Neuroinformatics (2019) 17:373–389
https://doi.org/10.1007/s12021-018-9405-x

ORIGINAL ARTICLE

An End-to-end System for Automatic Characterization of Iba1 Immunopositive Microglia in Whole Slide Imaging

Alexander D. Kyriazis[1] · Shahriar Noroozizadeh[1] · Amir Refaee[1] · Woongcheol Choi[1] · Lap-Tak Chu[1] · Asma Bashir[2] · Wai Hang Cheng[2] · Rachel Zhao[2] · Dhananjay R. Namjoshi[2] · Septimiu E. Salcudean[3] · Cheryl L. Wellington[2] · Guy Nir[4]

Published online: 8 November 2018

Abstract
Traumatic brain injury (TBI) is one of the leading causes of death and disability worldwide. Detailed studies of the microglial response after TBI require high throughput quantification of changes in microglial count and morphology in histological sections throughout the brain. In this paper, we present a fully automated end-to-end system that is capable of assessing microglial activation in white matter regions on whole slide images of Iba1 stained sections. Our approach involves the division of the full brain slides into smaller image patches that are subsequently automatically classified into white and grey matter sections. On the patches classified as white matter, we jointly apply functional minimization methods and deep learning classification to identify Iba1-immunopositive microglia. Detected cells are then automatically traced to preserve their complex branching structure after which fractal analysis is applied to determine the activation states of the cells. The resulting system detects white matter regions with 84% accuracy, detects microglia with a performance level of 0.70 (F1 score, the harmonic mean of precision and sensitivity) and performs binary microglia morphology classification with a 70% accuracy. This automated pipeline performs these analyses at a 20-fold increase in speed when compared to a human pathologist. Moreover, we have demonstrated robustness to variations in stain intensity common for Iba1 immunostaining. A preliminary analysis was conducted that indicated that this pipeline can identify differences in microglia response due to TBI. An automated solution to microglia cell analysis can greatly increase standardized analysis of brain slides, allowing pathologists and neuroscientists to focus on characterizing the associated underlying diseases and injuries.

Keywords Computer-aided detection and diagnosis · Whole slide imaging · Digital pathology · Image analysis · Classification · Traumatic brain injury

Introduction

Background and Motivation

Traumatic brain injury (TBI) is one of the leading causes of death and disabilities worldwide (Popescu et al. 2015). In the United States alone, the incidence of TBI in 2013 was approximately 2.8 million (Taylor et al. 2017), and the estimated annual economic cost associated with TBI in 2010 was approximately $60.4 billion (Finkelstein et al. 2006). Approximately 80% of TBI are mild when classified as mild, moderate, or severe based on the Glasgow Coma Scale (GCS), which was developed to assesses a patient's level of consciousness. Despite its wide historical adoption to grade TBI severity, the GCS has many well-documented limitations (Teasdale et al. 2014; Lieh-Lai et al. 1992). For example, facial injuries and sedation, which are common in neurotrauma patients, can compromise eye and verbal responses to GCS questions. Further, the GCS lacks sensitivity in the mild end of the TBI severity spectrum, to distinguish among patients who do or do not have a potentially life-threatening bleeding within the brain. Finally, as GCS scores can change as patients either improve or decline, there is inconsistent application of when the GCS is administered, for example, by first responders,

Alexander D. Kyriazis, Shahriar Noroozizadeh, Amir Refaee, Woongcheol Choi and Lap-Tak Chu contributed equally to this work.

✉ Guy Nir
guynir@ece.ubc.ca

Extended author information available on the last page of the article.

upon admission, or using the maximum GCS to define TBI severity. Recent clinical studies on contact sport athletes and military personnel have shown that repetitive exposure to mild TBI (which includes concussion) may have long-term consequences and may lead to a neurodegenerative condition known as chronic traumatic encephalopathy (McKee et al. 2009; McKee et al. 2013; Stein et al. 2014). These findings highlight TBI as an important public health concern, and underscore a need to improve understanding of its disease mechanisms using both clinical and animal model approaches.

Neuroinflammation is one of the key neuropathological features found in the post-mortem TBI brain. It has been consistently observed in clinical TBI ranging from severe to mild cases, and across multiple injury mechanisms ranging from cerebral contusion to concussion (Holmin et al. 1998; Li et al. 2013; Coughlin et al. 2017). More importantly, neuroinflammatory changes may persist in TBI patients up to ten years after their initial injury (Ramlackhansingh et al. 2011; Johnson et al. 2013). Chronic neuroinflammatory changes are associated with white matter degeneration (Johnson et al. 2013) and the development of tauopathy in chronic traumatic encephalopathy (CTE) cases (Cherry et al. 2016). Therefore, neuroinflammation plays a pivotal role in the pathogenesis of TBI at both acute and chronic recovery phases.

Microglia are the resident immune cells of the central nervous system (CNS) that mediate neuroinflammation (Kettenmann et al. 2011). Microglia are highly dynamic cells with many important roles in the brain. Under physiological conditions, "resting" microglia assume a ramified morphology with a small soma and fine elongated cellular processes, and actively survey the environment to maintain CNS homeostasis. Upon stimulation by cues (e.g. injury or infection), microglia become activated and undergo rapid and profound changes in morphology including enlargement of the soma and reduction of the complexity of cellular processes. Functionally, "activated" microglia have upregulated motility, proliferation, and phagocytic activities, which are important for clearance of cell debris and microbes, and also during brain development (e.g. synaptic pruning). The gene expression profile of microglial cell surface receptors and signaling molecules (e.g. cytokines and chemokines) also undergo substantial changes during microglial activation.

To enable detailed studies of the microglial response after TBI, we recently developed a mouse model of TBI called Closed Head Impact Model of Engineered Rotational Acceleration (CHIMERA) (Namjoshi et al. 2014; Namjoshi et al. 2017; Cheng et al. 2018) and have shown that mild CHIMERA TBI induces rapid changes to microglial morphology and microglial number (microgliosis), in multiple white matter regions of the murine brain. We have also recently demonstrated that the extent of post-injury microgliosis response is directly correlated to the level of mechanical input by the CHIMERA system (Namjoshi et al. 2017). CHIMERA generates diffuse traumatic axonal injury consistent with human mild TBI and its ability to produce TBI in a procedure that requires only ten minutes per animal greatly improves the ability to generate well-powered injury cohorts. However, one challenge is how to increase throughput in quantifying changes in microglial count and morphology in stained histological sections throughout the brain. Although we have employed objective quantification methods through the use of published analysis algorithms, these methods still require manual selection of regions of interest (ROIs) for cell count quantification and manual selection of representative microglial cells to quantify cell morphology. These image analysis constraints limit the number of brain sections that can be processed rapidly and accurately, thereby posing time and human error caveats toward unbiased analysis of the entire brain.

