows. Moreover, other external factors such as inconsistent environment light can lead to variation in the spectral domain. These effects can be reduced by using a pre-processing pipeline that normalizes the reflectance of the hyperspectral image. Our pre-processing pipeline is divided into three main steps: i) image calibration, ii) noise reduction, and iii) data normalization. For image calibration, the variations caused by the non-uniform illumination over the surface of the captured scene are adjusted. The hyperspectral image is calibrated using white and dark reference images. These reference images are acquired separately inside the operating theatre before the procedure. The white reference image is obtained from a standard white tile; the dark reference image is obtained instead, by keeping the camera shutter closed. The calibration is performed pixel-wise, and for each vector r_i on the i^{th} cell of an uncalibrated hyperspectral cube R, the corresponding calibrated vector y_i is obtained using the following equation:

$$y_i = 100\frac{r_i - d_i}{w_i - d_i} \tag{2}$$

where w_i and d_i are respectively the i^{th} vector in the white reference hyperspectral cube W and the i^{th} vector in the dark reference hyperspectral cube D. The second step of the pre-processing pipeline aims to remove the noise generated by common imperfections of the CCD (Charge Coupled Device) cells of the hyperspectral camera sensor. Fig. 3(a) shows the signature of a tissue sample captured by the VNIR camera. In the figure, we can see that there is a high fluctuation in the extreme bands. This fluctuation is due to the low ratio between



(a) (b)

Fig. 3. Spectral signatures in a tissue sample captured by the VNIR camera. 3(a) and 3(b) contain respectively the signature before and after applying our pre-processing pipeline.

the noise and the real signal in those bands. Due to the limited performance of the CCD sensor in these ranges, we decide to discard all the distorted bands in the extreme sides (before the 450 nm and after 950 nm for the VNIR camera and before the 950 nm and after 1650 nm for the NIR camera). For the remaining bands, the noise removal is obtained by using the HySIME algorithm by assuming that the reflectance at a given band is well modelled by linear regression [59]. The last consideration taken into account in this section is that, due to the scanning procedure of the hyperspectral camera and the non-uniformity of the brain surface, some areas of the image can be captured with a different illumination biasing the successive classification. In order to avoid this issue, a unit magnitude normalization of the brightness is performed for each pixel by using Eq. 3. The bright normalized value y_{ib_norm} of the i^{th} cell in band b is obtained through the division of initial value y_{ib} by the norm of the vector y_i. Fig. 3(b) illustrates the final spectral signature after applying the proposed pre-processing pipeline.

$$y_{ib\_norm} = \frac{y_{ib}}{||y_i||} \tag{3}$$

C. Manifold Embedding

1) Evaluation of a Dimensionality Reduction: Except for very simple cases, it is practically difficult to decide a priori which dimensionality reduction approach performs better for a specific problem, and different techniques can lead to contrastingly different results. On the other hand, it is also hard for human eyes to judge the quality of a given mapping and the suitability of a specific technique if the dimensionality of the original data is very high. Therefore, it is necessary to develop formal measures that can judge the quality of a given mapping [60]. Quality measures can be formulated by ranking all the differences between the distances from each point to all others and comparing the low-dimensional representation to the original data [61]. A criterion Q_Y for the assessment of dimensionality reduction methods was proposed in [62], where a global-structure-holding performance is defined. The method takes into consideration both local and global properties, with the aim to faithfully reflect the intrinsic capability of dimensionality reduction. Specifically, a shortest path tree is first generated from the k-neighborhood graph and then, the Spearman's rank order correlation is used to compute a global-structure assessment Q_{GB}. The overall result Q_Y is defined as a linear combination of this global assessment Q_{GB}, and other local assessments such as Q_k proposed in [63] and Q_t proposed in [64].

However, these criteria do not take into account the visual quality of the embedded image that is an important feature for the visualization of the output or for the automatic segmentation of the tissues. For this reason, we propose a new quality score, described in Eq. 4, that considers also some visual characteristics such as the Global Contrast Factor (GCF) [65] and a measure that depends on the shape of the histogram of the embedded image. Contrast in image processing is usually defined as a ratio between the darkest and the brightest spots of an image. The GCF uses contrast values at various resolution levels in order to compute the overall contrast value. Moreover, the GCF measures the richness of detail as perceived by a viewer. Histograms instead plot the number of samples for each tonal value. By looking at the histogram for a specific image, a human observer will be able to judge the entire tonal distribution at a glance. Specifically, with the last term of Eq. 4 we try to capture if the dimensionality reduction approach is capable to use all the possible values of the histogram to represent the embedded image and to do so, we count how many histogram bins in each channel overcome a threshold z (with z fixed to 10 in our experiments).

$$M(Y) = \frac{Q_Y}{k_1} + \frac{\sum_{i=1}^{d} GCF(Y_i)}{k_2} + \frac{\sum_{i=1}^{d} hist_i > z}{k_3} \tag{4}$$

where Q_Y is the global measure proposed in [62], whereas k_1, k_2, k_3 are the terms used to normalize each component of the formula in the range [0, 1]. Therefore, the range of the final score will be [0, 3].

By looking at the overall contrast and histogram in the embedded image, we aim to select the approach that can best separate different tissue classes. Visual characteristics are usually not considered by existing metrics designed to evaluate dimensionality reduction techniques. For this reason, we believe that Eq. 4 provides a better indication of which approach is more suitable for dimensionality embedding in this application.

2) t-SNE: The main contribution of this paper is the implementation of a new method of manifold embedding based on T-distributed Stochastic Neighbours (t-SNE) [66]. In this section, we highlight the key details involved, while our proposed method is explained later in section III-C.3. t-SNE is a machine-learning algorithm for dimensionality reduction. It is a nonlinear technique that is particularly well suited to embedding high-dimensional data into a space with few dimensions (i.e. two or three). Specifically, it models each high-dimensional object by a low-dimensional point such that similar objects are modelled by nearby points and dissimilar objects are modelled by distant points. The main advantage of t-SNE is that it captures the local structure of the high-dimensional data and at the same time reveals global structures such as clusters. For these reasons, t-SNE has been used in a wide range of applications, including cancer research, and bioinformatics [67]. The t-SNE algorithm consists of

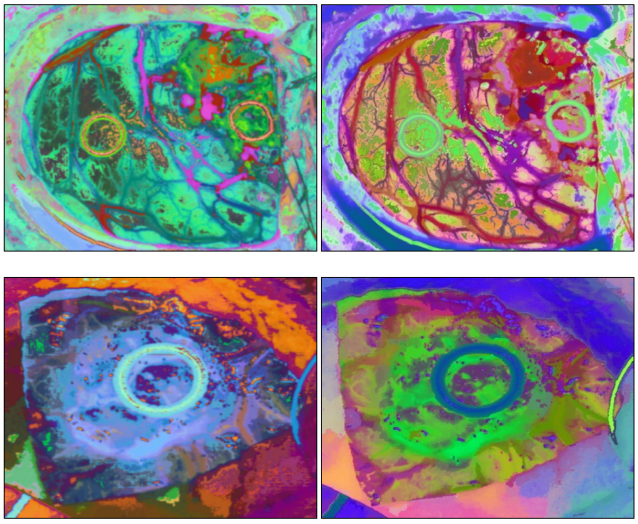


Fig. 4. Two examples generated by the t-SNE on two different hyperspectral images captured from the same subject. Here similar tissues are represented with inconsistent colors making the successive tissue characterization difficult to achieve.