Without an effective way to automate brain section processing, our laboratories, and others, could not feasibly accommodate very large histopathological demands: indeed, each mouse brain produces a minimum of five slides or 1GB of image data per mouse, and each study includes an average of 200 mice. Furthermore, automatic classification has the huge advantage of being reproducible and auditable especially for longitudinal studies. Methodological updates can be easily applied to the entire database with little additional overhead.

In this paper, we propose a fully automated end-to-end system to take an Iba1-stained brain slide consisting of several coronal cross sections, segment the white matter regions, and identify all microglial cells in these white matter regions to yield accurate analyses of both cell count and morphology. The proposed system uses a combination of signal processing and image analysis tools, including machine learning/deep learning methods, to achieve high counting and classification accuracies. The challenges we address include managing the large variation in specimen morphology, staining intensities, and the subjectivity associated with the data. The specific methodological contributions associated with this new system are described in "Contributions".

Previous Work

Prior to the advent of computer-aided diagnosis, large scale data analysis in the context of pathology and diagnosis would require manual compiling and review of large amounts of data. In recent years, computational advances have allowed for powerful automated approaches to analyze biomedical data (Balis et al. 2011; Doi 2007). Whole slide imaging (WSI) in particular involves capturing entire

specimens using a glass slide scanner, and appears to be the most promising modality for most digital pathology applications (Kaplan and Rao 2015). Scrutinizing WSI tissue samples using image analysis and supervised machine learning methods has shown success in many areas including detection and grading of breast and prostate cancers (Petushi et al. 2006; Ghaznavi et al. 2013; Cruz-Roa et al. 2014; Madabhushi and Lee 2016; Rashid et al. 2018; Nir et al. 2018). Furthermore, Chen et al. (2014) validated the feasibility of providing an objective approach to quantify neurological events of brain injury in mice using image analysis based on statistical machine learning methods.

Deep learning is a very active field of research in machine learning that has shown strong promise in general image classification tasks (Krizhevsky et al. 2012). These techniques are becoming extremely abundant in medical image analysis (Litjens et al. 2017). Previous applications of deep learning in digital pathology include detecting nuclei in colon cancer slides (Sirinukunwattana et al. 2016), grading gliomas (Ertosun and Rubin 2015), and neuron segmentation (Cireşan et al. 2012). In general, deep learning frameworks are used in digital pathology for segmentation, detection, or classification tasks and can often report F1 scores, defined as the harmonic mean of precision and sensitivity, above 0.8 in several medical imaging domains (Janowczyk and Madabhushi 2016).

Segmentation of white matter using machine learning has been a prevalent investigation topic for many research groups (Bae et al. 2009; Wu et al. 2012). However, most of the previous work focuses on segmentation for magnetic resonance images. These are low resolution, three dimensional images that have a different contrast mechanism from histopathology images, which are a sparse sampling of the brain specimen by two-dimensional high resolution tissue slides. To the best of our knowledge, recent advances in image analysis have not yet been fully taken advantage of in automatic segmentation of histological images of the brain, even though WSI is a common modality for pathology research (Farahani et al. 2015). In the case of mild TBI, white matter tracts are more susceptible to damage and represent major ROIs for microglial analysis (Namjoshi et al. 2017).

Automatic microglia detection in brain matter has been approached by several research groups. In Valous et al. (2013), the problem was approached using techniques including contrast boosting quaternion operations, non-local total variation denoising, and homomorphic filtering. They report 80-90% detection accuracies; however, they state systematic difficulties including the appearance of large complex artifacts, texture variabilities, and variability in the contrast between cell bodies and background tissue for Iba1 staining. Additionally, the authors of Johnson and Walker (2015) demonstrate that standard thresholding techniques in Iba1 staining may lead to incorrect cell segmentation results between different stain groups if the data is assumed to be standardly stained. Instead, they suggest using information from the Cumulative Threshold Spectra to give more insight into threshold selection. The authors of Kozlowski and Weimer (2012) used peak detection and thresholding methods to identify individual cells in their z-stacked fluorescently labelled microglia dataset. However, as outlined in Ding et al. (2017), methods that rely on thresholding are not suitable for dealing with noisy, inhomogeneous histological images of microglia. In Ding et al. (2017), it is suggested that many of these specific image analysis challenges can be remedied by approximating tissue images with piecewise representations, by minimization of the Mumford-Shah functional. However, it is mentioned that the algorithm's accuracy may still be compromised by artifacts.

There are similarities between the problem of microglia detection and detecting mitosis in breast cancer histology images; mitosis detection was addressed with a deep learning approach in Cireşan et al. (2012) and Cireşan et al. (2013). The approach is to train a deep convolutional neural network (CNN) to predict the probability that a patch centered around an input pixel is an instance of mitosis. Training examples are obtained from a ground truth labelled set of mitosis occurrences. During testing, the neural network predicts the probability associated with every pixel, where the highest probabilities are the predicted instances of mitosis. Considering individual pixel classification, they claim their neural network achieves superhuman performance.

In addition to microglia detection, considerable research has been invested into microglia morphology analysis, as this offers a glimpse into microglial functional state. Multithresholding techniques are found to yield accurate branching structures for individual cells (Kozlowski and Weimer 2012; Abdolhoseini et al. 2016). In Ding et al. (2017), supervised learning techniques were employed using fractal and multifractal features of branching structures yielding 94.27% morphology classification accuracy. Rouchdy and Cohen (2013) uses the fast-marching method and Harris points to extract tree structures from microglia, and uses the identified branching structure to classify morphologies. Both lacunarity of the branching structure and box-counting fractal dimension have been shown to be good indicators of microglia morphology, with both generally increasing as the cell shifts to a more ramified state (Karperien et al. 2013; Karperien and Jelinek 2015).

Contributions

To the best of our knowledge, this is the first work to propose a robust end-to-end automated pipeline to analyze microglia characteristics in automated segmented ROIs in WSI. From the literature review, white matter segmentation has not been

performed on Iba1 stained images, and previous methods for white matter segmentation on other stains were found to be highly sensitive to the varying intensities Iba1-stained brain tissue produces. In this work, we designed features that are robust against the varying intensities.

As for microglia detection, we developed a deep learning filtering step that casts out image artifacts and rejects false positives on a cell-by-cell basis using a CNN framework for an increased specificity. We showed that this deep learning classification outperforms other popular image classifiers, thus improving upon the shortcomings of the methods outlined in Ding et al. (2017) and Kozlowski and Weimer (2012) and allowing a robust cell detection in noisy images.

The rest of the manuscript is organized as follows. In "Methods" we describe the methods and algorithms used to build our pipeline. In "Results" we present the results obtained when comparing to ground truth test sets. In "Discussion and Conclusion" we discuss the significance and limitations of our work, and propose future research directions.

Methods

Dataset

The pathology slides used in this study were Iba1 immunohistochemistry of brain sections collected from C57Bl/6 mice at various acute time-points following a CHIMERA TBI procedure, as described in Namjoshi et al. (2014). All animal studies were carried out in accordance with guidelines stipulated by the Canadian Council on Animal Care (CCAC) and the UBC Animal Care Committee (ACC; protocol #A15-0096).

Four to five 40μm coronal sections per brain, spanning the olfactory bulb to the posterior hippocampus, were prepared using a Leica CM3050 S Research Cryostat (Namjoshi et al. 2014). A subset of our sections was digitally imaged using an in-house Zeiss Axio Scan (Carl Zeiss Microscopy) slide scanner using a 20x magnification objective lens. From the whole-mount images, 20x magnified images containing ROIs were extracted using Zen Pro Image Viewer v. 10.0.15063, Zen 2.3, Blue Edition software (Zeiss, Oberkochen, Germany). The remainder of our sections were digitally imaged at a commercial histology service (Wax-it Histology Services Inc., Vancouver, BC) with a ScanScope CS-R scanner (Aperio Technologies) using a 20x magnification objective lens. From the whole-mount images, 20x magnified images containing ROIs were extracted using the Aperio ImageScope Viewer (v. 12.2.2.2015, Aperio Technologies) software.

For this pipeline a total of thirty brain slides were used. In order to properly sample the slides, we imposed the condition that we wanted an even distribution of three different time-after-injury classes. Furthermore, we imposed that we wanted approximately one third of each class to be negative control images (shams). With these slots in place, we randomly shuffled and sampled the available brains, in addition to randomly sampling the particular section of each brain used. These were sampled from a large collection of over 400 brains from three completed studies for which data had already been generated (Namjoshi et al. 2014; Namjoshi et al. 2017; Cheng et al. 2018). The collection included slides prepared by different investigators in separate batches of staining. A final imposed condition was to only use slides that we considered 'darker stainings'. This was considered acceptable because it is influenced by duration that the tissue is kept in the Diaminobenzidine (DAB) solution, which is an entirely controllable characteristic.

The resulting dataset consisted of slides that each belonged to a different brain, and varied by time-after-injury, impact energy applied, stain set, neuroscientist who prepared them and whether they belonged to the sham group or the experimental group. In particular, the dataset used comprised of nine slides sampled six hours after injury, twelve slides sampled two days after injury and nine slides sampled fourteen days after injury. Additionally, the thirty slides consisted of eleven sham images and nineteen experimental images. These variations ensure we can get true insight into how the method copes with typical variability that a pathologist may encounter.

The neuroscientists involved in the project were Dr. Cheryl L. Wellington (25 years of experience in neuropathology), Dr. Dhananjay R. Namjoshi (9 years), Dr. Wai Hang Cheng (6 years), Asma Bashir (3 years), and Rachel Zhao (2 years).

For white matter segmentation, the white matter sections of these thirty slides, including regions such as the optic tract, and the corpus callosum, were manually marked by a single experienced neuroscientist using ImageScope by drawing boundaries around these regions.

For microglia detection, the same thirty slides were analyzed using our automated white matter segmentation process. A collection of 300 white matter patches were obtained accordingly, with ten patches sampled from each slide. These patches were all independently analyzed by two neuroscientists to determine microglia cell positions. The two neuroscientists were senior PhD students with three and six years of experience. Using a point-and-click interface, the neuroscientists annotated each image by selecting the position of microglia cells. This is used as the basis of our microglia detection accuracy measurement.

For microglia morphology characterization, 300 cells were randomly sampled from the aforementioned white matter patches. For each cell, a 40×40 pixel image centered

about each cell centroid was sampled, and sent to a trained neuroscientist. These images were classified by trained neuroscientists as being 'Ramified' or 'Amoeboid' in morphology. This is used as the ground truth for the morphology detection component of the pipeline.

The proposed automatic processing pipeline is comprised of three distinct tasks. Firstly, the white matter regions are segmented from the rest of the brain slide. Secondly, microglial cells are detected in these white matter patches. Finally, cell morphology analysis is performed on a cell-by-cell basis on the detected microglia cells. The proposed processing and classification workflow is illustrated in Fig. 1.

White Matter Segmentation

White matter segmentation of the images involves dividing the high-resolution pathology slide to a grid of smaller image patches of size 256×256 pixels. Some of the slides exhibited air bubble artifacts due to the preparation methods. These artifacts were always found to be very close to a perfect shade of grey, exhibiting R,G, and B intensities that are within five units of each other on a scale of 0–255. Preprocessing was performed on every image patch to invalidate patches that had this discriminating greyness characteristic, effectively excluding artifacts from training and evaluation. It was found that white and grey matter tissue patches never have this characteristic, so they were always preserved. In the case that the air bubbles overlap with tissue (which was not encountered), tissue will still not be falsely discarded since the air bubbles are translucent. Regardless, the number of such artifact occurrences were low and the preprocessing was implemented as a safe-guard for possible use of the algorithm on unseen test data. As a limitation, our work currently addresses only air bubbles as artifacts. However, the processing could be extended to account for other artifacts such as the rare occurrence of fiber strings, by using similar image processing techniques.

The ROIs are defined to be areas of white matter, as these areas are more susceptible to damage than grey matter in mild CHIMERA TBI under these experimental conditions (Namjoshi et al. 2017). The problem was approached from a machine learning perspective, where each valid 256×256 pixel image patch is considered as a sample. For the training stage, expert labelled slides were used. Based on manual segmentation, each patch within an annotated contour was assigned a white matter label, whereas the rest of the valid patches were assigned with a grey matter label. An example of a segmented slide and its corresponding processed label map is illustrated in Fig. 2.

A set of features was then extracted from the image patches. Some features, e.g., color and edge statistics, were designed to follow the visual cues suggested by the expert annotators, whereas other features were extracted using high-level image-based techniques, as described below. The features were assessed and selected based on the classification performance measured during cross-validation. Table 1 summarizes the set of features that were found to minimize classification error and that were eventually used in our experiments.