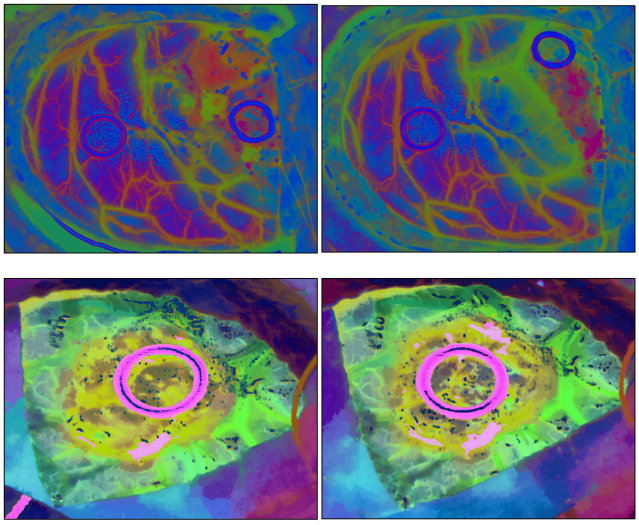


Fig. 5. Two examples generated by the proposed approach on two different hyperspectral images captured from the same subject. In this case, similar tissues are represented with similar colors.

two main stages. Given a set of N high-dimensional objects $x_1, \ldots, x_n$, t-SNE first constructs a probability distribution p_{ij} over pairs x_i and x_j in such a way that similar objects have a high probability of being picked, whilst dissimilar points have an infinitesimal probability of being picked.

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \tag{5}$$

where $p_{j|i}$ is defined as follows:

$$p_{j|i} = \frac{exp(-\|x_i - x_j\|^2/2\alpha_i^2)}{\sum_{k \neq i} exp(-\|x_i - x_k\|^2/2\alpha_i^2)} \tag{6}$$

and α_i is the variance of the Gaussian that is centred on data point x_i and its value depends on the density of the data. Specifically, smaller values of α_i are used in denser parts of the space and large values otherwise. Second, t-SNE aims to learn a d-dimensional map of $y_1, \ldots, y_N$ that reflects the similarities p_{ij} as well as possible. To this end, it defines a similarity measure q_{ij} computed over the points y_i and y_j that represent the low-dimensional counterparts of the high-dimensional data points x_i and x_j.

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq i}(1 + \|y_k - y_i\|^2)^{-1}} \tag{7}$$

Finally, the locations of points y_i in the map are determined by minimizing Eq. 8, which represents the Kullback-Leibler divergence of the distribution Q from the distribution P.

$$KL(P\|Q) = \sum_{i \neq j} p_{ij} log \frac{p_{ij}}{q_{ij}} \tag{8}$$

This minimization is obtained through the gradient descent function initialized by extracting each y_i from a random sampling of an isotropic Gaussian with small variance that is centered around the origin. Embedding a hyperspectral image using t-SNE may not guarantee a consistent embedding due to i) the random nature of the approach and ii) the lack of a fixed coordinate system, which prohibits the comparison of the embedded results across different tissue samples [68]. Fig. 4 shows two examples of outputs generated by the t-SNE on two different images captured on the same subject. As we can see, in the embedded images similar tissues are represented with different colors making the successive tissue characterization difficult to achieve.

3) Proposed Embedding Method: To address the problem mentioned above, we have proposed a new manifold embedding approach called Fixed Reference T-distributed Stochastic Neighbours (FR-t-SNE) that includes the following four steps. The first three steps are used to train the manifold offline and the last to embed a new image. The details of the proposed approach are described in Algorithm 1. Specifically, in Step 1, an optimal reference system is fixed to maintain a consistent manifold embedding along with all the images and circumvent the lack of a fixed coordinate system. In Step 2, the manifold is gradually discovered on the training set using the predefined fixed reference. In Step 3, a DNN is trained to learn how to map the samples from the high dimensional space into the low dimensional space. Finally, a hyperspectral image H is embedded in real-time using the trained DNN. In order to obtain the optimal reference coordinate system in Step 1, the spectral samples collected from all the training images are clustered and K centroids $Cx = (Cx_1, Cx_2, \ldots Cx_K)$ are computed. Considering these centroids as a synthetic hyperspectral cube, the standard t-SNE routine is performed on it, producing the embedded output $Cy = (Cy_1, Cy_2, \ldots Cy_K)$. Cx and Cy are vectors that represent respectively the fixed coordinate system in the original space and in the embedded space. These will be used in Step 2 to predict the embedded value of each spectral sample of the training hyperspectral images. This prediction is obtained by initializing point y_i in Eq. 7 using a lookup table that contains the vectors Cx as indexes and the vector Cy as values. More specifically,

Algorithm 1 Proposed Manifold Embedding

Input:

$TrainSet$ ▷ Hyperspectral images used for training

H ▷ Hyperspectral image to embed

Variables:

S_i ▷ Set of high dimensional samples obtained on hyperspectral image I

T_i ▷ Set of embedded training samples obtained on hyperspectral image I

C_x ▷ Centroids obtained in the high dimensional space

C_y ▷ Counterpart for centroids C_x living in the low dimensional space

K_i ▷ Closest centroids for each sample in S_i obtained by the 1-NN classifier

D_o ▷ Set of high dimensional samples obtained on the entire TrainSet

D_e ▷ Set of embedded training samples obtained on the entire TrainSet

DNN_Emb ▷ Deep neural network used for embedding

Output:

E ▷ Embedded image

procedure $D_e, D_o \leftarrow$ TRAINING($TrainSet$)
  $D_o \leftarrow \emptyset$
  ForEach I in $TrainSet$ ▷ Step 1: Find Fixed Reference System
    $S_i \leftarrow Extract\_Samples(I)$;
    $D_o \leftarrow D_o \cup S_i$;
  End
  $C_x \leftarrow$k-means(D_o);
  $C_y \leftarrow tSNE(C_x)$; ▷ End Step 1
  $D_e \leftarrow \emptyset$; $D_o \leftarrow \emptyset$
  ForEach I in $TrainSet$ ▷ Step 2: Discover Manifold
    $S_i \leftarrow Extract\_Samples(I)$;
    $K_i \leftarrow$1-NN(S_i, C_x, C_y);
    $Initialize\_tSNE(K_i)$;
    $T_i \leftarrow tSNE(I)$;
    $D_e \leftarrow D_e \cup T_i$;
    $D_o \leftarrow D_o \cup S_i$;
  End ▷ End Step 2
  $DNN\_Emb \leftarrow$ Train_DNN(D_o,D_e) ▷ Step 3: Training the model
end procedure
procedure $E \leftarrow$ EMBEDDING(H,DNN_Emb) ▷ Step 4: Embedding
  $S_h \leftarrow Extract\_Samples(H)$;
  $E \leftarrow DNN\_Emb(S_h)$;
end procedure

for each sample within a hyperspectral image, the closest centroid Cx_s (according the Euclidean distance) is selected and the corresponding counterpart Cy_s is used to initialize y_i. In this way, the gradient descent routing, required to minimize Eq. 8, will start from the fixed reference value of that point and just few rounds of update will be performed to find the embedded result.

In Step 2, for each training image, the input samples S_i and the corresponding counterpart T_i are collected in a final selection D_o and D_e of original and embedded samples. Once all the training images are processed and the entire manifold is discovered, the collected vectors D_o and the target values D_e are used to train a DNN to predict the embedded values of each sample of a new hyperspectral image H. During the training of our system, the number of vectors collected in D_o and D_e reaches approximately 1 million, which is a relatively high number to avoid overfitting issues. Although training the FR-t-SNE may take days when the number of hyperspectral images becomes large, we believe that this is not a real limitation since this procedure only needs to be performed once offline. There are also extensive means of hardware acceleration, e.g., with the use of GPU architecture.

From a mathematical point of view, the objective function of our dimensionality reduction approach is an approximation of the t-SNE Kullback-Leibler divergence of Eq. 8, obtained by using a DNN so that the embedded images are always consistent. Evidence for this consistence property obtained by our approach is shown in Fig. 5, where two example results obtained from two different hyperspectral images of the same subject are displayed. This is in contrast to Fig. 4, where similar tissues are represented using different colors. As will be shown later in the experimental section of this paper, the different tissues are better recognized with the proposed method and there is a significant classification improvement over the standard t-SNE approach.

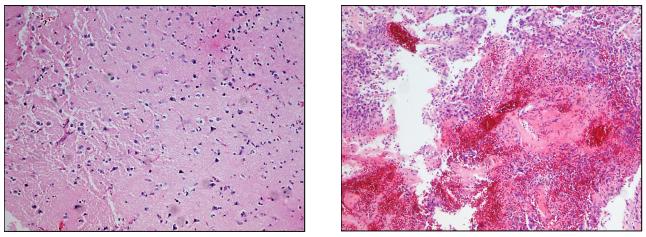
A further improvement of t-SNE is the resources efficiency for real-time processing. Due to the extremely high spectral resolution of a hyperspectral image, it has been observed that consecutive bands are correlated and for this reason some of them can be removed before applying manifold embedding. Specifically, each set of adjacent bands $[(z-1)t+1, zt]$ is replaced by their averaged band Y_z. The equation used for this process is described as follows:

$$Y_z = \frac{\sum_{i=(z-1)t+1}^{i=zt} Y_i}{t} \text{with } z \in \left[1, 2, \ldots, \frac{B}{t}\right] \tag{9}$$

where t is the number of averaged bands and B is the total number of hyperspectral bands. Although the final number of bands considered by the dimensional reduction approach is inversely proportional to this parameter t, a large value of it can discard discriminative information from the spectral signature. The optimal value is selected empirically in this study as $t = 6$ for the VNIR camera and $t = 1$ for the NIR camera. Eq. 9 is a filter that allows not only to speed-up the embedding process, but also to reduce noise effectively.

D. Semantic Segmentation

To obtain the tumour classification map, a semantic segmentation approach is performed on the obtained embedded images. One of the most popular approaches designed for this task is based on random forests and is called Semantic Texton Forest (STF) [69]. Random forests approaches have been used already for many medical applications such as for the detection of Alzheimer's disease [70], delineation of myocardium in real-time [71], and lesion segmentation from multimodal MRI [72]. These approaches are effective as they also consider the fact that within an image, similar low-level feature responses can represent different objects and each single feature is not sufficient for characterizing the object to which it belongs. Therefore, approaches like STF tries to exploit the spatial arrangement of these low-level features to



(a) (b)

Fig. 6. Examples of histological images obtained from normal brain tissue in 6(a) and glioblastoma tumour brain tissue in 6(b).

increase the object discrimination. In our case, the spatial information is useful to distinguish, for example, blood vessels in healthy tissues with respect to those surrounding the tumour as they have similar spectral responses, but the former have elongated shapes and the latter agglomerated shapes. These differences in the appearance can help to discriminate different tissue classes. For these reasons, for the semantic segmentation task, we decided to exploit the STF approach and also a more recent method called DCT-STF [73] that has shown more accurate results when the classes of the database are un-balanced.

IV. Validation Data

A total of 33 hyperspectral brain images have been captured from 18 different patients during neurosurgical operations. Local Ethics Committee approvals were obtained from the National Health Service (REC reference 14/SC/0108) and from the Servicio Canario de la Salud (REC reference 130069). During each surgical procedure, two sterilised fiducial markers are placed in the brain as shown in Fig. 2: one in the area that belongs to a tumour and the other in the healthy brain tissue. The evaluation of the brain tissue is performed by the surgeon based on visual appearance, anatomical relationship of the sulci and gyri, and on the MRI or CT performed on the patient before the surgical procedure. The precise location of the markers is also noted using the Image-Guided Surgery (IGS) system to identify the location of the markers on the brain. This pointer defines the position of the fiducial markers with respect to the pre-operative MRI or CT data. Afterwards, the hyperspectral image is captured with our proposed system and the tissue samples contained inside the markers are resected from the patient and sent to the pathologist. Pathologic diagnosis is carried out by a professional pathologist that undergoes standard hematoxylin and eosin staining to establish a definitive histopathological diagnosis. Fig. 6 shows two examples of the histological images obtained from normal brain tissue and tumour, respectively. The results of these diagnoses are used as the ground truth reference for the training process. After the hyperspectral images have been captured, the system is moved out of the surgical zone and the neurosurgeon continues with tumour resection until a new image can be captured depending on the nature of the surgical procedure. Table I summarizes the obtained database describing in detail the type of tumours diagnosed in different neurosurgical operations.