In particular, we found features that were adapted from an aerial image segmentation method (Li et al. 2000) to be useful, and the approach was adapted to account for the different image patch size used for our system. The discrete cosine transform (DCT) coefficients were calculated for patches of size 256×256 pixels, instead of 4×4 pixels. Every

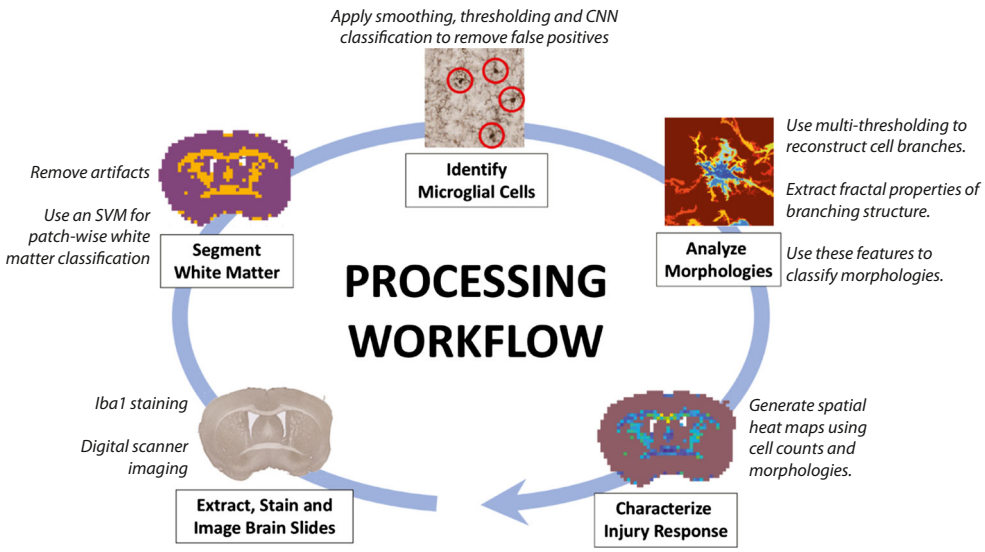
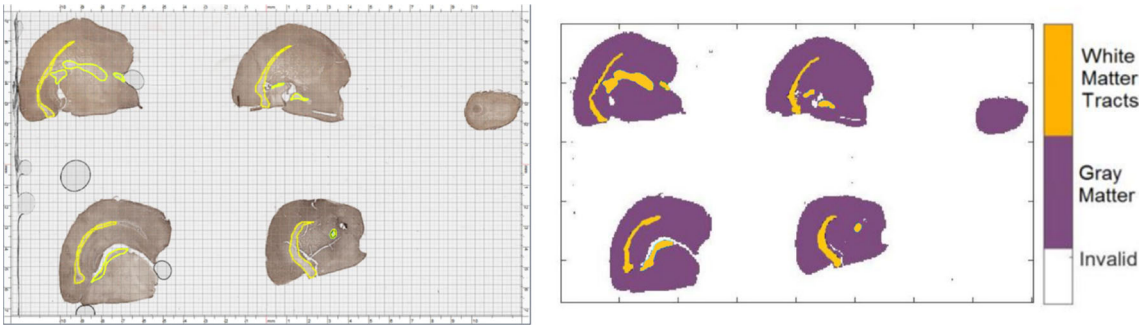


Fig. 1 Processing and classification workflow



(a) Labeled Iba1 slide (b) Processed label map

Fig. 2 A sample collection of coronal section. **a** The division into patches (black grids) and the manually marked ROIs (yellow contours) are overlaid on the image. **b** Artifacts removed. Such maps are used in training and evaluation of the classifier

patch is divided into 16 sub-patches of size 64×64 pixels each. By calculating the mean of the DCT coefficients of each sub-patch region, 16 mean DCT coefficients of $D_{0,0}$, $D_{0,1}$, ..., $D_{3,2}$, $D_{3,3}$ were obtained. We then followed (Li et al. 2000) for acquiring six features based on linear combinations of the mean sub-patch DCT coefficients.

For every slide, the feature values obtained from each patch were normalized with respect to the distribution of its corresponding feature value in the set of all patches combined from the entire set of training slides. The normalized features of each patch and the corresponding label of the patch were subsequently fed into our machine learning model, resulting in every patch of our overall brain slide being classified. We have found Support Vector Machines (SVM) (Bishop 2006) to provide a good trade-off in terms of classification performance and speed.

The SVM classifier had the highest performance on the training data based on metrics defined in "Results". The classifier was trained on ten Iba1 stained slides each containing four to six coronal slices of the brain that had the ROIs marked by a trained neuroscientist experienced in Iba1 histology. The hyperparameters of the model were optimized using a leave-one-out cross-validation. In general, a range of 69-365 white matter patches were found (mean and standard deviation of 183 ± 71 white matter patches), but this varies depending on the position of the cut along the brain.

Additionally, as the approach used for segmentation is based on small image patches and does not use information about the overall geometry of the White Matter tracts, it is slice-plane invariant. In this study, coronal slices were used, however we expect similar results in the sagittal and transverse planes.

Table 1 List of features for the ROI detection step of the pipeline (Gonzalez and Woods 2008; Van Gool et al. 2014; Li et al. 2000)

Feature	Details
1–3	Mean values of RGB intensities.
4–6	Standard Deviation values of RGB intensities.
7–9	Standard deviation of 10-bin histogram for each color channel.
10–11	Gradient magnitude and direction of the greyscale image.
12–17	Linear combinations of mean DCT coefficients of sub-patches.

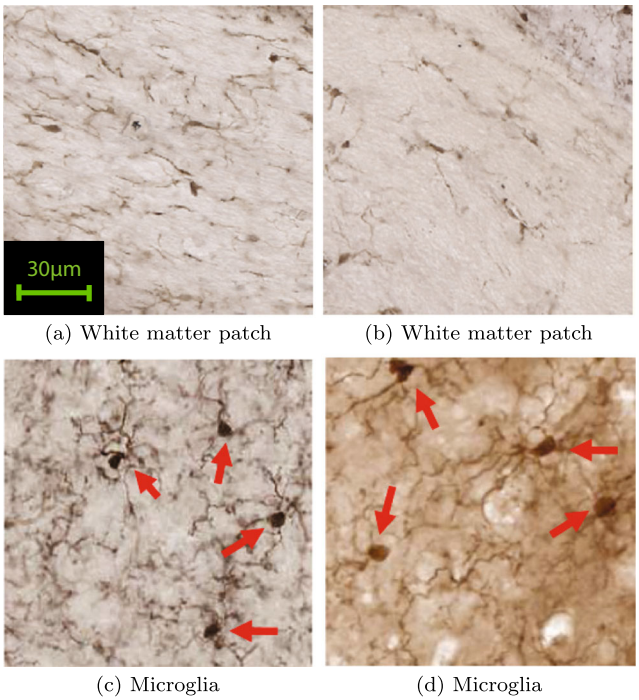
Microglia Detection

The white matter image patches extracted from the previous step are passed to the next stage of the pipeline. Since intensities may vary significantly between slides due to differences in slide preparation and staining, as demonstrated in Fig. 3, they are converted to greyscale and normalized to have the same mean intensity. Patches that are found to be more than 15% white pixels (e.g., patches around tissue edges) are discarded due to abnormalities.