TABLE I
Summary of the Type of Tumours Diagnosed in Different Neurosurgical Operations

Tissue	Description	Number of Patients	Number of Images
Healthy	Normal	11	17
Primary Tumour	Glioblastoma (Grade IV)	8	12
	Anaplastic Oligodend. (Grade III)	1	4
	Ganglioglioma (Grade I)	1	2
	Meningioma (Grade I)	1	1
Secondary Tumour	Lung Carcinoma	2	2
	Lung Adenocarcinoma	1	1
	Renal Carcinoma	1	1
	Breast Carcinoma	1	3

A. Segmented Reference Maps Generation

In order to be robust against imaging artifacts, intensity changes and shape variations, supervised machine learning methods typically require a substantial amount of training data. Building the training database for our proposed platform is an important step. Due to the nature of the clinical setting, limited training images are available in our case. Another issue of having a small database is that it can affect the results in terms of overfitting. Overfitting occurs when a statistical model describes random error or noise instead of the underlying relationship. A model that has been overfitted will generally have poor predictive performance, as it can exaggerate minor fluctuations in the data. A second problem with the generation of the current database is related to the label assigned to each sample inside a hyperspectral image. These labels are important for creating the segmented reference maps required for supervised training of the semantic segmentation classifier. However, the real labels for the entire brain are unknown and only a small portion of tissue is extracted and analysed from the pathologists. Moreover, these tissue samples are not always 100% homogeneous and may contain mixed tissues. Having mixed tissue samples can compromise tissue characterization due to inconsistent labelling. Also, some image artifacts, such as specular highlights due to non-polarized light, can bias the classification step and they need to be removed from the region under consideration. To overcome these issues, we propose a semi-automatic reference maps generation pipeline that combines the results provided by the pathologists together with two other segmented image outputs. The details of this process are explained in Fig. 7. Specifically, the diagnosis provided by the pathologist in step a) is used to create a manually segmented image in step b). Here, a predefined color table (red for healthy tissue, green for tumour tissue and white for the background) is used to draw a 2D bitmap image and paint regions around each of the available marker. The remaining area of the brain is represented with an unlabelled region marked as black. To eliminate subjective errors during manual segmentation, an automatic segmentation is performed in parallel on the embedded image, steps c) and d). This automatic segmentation is computed using k-means (k=15 is selected experimentally). Both the manually and automatically segmented images are combined together in step e) to remove all the possible outliers (image artifacts, mixed tissues, and human error). Specifically, all the automatically generated

TABLE II
PIXELS DISTRIBUTION IN EACH CLASS TYPE

Class	Number of pixels	(%) of pixels
Unlabelled Brain Tissue	3.619.611	52.39%
Tumour	65.580	0.95%
Healthy	57.453	0.83%
Others	3.166.818	45.83%

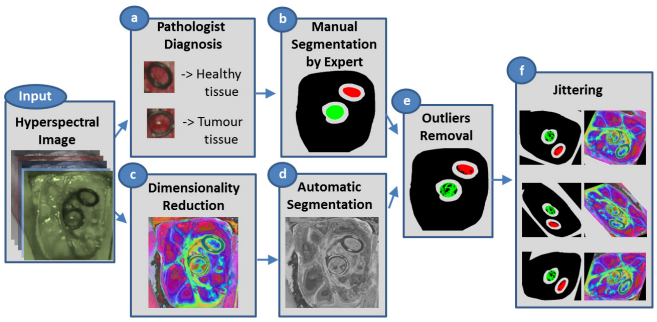


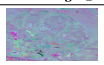
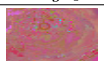
Fig. 7. Proposed semi-automatic segmented reference maps generation pipeline.

segments that belong to a specific marker are arranged in a descending order according to their dimensionality and merged together in a greedy strategy until 60% of the total area is covered. All the remaining small segments are removed since they can represent possible outliers. When all the regions have been processed, the segmented reference map of the current hyperspectral cube is completed. The last step f) of the pipeline is then applied. This block generates synthetic images through a procedure called jittering. It augments the training database by exploiting a sequence of random transformations applied to the initial images. These transformations include rotations, crops, geometric distortions and scaling. For each pair of embedded image and the corresponding segmented reference map, we augment them up to six times. This allows us to extend the database from 33 images to 231 images. As described later in Section V, the synthetic images are just used for training the semantic segmentation model, instead the initial 33 images are used for testing. The distributions of the pixels in the different classes, before the jittering, are provided in Table II.

V. RESULTS

To evaluate the proposed manifold embedding approach, we have compared our solution against 22 other different state-of-the-art methods. The methods considered in these experiments were implemented using the Matlab Toolbox for Dimensionality Reduction [87]. Three hyperspectral images, captured by our system, are used for validation. Table III shows the outputs of these methods by projecting the hyperspectral images in a low dimension space consisting of three channels. The results obtained by Eq. 4 on these three embedded images are averaged and used for rank ordering the different methods. As we can see from the results, the more we go down in Table III, the more visual structures of the brain are visible in the embedded images and this supports the meaning of Eq. 4. This table gives us a first indication of which dimensional reduction approaches are suitable for this application. Since the final experiments related to tissue classification are CPU intensive, exhaustive consideration of all the 22 approaches is not practical. Therefore, from now,

TABLE III
OUTPUTS OF THE DIFFERENT EMBEDDED APPROACHES SORTED ACCORDING THE PROPOSED QUALITY SCORE

Approach	$Image_1$	$Image_2$	$Image_3$	Score
LLC [74]				1.18
Diffusion Maps [75]				1.47
Landmark MVU [76]				1.50
LLE [42]				1.53
Hessian LLE [43]				1.53
NPE [77]				1.57
LTSA [78]				1.58
PCA [39]				1.63
Autoencoder [49]				1.66
Isomap [41]				1.71
LLTSA [79]				1.73
Landmark Isomap [80]				1.76
CCA [81]				1.78
SPE [82]				1.85
Prob. PCA [83]				1.90
Sammon [48]				1.98
Factor Analysis [84]				2.06
LPP [85]				2.08
MVU [45]				2.25
Fast MVU [86]				2.33
Laplacian [46]				2.38
FR-t-SNE				2.61
t-SNE [66]				2.66

just the best two approaches – t-SNE and FR-t-SNE – are taken into account.

In order to evaluate the next experiments, the following metrics are analyzed:

1) Mean-Class Accuracy: Is obtained averaging the accuracies achieved in each of the classes. It is a more reliable measure than the overall accuracy when, as in this case, the sample distributions for the same classes are limited in number, causing an unbalanced dataset.
2) Sensitivity: Is the proportion of actual positives which are correctly identified as positives by the classifier.
3) Specificity: Is the proportion of the actual negatives which the classifier successfully identifies as negative.

Mathematically, sensitivity and specificity are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{10}$$

$$Specificity = \frac{TN}{TN + FP} \tag{11}$$

where TP is the number of true positives, FN is the number of false negatives, TN is the number of true negatives and FP the number of false positives.

The validation of these experiments is instead based on two case studies, CS_1 and CS_2. These case studies differ in the type of images used for training the classifiers and they allow us to evaluate different properties of the system. Specifically, in CS_1 all the synthetic images (6×33 images) are considered in the training-set and the original 33 embedded images are used in the test-set. This case study allows us to understand if the classifier is capable of learning the inter-patient variability. Instead, in CS_2, each surgical procedure is used separately in the test-set and the classification model is trained using the embedded images belonging to all the other patients. The case study CS_2 describes the results in a real scenario and it analyses the capability of the system to transfer the inter-patient tumour classification to new patients that have never been seen before. In our experiments, the results in CS_2 are obtained by training six models for six different patients and averaging all the results obtained. Of note is that the obtained results are computed by taking into account just the labelled pixels.

The first experiment is performed to define some basic parameters such as the classifier and the intrinsic dimension d to consider. Intrinsic dimension larger than three are not considered due to memory limitation and real-time considerations. The obtained results are plotted in Fig. 8 and they show that, for case study CS_1, the camera NIR with an intrinsic dimension equal to three is the configuration that provides the best performance. Instead, for case study CS_2, the camera VNIR with intrinsic dimension equal to one, is the one selected for the best performance. These configurations are therefore fixed respectively for CS_1 and CS_2 in our final algorithm. From this experiment, we can also see that classifier DCT-STF always obtains better performance with respect to STF. Therefore, DCT-STF is selected for final implementation.

The second experiment in this work is to compare our method against t-SNE. The results are reported in Table IV where mean-class accuracy, sensitivity, and specificity

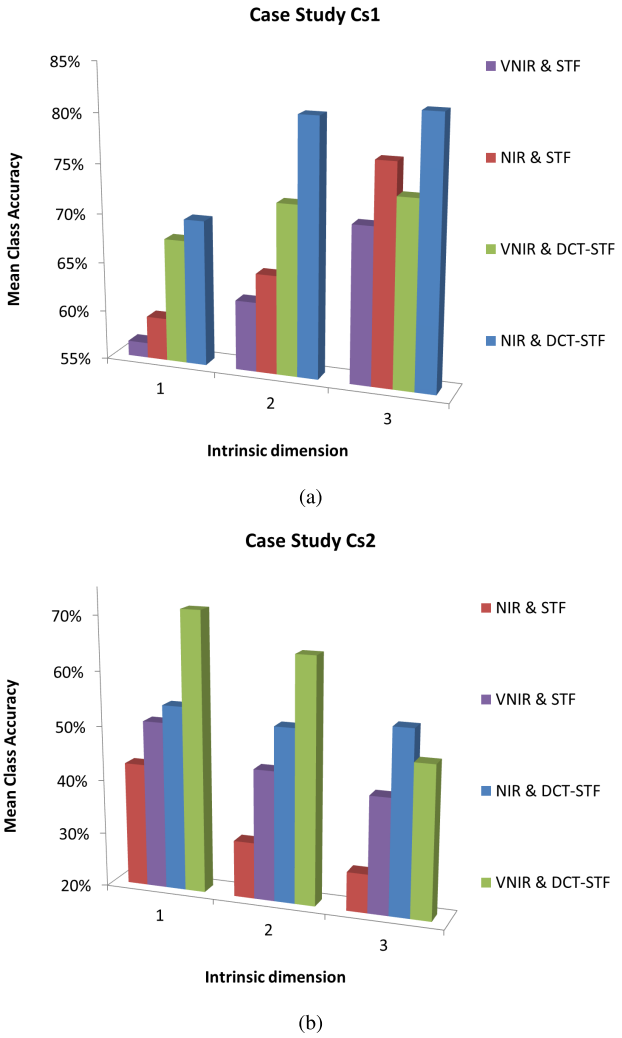


Fig. 8. Mean-class accuracy obtained by the proposed approach on two different case studies when the intrinsic dimension and the classifier are systematically verified.

obtained by our proposed approach and t-SNE are listed for each case study. For case study CS_1, namely when the classifier is trained with the synthetic images and tested on the original embedded images, our approach obtains 2.03% more than t-SNE on mean-class accuracy. Whereas in case study CS_2, when the classifier is trained by leave-one-out test, the results on the mean-class accuracy are 19.94% better than t-SNE. These results suggest that the the fixed reference coordinate system proposed in our approach is an important step to improve the classification results.

The final study of this paper is to show the feasibility of performing our proposed method in real-time. First, we want to highlight that for a hyperspectral image with N rows and M columns, the time complexity of the t-SNE is $\mathcal{O}(N * M)^2$ [66]. This makes the t-SNE not suitable for real time processing. Instead, the DNN proposed in our FR-t-SNE approach that consists of a 4-multilayer perceptron with around 80 nodes in each layer, has a computational time of $\mathcal{O}(N * M)$ resulting in a faster solution. To further emphasize this aspect, we have recorded the averaged computational time required by each block of the proposed system. The CPU

TABLE IV
RESULTS OBTAINED BY THE PROPOSED APPROACH AGAINST T-SNE ON THE TWO DIFFERENT CASE STUDIES

					Specificity			Sensitivity		
Case Study	**Camera Type**	**Approach**	**Intr. Dimen.**	**Mean-Class Acc.**	**Tumour**	**Healthy**	**Others**	**Tumour**	**Healthy**	**Others**
CS_1	NIR	t-SNE	3	79.87%	**86.56%**	83.97%	97.26%	**86.88%**	78.45%	74.29%
	NIR	FR-t-SNE	3	**81.90%**	86.08%	**89.15%**	**98.79%**	80.91%	**86.80%**	**78.00%**
CS_2	VNIR	t-SNE	1	51.56%	90.58%	**94.05%**	86.86%	39.47%	27.28%	**87.94%**
	VNIR	FR-t-SNE	1	**71.50%**	**91.66%**	92.80%	**86.84%**	**52.95%**	**74.80%**	86.76%

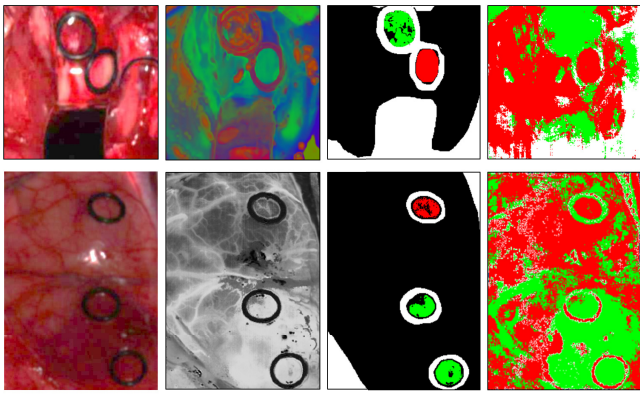


Fig. 9. Two example clinical cases by using the proposed hyper-spectral system and analysis framework. The first row shows a NIR camera image with intrinsic dimension equal to three whilst in the second row a VNIR camera image with intrinsic dimension of one is considered. In (a), the RGB image, (b) the obtained embedded output, (c) the segmented reference maps and (d) the classification map obtained by the DCT-STF classifier.