Imaging artifacts and noise may lead to misdetection of cells when using standard thresholding. To solve this problem, we employed the Mumford-Shah total variation smoothing as described in Ding et al. (2017). This smoothing acts as an effective noise removal tool that has edge-preserving properties. This characteristic is critical for retaining cell masses and morphologies that characterize microglia.

The Mumford-Shah total variation smoothing algorithm (Strekalovskiy and Cremers 2014) is applied to each normalized patch to yield a piecewise representation of the image. Following Ding et al. (2017), this entails the minimization of the Mumford-Shah functional, as outlined in Strekalovskiy and Cremers (2014),

$$\min_{u,K}\left\{\int_{\Omega}|u-f|^2\,dx+\alpha\int_{\Omega\setminus K}|\nabla u|^2\,dx+\lambda|K|\right\}, \qquad (1)$$



(a) White matter patch (b) White matter patch

(c) Microglia (d) Microglia

Fig. 3 (3a) and (3b) are representative white matter patches that were successfully identified in the white matter segmentation part of the process, and will subsequently undergo microglia detection. (3c) and (3d) are examples of microglia appearing in the tissue. The red arrows indicate the positions of the microglial cell bodies

where f represents the original image intensity, u is the piecewise representation of f, and K represents the set of boundaries of the piecewise function u. The parameters α and λ penalize the (lack of) smoothness and discontinuities imposed on the resulting piecewise image, respectively.

In Strekalovskiy and Cremers (2014), an efficient and scalable algorithm was proposed for approximately minimizing this functional using primal-dual methods. Using an implementation of this algorithm made publicly accessible on github (Strekalovskiy and Cremers 2014), we performed effective denoising of our tissue images.

We opted for maximum smoothness ($\alpha = \infty$) resulting in constant intensities in the piecewise regions. This is to ensure that cell bodies are homogeneous in intensity, so they are more robust to subsequent peak detection steps. The λ parameter was left as a variable, and tuned in our optimization stage.

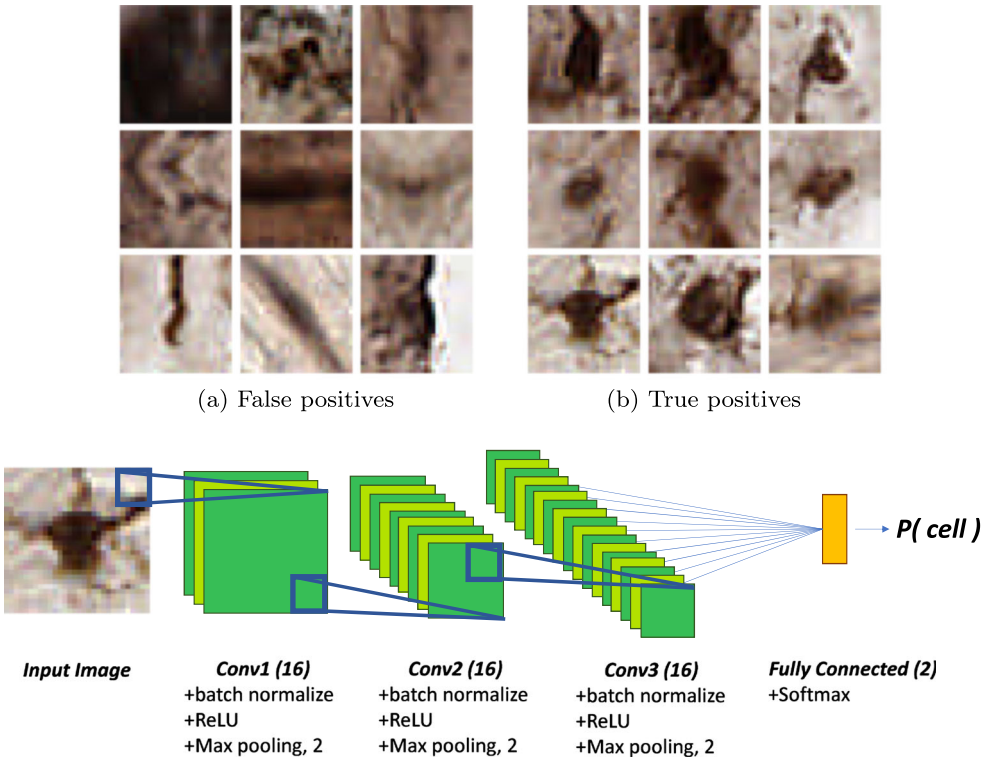
Following smoothing, images are binarized via a peak detection algorithm, which selects components of constant intensity whose boundary comprises pixels of lower intensity (The MathWorks 2017b). This is found to work well in situations where conglomerate cell masses would otherwise be interpreted as a single cell. Instead, constituent cell peaks are individually detected. All connected components of the resulting binary image are easily detected. Components with an area less than a set threshold are then discarded. This size threshold is left as a parameter and tuned in our optimization step.

The masses of the residual pixels marked as true detections following the previous steps are considered as detected cells. Detected cells with a total area above 5000 pixels (1230 μm^2) are scrutinized with an additional application of the Mumford-Shah smoothing, to increase detection sensitivity as these clumps may still consist of multiple components. This particular size threshold was determined empirically based on typical size variation of grouped cells. Figure 4 illustrates the full microglia detection pipeline until this point.

Naturally, there will be false positives detected along with the true positives. Dark, non-cellular objects are often found. We proceed through a second layer of screening using supervised learning techniques, as follows. A 30×30 pixel image is cropped around the center of mass of each positive detection, and passed as an input to a classification model. Cells are classified into a false positive category and a true positive category, in order to weed out false detections. Sample images from each class are illustrated in Fig. 5. The classification model is trained on an image data set of false positive and true positive cells that are labelled by considering the intersection and difference of

(a) Original image (b) Normalized (c) Smoothed (d) Peak detection (e) Detected cells

Fig. 4 The various stages of the proposed microglia detection algorithm. (4a) Original image, (4b) Normalized image after preprocessing, (4c) Smoothed image after the Mumford-Shah functional is minimized (4d) Peak detection applied (4e) Detected cells after size filtering and deep learning filtering



(a) False positives (b) True positives

(c) Our convolutional neural network architecture

Fig. 5 The components of the deep filtering stage. Many false positives (5a) are detected alongside the true positives (5b). Machine learning classification models are trained to discard false positives. A convolutional network architecture used in the classification of detected cells as shown in (5c). Raw RGB input images are passed through three convolutional layers with batch normalization, ReLU and max pooling. The probability of the image being a valid cell is outputted

cell detections between a naive application of our proposed method, and the ground truth detection set. To make classification more robust, the training set is augmented with rotated versions of each cell image (90°, 180° and 270°).

The model is used to predict the probability that a positive detection is a false positive or a true positive. A decision threshold that can bias the classifier is left as a parameter. Cells predicted as false positives at this stage are then discarded and do not appear in the final detection set.

Several classification models were examined, namely, logistic regression (Hosmer et al. 2013), support vector machines (Steinwart and Christmann 2008), RDF (Breiman 2001), Adaboost (Freund et al. 1999), and the CNN described below. A comparison between the different models is provided in "Results". Among the tested models, the deep learning approach using a custom CNN leads to the best results. CNNs automatically identify their own discriminating features of cells by looking at the differences between large numbers of pre-labelled 'cell' and 'non-cell' objects. The features could be related to textures, patterns, intensities, shapes, sizes, etc. These features are generally more abstract, and not easily interpretable by humans as CNNs operate in a black-box fashion. The particular architecture we used for the network was motivated by an implementation of a CNN used to classify handwritten digits from the MNIST dataset using batch normalization, ReLU activation, and max pooling layers (The MathWorks 2017a). The architecture was optimized using cross-validation techniques on our labelled set. A visualization of the architecture used can be seen in Fig. 5c.

The entire microglia detection process has three parameters that need to be adjusted: the cell size threshold, a Mumford-Shah functional minimizer λ parameter, and a classifier decision threshold. These parameters were optimized using a gradient descent method to minimize cell detection error. With a parameterized segmentation, cell locations in each patch can be identified, and cell counts can be calculated on a per patch basis. All cells detected at this stage undergo microglia morphology analysis which is outlined in the following section.

Microglia Tracing and Analysis

In order to properly analyze morphology, it is necessary to properly segment the branching structure of the cells. Due to intensity variations in the background and the pixels belonging to the cells in our dataset, methods outlined in Kozlowski and Weimer (2012) and Ding et al. (2017) were not sufficiently effective at capturing the complex branch-

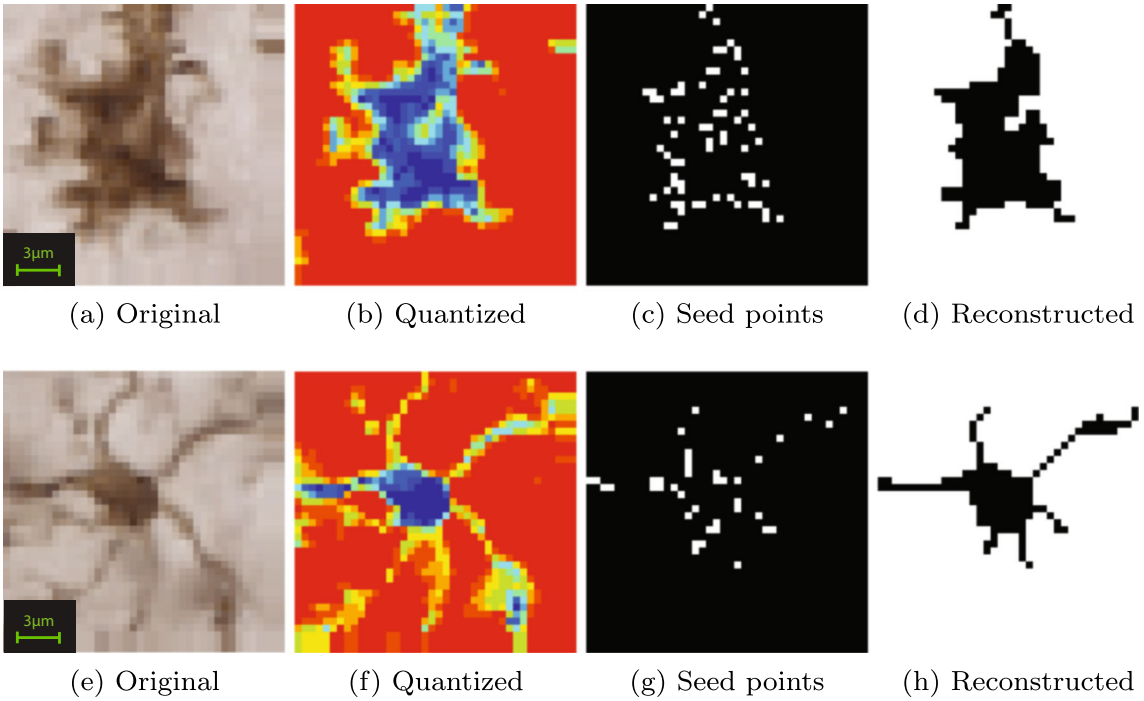


Fig. 6 Overview of automatic tracing of microglia. (6a - 6d) shows the analysis performed on a relatively amoeboid specimen. (6e - 6h) shows the analysis performed on a relatively ramified specimen. (**a/e**) Original image. (**b/f**) Quantized image using Otsu's multi-thresholding. (**c/g**) Sampled seed points. (**d/h**) Reconstructed microglia using minimum spanning tree to connect the seed points

ing structures. Thus, for automatic tracing of microglia we adopted the approach outlined in Abdolhoseini et al. (2016). The algorithm involves quantizing the raw image (Fig. 6a) into discrete levels using Otsu's multi-thresholding (Fig. 6b) and sampling seed points (Fig. 6c) from each level with successively decreasing rates. The cells are then reconstructed from the resulting seed points using a minimum spanning tree as shown in Fig. 6d. The proposed routine is fast and robust to intensity variation throughout the image. Another example of this process is demonstrated for a ramified cell in Fig. 6e-g.