TABLE V
COMPUTATIONAL TIME OBTAINED BY EACH BLOCK OF THE PROPOSED SYSTEM WHEN THE T-SNE AND THE PROPOSED APPROACH ARE USED

Camera Type	Pre-Proces.	Embedding		Classif.	Total	
		t-SNE	FR-t-SNE	DCT-STF	t-SNE	FR-t-SNE
VNIR	22(s)	14532(s)	2(s)	16(s)	14556(s)	40(s)
NIR	5(s)	14770(s)	2(s)	5(s)	14777(s)	12(s)

timing is recorded on an Intel Xeon E7-8890 v4 with 24 cores running at 2.2Ghz and with 64 GB of DDR memory, and the results are reported in Table V. Although these results show online processing with a total computational time that requires several seconds for each hyperspectral image, with cluster computing and further software optimization, they can be accelerated to provide real-time results. Compared to t-SNE, our approach is ∼7000 times faster, requiring just 2 seconds to process an image, against the 4 hours required by t-SNE, which demonstrates again the relative merit of our solution.

For completeness, in Fig. 9 we show two examples of visual output generated by our system. This figure includes the RGB image, the embedded output, the segmented reference maps and the classification map obtained by the DCT-STF classifier for the case studies. Fig. 10 shows an example of the hyperspectral image samples where the healthy and tumour tissues are compared in the high dimensional and embedded space, respectively.

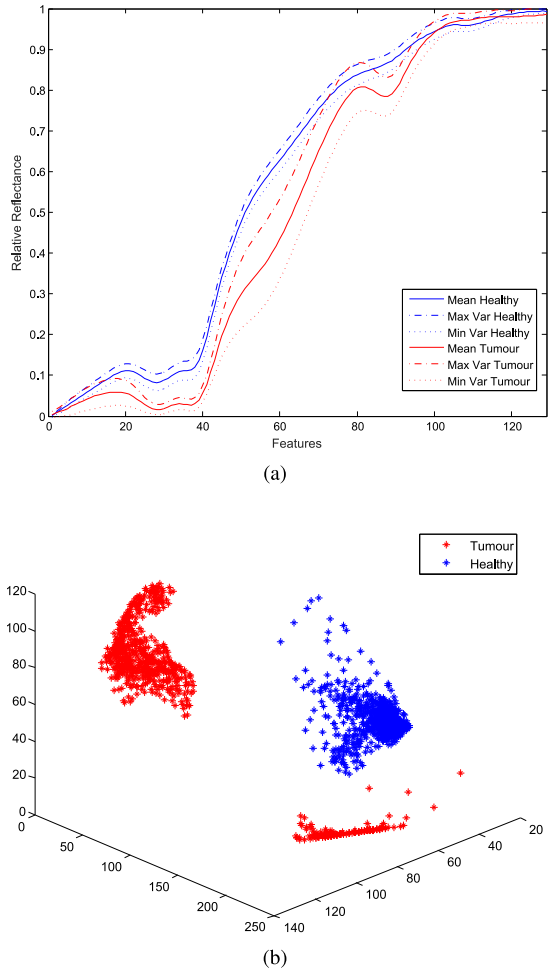


Fig. 10. Comparison of healthy tissue and tumour tissue represented in: (a) the high dimensional space, and (b) the embedded 3D space.

VI. CONCLUSION

In this paper, we have proposed a novel manifold embedding framework FR-t-SNE with which the output generated from a hyperspectral image can be used as input for a semantic segmentation classifier of brain tissues in vivo, in situ. The proposed method aims to determine the boundaries of tumours, saving healthy brain tissue and allowing a complete resection of the malignant cells. Conventional diagnoses of internal tumours are based on excisional biopsy followed by histology or cytology. The main weaknesses of the traditional approach are twofold. Firstly, it is invasive with many potential side effects and complications; and secondly, diagnostic information is not available in real-time and requires off-line histopathology sample preparation and analysis. The proposed

system can overcome these problems and tumour resection can be greatly improved during surgical procedures, thus reducing the risk of disease recurrence. Moreover, the proposed system can help in understanding the cancer progression. The real-time nature of the techniques can improve surgical accuracy, providing additional information that can also reduce the probability of erroneous resectioning of healthy tissue. From the results presented in this paper, it can be seen that the high quality and accuracy of the obtained tumour maps can be achieved by using a suitable embedding approach combined with a classifier that takes into account not only the spectral value of each sample, but also their spatial context. The relative merits of our approach against other existing methods are also demonstrated in this paper, highlighting the strengths of our proposed framework in terms of processing speed and accuracy.

ACKNOWLEDGMENT

The authors would like to thank Samuel Ortega from ULPGC and Dr. Adam Szolna from the Hospital Doctor Negrín for their valuable contributions.

REFERENCES

[1] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[2] M. T. Eismann, *Hyperspectral Remote Sensing*. Bellingham, WA, USA: SPIE, 2012.

[3] N. Ramanujam, "Fluorescence spectroscopy of neoplastic and non-neoplastic tissues," *Neoplasia*, vol. 2, nos. 1–2, pp. 89–117, 2000.

[4] H. Akbari *et al.*, "Hyperspectral imaging and quantitative analysis for prostate cancer detection," *J. Biomed. Opt.*, vol. 17, no. 7, pp. 076005-1–076005-10, 2012.

[5] N. Amharref *et al.*, "Discriminating healthy from tumor and necrosis tissue in rat brain tissue samples by Raman spectral imaging," *Biochim. Biophys. Acta (BBA)-Biomembranes*, vol. 1768, no. 10, pp. 2605–2615, 2007.

[6] A. Cerussi, N. Shah, D. Hsiang, A. Durkin, J. Butler, and B. J. Tromberg, "*In vivo* absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," *J. Biomed. Opt.*, vol. 11, no. 4, p. 044005, 2006.

[7] C. V. Dinh, M. Loog, R. Leitner, O. Rajadell, and R. P. Duin, "Training data selection for cancer detection in multispectral endoscopy images," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 161–164.