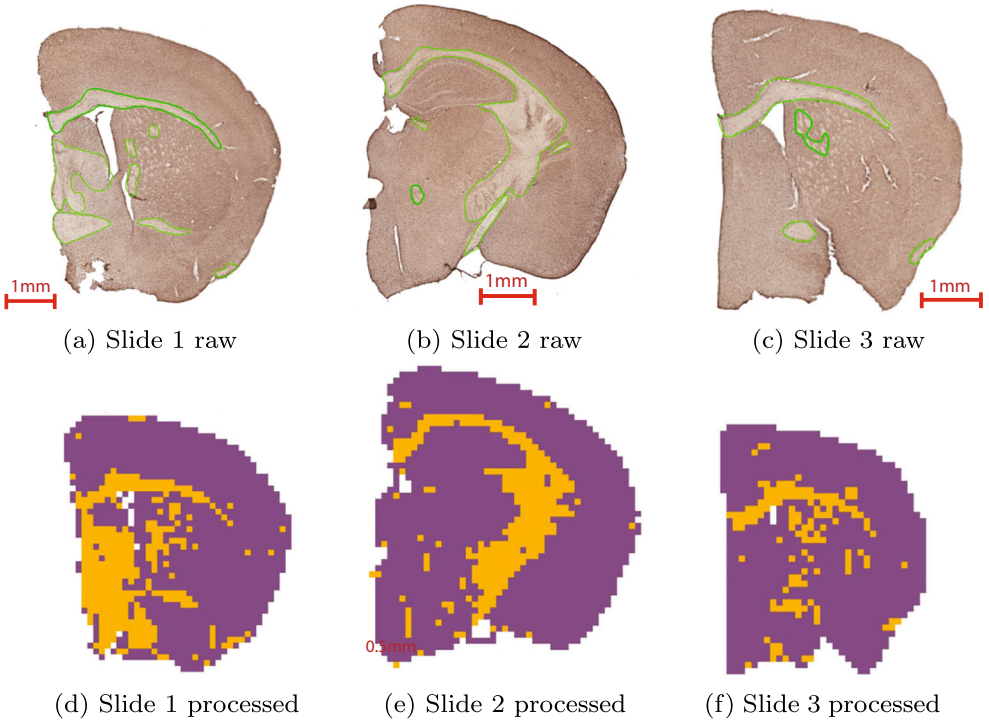
To quantitatively assess the shift in morphology, the resulting binary image of the cell is processed further to determine the box-counting fractal dimension and lacunarity, which are both shown to be good indicators of microglia morphology (Karperien et al. 2013; Karperien and Jelinek 2015). These quantities are calculated using the libraries (Costa 2013) and (Vadakkan 2009). Through our testing, we additionally found the second moment of area of the reconstructed skeleton image with respect to the centroid, cell circularity (ratio of perimeter and area) and cell area to be effective discriminating features as well. The reconstructed skeleton image is obtained using the library (Kollmannsberger 2009). These features can be used as the inputs to a trained machine learning based classification system to characterize morphology. In particular, we used an SVM classifier to determine whether cells fall closer on the spectrum to being contained amoeboid bodies, or branched ramified structures.

The entire routine was applied to each cell in the detection set, using parallel programming to reduce the computational time of the algorithm.

Results

White Matter Segmentation

To evaluate the performance of the white matter segmentation, this step of the pipeline was tested on the thirty slides described in "Dataset". These slides were previously unseen by the algorithm. Performance was assessed by generating confusion matrices from the comparison between the classifier's output to the ground-truth labelling. Figure 7 shows examples of the classification output in comparison to the labelling by the neuroscientist. Quantitative results for each of the slides are represented in Fig. 8. The mean and standard deviation of sensitivity, specificity, and accuracy over the thirty slides were 76 ± 18%, 84 ± 18%, and 84 ± 14%, respectively. Standard deviation values in these results are disproportionately worsened by five particular outlier slices- mainly due to differences in their staining intensity.



(a) Slide 1 raw (b) Slide 2 raw (c) Slide 3 raw

(d) Slide 1 processed (e) Slide 2 processed (f) Slide 3 processed

Fig. 7 Example of different outputs compared to the true labeling. First row includes the raw pathology slides. Second row is the white matter classification output

Indeed, if we exclude these five slices, the sensitivity, specificity, and accuracy are 75 ± 12%, 90 ± 6%, and 89 ± 4%, respectively.

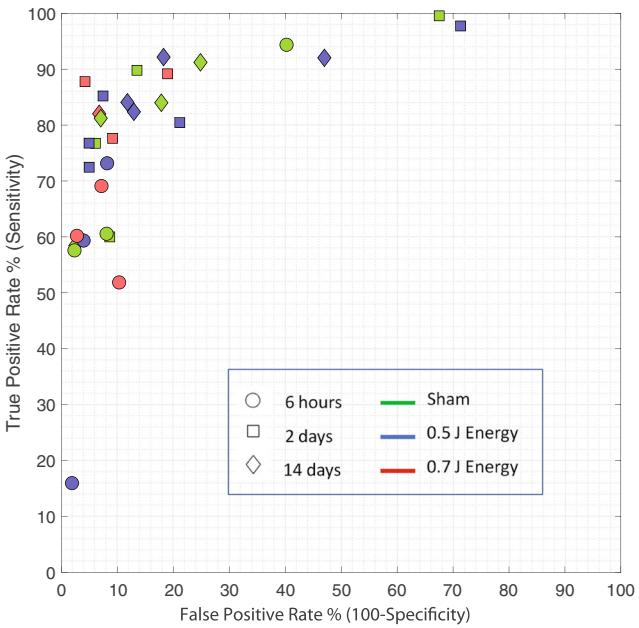


Fig. 8 White matter segmentation results. Each point represents one of the thirty slides tested. They are distinguishable by time-after-injury class (shape), and impact energy (colour)

Microglia Detection

Microglial cell detection was evaluated by comparing the algorithm output to a labelled, ground truth set. A collection of 300 white matter patches were evenly obtained from the thirty brain slides, with each brain slide having a slightly different stain intensity. These patches were sent to the lab to be analyzed by neuroscientists. Using a point-and-click interface, two neuroscientists independently annotated the images by selecting the position of microglia cells. The 300 patches were evaluated and used as the basis of our accuracy measurement.

With regards to the final deep learning filtering stage, three different methods of feature extraction are compared:

(i) manually calculating various image features,
(ii) features extracted from a pre-trained neural network, in our case AlexNet as discussed in Janowczyk and Madabhushi (2016),
(iii) a custom CNN designed and trained based on our data.

For (i) and (ii), features are used to train SVM, Logistic Regression, Random Forest, and Adaboost models. Hyperparameters for each model are optimized using cross-validation. It is found for all models that the features extracted using AlexNet managed to classify with a higher

accuracy than using the manually calculated image features such as the mean, standard deviation, DCT coefficients, circularity and second moment of the cell. We therefore focus on the comparison of feature extraction using a pre-trained neural network (ii) vs. a custom trained network (iii).

For (iii), our network architecture is optimized by cross-validation and consists of three convolutional layers, each followed by a batch normalization layer, ReLU activation layer and a max pooling layer. Highest accuracies were found using only the blue colour channel, rather than using all three channels.

As Fig. 9b shows, a custom trained convolutional neural network has the best performance in the high sensitivity range where we intend to operate. It is used as a final stage to our microglia detection algorithm, and any positives predicted as false positives are discarded from the final cell set.

Considering the whole process starting from the Mumford-Shah smoothing, we analyzed the precision and sensitivity of our cell identification with respect to the labelled sets. The parameter varied to generate the plot is the decision threshold used by our neural network. Since this parameter only influences cells that have proceeded to the deep filtering stage, it is not possible to obtain 100% sensitivity by setting the decision threshold to an extreme value (although 100% precision is possible.)