[8] H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka, "Cancer detection using infrared hyperspectral imaging," *Cancer Sci.*, vol. 102, no. 4, pp. 852–857, 2011.

[9] (2017). *HELICoiD*. [Online]. Available: http://helicoid.eu/

[10] H. Fabelo *et al.*, "HELICoiD project: A new use of hyperspectral imaging for brain cancer detection in real-time during neurosurgical operations," *Proc. SPIE*, vol. 9860, p. 986002, May 2016.

[11] S. Kabwama *et al.*, "Intra-operative hyperspectral imaging for brain tumour detection and delineation: Current progress on the HELICoid project," *Int. J. Surgery*, vol. 36, no. 2, p. S140, 2016.

[12] E. D. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, and H. Duffau, "Glioma dynamics and computational models: A review of segmentation, registration, and in silico growth algorithms and their clinical applications," *Current Med. Imag. Rev.*, vol. 3, no. 4, pp. 262–276, 2007.

[13] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, vol. 58, no. 13, p. R97, 2013.

[14] M. Kaus *et al.*, "Segmentation of meningiomas and low grade gliomas in MRI," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Cambridge, U.K.: Springer, 1999, pp. 1–10.

[15] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. R. Murtagh, "Automatic segmentation of non-enhancing brain tumors in magnetic resonance images," *Artif. Intell. Med.*, vol. 21, nos. 1–3, pp. 43–63, 2001.

[16] Y.-F. Tsai, I.-J. Chiang, Y.-C. Lee, C.-C. Liao, and K.-L. Wang, "Automatic MRI meningioma segmentation using estimation maximization," in *Proc. 27th Annu. Int. Conf. Eng. Med. Biol. Soc. (IEEE-EMBS)*, Jan. 2005, pp. 3074–3077.

[17] E. Konukoglu *et al.*, "Monitoring slowly evolving tumors," in *Proc. 5th IEEE Int. Symp. Biomed. Imag., Nano Macro (ISBI)*, May 2008, pp. 812–815.

[18] M. B. Cuadra, M. D. Craene, V. Duay, B. Macq, C. Pollo, and J.-P. Thiran, "Dense deformation field estimation for atlas-based segmentation of pathological MR brain images," *Comput. Methods Programs Biomed.*, vol. 84, no. 2, pp. 66–75, 2006.

[19] L. Weizman *et al.*, "Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in MRI," *Med. Image Anal.*, vol. 16, no. 1, pp. 177–188, 2012.

[20] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 885–896, Oct. 1999.

[21] M. R. Kaus, S. K. Warfield, A. Nabavi, P. M. Black, F. A. Jolesz, and R. Kikinis, "Automated segmentation of MR images of brain tumors 1," *Radiology*, vol. 218, no. 2, pp. 586–591, 2001.

[22] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Med. Image Anal.*, vol. 8, no. 3, pp. 275–283, 2004.

[23] K. M. Pohl *et al.*, "A unifying approach to registration, segmentation, and intensity correction," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI*. Palm Springs, CA, USA: Springer, 2005, pp. 310–318.

[24] F. O. Kaster, B. H. Menze, M.-A. Weber, and F. A. Hamprecht, "Comparative validation of graphical models for learning tumor segmentations from noisy manual annotations," in *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Beijing, China: Springer, 2010, pp. 74–85.

[25] B. Fischl *et al.*, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.

[26] J. Ashburner and K. J. Friston, "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.

[27] D. Cobzas, N. Birkbeck, M. Schmidt, M. Jagersand, and A. Murtha, "3D variational brain tumor segmentation using a high dimensional feature set," in *Proc. IEEE 11th Int. Conf. Comput. Vis. (ICCV)*, Oct. 2007, pp. 1–8.

[28] A. E. Lefohn, J. E. Cates, and R. T. Whitaker, "Interactive, GPU-based level sets for 3D segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Montreal, QC, Canada: Springer, 2003, pp. 564–572.

[29] S. Ho, L. Bullitt, and G. Gerig, "Level-set evolution with region competition: Automatic 3-D segmentation of brain tumors," in *Proc. 16th Int. Conf. Pattern Recognit.*, vol. 1. Aug. 2002, pp. 532–535.

[30] D. Zikic *et al.*, "Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Nice, France: Springer, 2012, pp. 369–376.

[31] E. Geremia, O. Clatz, B. H. Menze, E. Konukoglu, A. Criminisi, and N. Ayache, "Spatial decision forests for ms lesion segmentation in multi-channel magnetic resonance images," *NeuroImage*, vol. 57, no. 2, pp. 378–390, 2011.

[32] S. Bauer, L.-P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Toronto, ON, Canada: Springer, 2011, pp. 354–361.

[33] W. Wu, A. Y. C. Chen, L. Zhao, and J. J. Corso, "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 9, no. 2, pp. 241–253, 2014.

[34] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," *J. Biomed. Opt.*, vol. 19, no. 1, p. 010901, 2014.

[35] T. Vo-Dinh, *Biomedical Photonics Handbook: Biomedical Diagnostics*. Boca Raton, FL, USA: CRC Press, 2014.

[36] S. L. Jacques, "Optical properties of biological tissues: A review," *Phys. Med. Biol.*, vol. 58, no. 11, p. R37, 2013.

[37] V. V. Tuchin and V. Tuchin, *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*, vol. 642. Bellingham, WA, USA: SPIE, 2007.

[38] C. Balas, "Review of biomedical optical imaging—A powerful, non-invasive, non-ionizing technology for improving *in vivo* diagnosis," *Meas. Sci. Technol.*, vol. 20, no. 10, p. 104020, 2009.

[39] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2002.

[40] L. J. P. van der Maaten, E. O. Postma, and H. J. van den Herik, "Dimensionality reduction: A comparative review," *J. Mach. Learn. Res.*, vol. 10, nos. 1–41, pp. 66–71, Oct. 2009.

[41] M. Balasubramanian and E. L. Schwartz, "The isomap algorithm and topological stability," *Science*, vol. 295, no. 5552, p. 7, 2002.

[42] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[43] D. L. Donoho and C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 10, pp. 5591–5596, 2003.

[44] S. Mika, B. Schólkopf, A. J. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, "Kernel pca and de-noising in feature spaces," in *Proc. NIPS*, vol. 11. 1998, pp. 536–542.

[45] K. Q. Weinberger and L. K. Saul, "An introduction to nonlinear dimensionality reduction by maximum variance unfolding," in *Proc. AAAI*, vol. 6. 2006, pp. 1683–1686.

[46] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Proc. NIPS*, vol. 14. 2002, pp. 585–591.

[47] A. K. Ziemann and D. W. Messinger, "An adaptive locally linear embedding manifold learning approach for hyperspectral target detection," *Proc. SPIE*, vol. 9472, p. 94720O, May 2015.

[48] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Trans. Comput.*, vol. 18, no. 5, pp. 401–409, May 1969.

[49] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[50] D. Ravi *et al.*, "Deep learning for health informatics," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 4–21, Jan. 2017.

[51] C. Poultney *et al.*, "Efficient learning of sparse representations with an energy-based model," in *Proc. NIPS*, 2006, pp. 1137–1144.

[52] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. ICML*, 2008, pp. 1096–1103.

[53] S. Rifai, P. Vincent, X. Muller, X. Glorot, and Y. Bengio, "Contractive auto-encoders: Explicit invariance during feature extraction," in *Proc. ICML*, 2011, pp. 833–840.

[54] D. Gillis *et al.*, "Manifold learning techniques for the analysis of hyperspectral ocean data," *Proc. SPIE*, vol. 5806, pp. 342–351, 2005.

[55] C. M. Bachmann *et al.*, "Bathymetric retrieval from hyperspectral imagery using manifold coordinate representations," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 884–897, Mar. 2009.

[56] C. Olson and T. Doster, "A parametric study of unsupervised anomaly detection performance in maritime imagery using manifold learning techniques," *Proc. SPIE*, vol. 9840, p. 984016, 2016.

[57] J. Cheng *et al.*, "Peripapillary atrophy detection by sparse biologically inspired feature manifold," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2355–2365, Dec. 2012.

[58] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Sparse transfer manifold embedding for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1030–1043, Feb. 2014.

[59] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[60] B. Mokbel, W. Lueks, A. Gisbrecht, and B. Hammer, "Visualizing the quality of dimensionality reduction," *Neurocomputing*, vol. 112, pp. 109–123, Jul. 2013.

[61] W. Lueks, B. Mokbel, M. Biehl, and B. Hammer. (Oct. 2011). "How to evaluate dimensionality reduction?—Improving the co-ranking matrix." [Online]. Available: https://arxiv.org/abs/1110.3917

[62] D. Meng, Y. Leung, and Z. Xu, "Evaluating nonlinear dimensionality reduction based on its local and global quality assessments," *Neurocomputing*, vol. 74, no. 6, pp. 941–948, 2011.

[63] L. Chen and A. Buja, "Local multidimensional scaling for nonlinear dimension reduction, graph drawing, and proximity analysis," *J. Amer. Statist. Assoc.*, vol. 104, no. 485, pp. 209–219, 2009.

[64] J. Venna and S. Kaski, "Local multidimensional scaling," *Neural Netw.*, vol. 19, nos. 6–7, pp. 889–899, 2006.

[65] K. Matković, L. Neumann, A. Neumann, T. Psik, and W. Purgathofer, "Global contrast factor-a new approach to image contrast," *Computational Aesthetics*. Aire-la-Ville, Switzerland: Eurographics Association, 2005, pp. 159–168, 2005.

[66] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, nos. 2579–2605, p. 85, 2008.

[67] A. R. Jamieson, M. L. Giger, K. Drukker, H. Lui, Y. Yuan, and N. Bhooshan, "Exploring nonlinear feature space dimension reduction and data representation in breast CADx with Laplacian eigenmaps and t-SNE," *Med. Phys.*, vol. 37, no. 1, pp. 339–351, 2010.

[68] K. Lekadir *et al.*, "Tissue characterization using dimensionality reduction and fluorescence imaging," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Copenhagen, Denmark: Springer, 2006, pp. 586–593.

[69] J. Shotton, M. Johnson, and R. Cipolla, "Semantic texton forests for image categorization and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.

[70] A. V. Lebedev *et al.*, "Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness," *NeuroImage, Clin.*, vol. 6, pp. 115–125, Aug. 2014.

[71] V. Lempitsky, M. Verhoek, J. A. Noble, and A. Blake, "Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography," in *Functional Imaging and Modeling of the Heart*. Nice, France: Springer, 2009, pp. 447–456.

[72] J. Mitra *et al.*, "Lesion segmentation from multimodal MRI using random forest following ischemic stroke," *NeuroImage*, vol. 98, pp. 324–335, Sep. 2014.

[73] D. Ravì, M. Bober, G. M. Farinella, M. Guarnera, and S. Battiato, "Semantic segmentation of images exploiting DCT based features and random forest," *Pattern Recognit.*, vol. 52, pp. 260–273, Apr. 2016.

[74] Y. W. Teh and S. T. Roweis, "Automatic alignment of local representations," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 841–848.

[75] S. Lafon and A. B. Lee, "Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1393–1403, Sep. 2006.

[76] K. Q. Weinberger, B. Packer, and L. K. Saul, "Nonlinear dimensionality reduction by semidefinite programming and kernel matrix factorization," in *Proc. AISTATS*, 2005, pp. 1–8.

[77] X. He, D. Cai, S. Yan, and H.-J. Zhang, "Neighborhood preserving embedding," in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2. Oct. 2005, pp. 1208–1213.

[78] L. Teng, H. Li, X. Fu, W. Chen, and I.-F. Shen, "Dimension reduction of microarray data based on local tangent space alignment," in *Proc. 4th IEEE Conf. Cognit. Inform. (ICCI)*, Aug. 2005, pp. 154–159.

[79] T. Zhang, J. Yang, D. Zhao, and X. Ge, "Linear local tangent space alignment and application to face recognition," *Neurocomputing*, vol. 70, nos. 7–9, pp. 1547–1553, 2007.

[80] V. D. Silva and J. B. Tenenbaum, "Global versus local methods in nonlinear dimensionality reduction," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 705–712.

[81] F. Sha and L. K. Saul, "Analysis and extension of spectral methods for nonlinear dimensionality reduction," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 784–791.

[82] D. K. Agrafiotis, "Stochastic proximity embedding," *J. Comput. Chem.*, vol. 24, no. 10, pp. 1215–1221, 2003.

[83] S. Roweis, "EM algorithms for PCA and SPCA," in *Proc. Adv. Neural Inf. Process. Syst.*, pp. 626–632, 1998.

[84] C. Spearman, "'General intelligence,' objectively determined and measured," *Amer. J. Psychol.*, vol. 15, no. 2, pp. 201–292, 1904.

[85] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Neural Inf. Process. Syst.*, vol. 16. 2004, p. 153.

[86] K. Q. Weinberger, F. Sha, Q. Zhu, and L. K. Saul, "Graph laplacian regularization for large-scale semidefinite programming," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 1489–1496.

[87] L. van der Maaten. (2013). *Matlab Toolbox for Dimensionality Reduction (v0.8.1b)*. [Online]. Available: https://lvdmaaten.github.io/drtoolbox/