Our 300 labelled microglia-detection patch dataset was immediately divided into a 200 image training set, and a 100 image test set, with very few tests performed on the test set until the results are reported, so to not induce optimization bias. The test set data furthermore comes from completely different brain specimens, so there is no possibility of overfitting to the same staining intensities, as they can vary considerably between samples.

The ground truth is considered to be the union of the positive detections between the two neuroscientists that independently labelled the same data set, since there is considerable variability between their markings. This is a testament to the fact that microglia detection in our dataset can be a very challenging problem, due to blurring of features outside of the focal plane, in addition to human subjectivity. The union was chosen since it is likely that each neuroscientist's labelled set would suffer from false negative detections rather than false positives (based on our experience, it is more common for labellers to skip a cell than to mark one wrongly.)

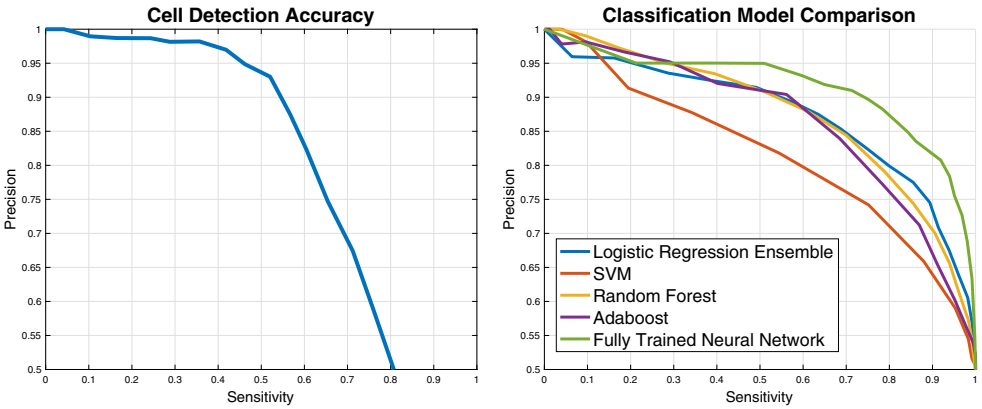
The training set was used to optimize segmentation parameters in addition to training the neural network, both optimized against cross-validation accuracy. After optimizing our procedure on validation sets, we carried out a pass through the test set to report the precision-sensitivity curve in Fig. 9a.

The results show that we can target approximately 69% precision and 70% sensitivity, which corresponds to an F1 score of 0.70. When comparing the ground truth sets that the neuroscientists independently labelled, between each neuroscientist, we find 44% precision and 96% sensitivity (or conversely switching the two values depending on which labeler perspective is being considered). These particular values indicate that one labeller was more much more selective than the other, and one detection set was roughly a subset of the other. Furthermore, this can be interpreted as an inter-observer variability F1 score of 0.60. We expect a third labeller to have similar inter-observer accuracies because of the subjective nature of classifying the ambiguous looking bodies. This comparison shows that our system can autonomously achieve a level of performance that would not be inferior to another human labeller.

Morphology Classification

As described in "Dataset", 300 cells are scrutinized by trained neuroscientists and classified as being 'Ramified' or 'Amoeboid' in morphology.

From each of these cells, we performed the microgila multi-threshold tracing technique described in "Methods".



(a) Cell detection accuracy (b) Classification model comparison

Fig. 9 Microglia detection results. (9a) A precision-sensitivity plot of microglia cell detection accuracy. The parameter being varied is the decision threshold of the deep learning classifier. (9b) A comparison of several classification models for the task of classifying a positive cell detection as a true positive or a false positive. Each model is optimized through cross-validation

Using the obtained branching structure, we extracted the fractal dimension, lacunarity, second moment of area about the centroid of the cell body, cell circularity and cell area. Using an SVM classifier, we used these features to predict the morphology class as defined by the labels we obtained from the neuroscientists.

To assess accuracy, we randomly sampled a subset of cells that were evenly distributed by morphology class. As the distribution of the classes in the original 300 cells was not uniform, and some bodies were too obscure to have a reasonable morphology label, 204 were sampled to achieve an even distribution. A binary classification test was performed (either 'Ramified' or 'Amoeboid') and then a leave-one-out cross validation was used to assess performance. It was found that the classifier could perform with 70% accuracy on both classes.

Injured vs. Sham Study

A preliminary analysis comparing three TBI affected brains with three Sham brains was conducted using the entire described pipeline. The microglial densities in the white matter tracts can be found in Table 2, where the mean and standard deviation of cell density is reported across every white matter 256×256 pixel ($\approx 16000 \mu m^2$) patch extracted from that particular brain. In this analysis, each brain yielded between 61-180 white matter patches, with mean and standard deviation of (93±44). The average density increase between the sham and injured specimens was 35%.

Discussion and Conclusion

In this paper we present a novel and robust end-to-end pipeline that serves as a tool for pathologists and neuroscientists to perform automated microglia morphology analysis from WSI. The pipeline comprises of an automatic segmentation of white matter in WSI, and a subsequent automatic detection and morphology analysis of microglia cells in these identified ROIs. Using the parameters such as microglia prevalence and microglia morphology obtained through this process, we can quantitatively assess the microglial activity throughout white matter regions in the brain.

A representative output of the pipeline is shown in Fig. 10 where microglia cell count density is illustrated in the white matter regions of a slide using a heat-map. The density generally varies from 0 cells to up to 2000 cells per mm^2 in the high density regions. We found our pipeline to be robust to variations in image intensities, based on slides randomly sampled from our lab that represent diversity in Iba1 staining. Furthermore, we can operate with acceptable accuracies on slides from brain specimens that our algorithm never trained on.

The results obtained from our white matter segmentation component were deemed acceptable as per the feedback obtained from the consulted neuroscientists for two reasons. Firstly, there is inevitable inconsistency of hand-labelled slides, which are used as inputs to train and evaluate the algorithm. Reaching accuracies closer to 100% on a patch-by-patch basis would indicate we have overfitted to the hand-labelled set, which imply a classification model that is not easily generalizable. As future steps, this issue can be addressed by exploring unsupervised learning methods such as clustering algorithms that do not depend on user input and labelling such as in k-means (Duda et al. 2012). Secondly, for the performance of the overall pipeline, perfect white matter segmentation is not critical. Although white matter tissue is more susceptible than grey matter to CHIMERA-induced injury under the experimental conditions used in these studies, and provides more insight

Table 2 A comparison of the mean and standard deviation of microglial density between Injured and Sham brain specimens

	Microglial density (1/mm^2)			
	6 Hours	2 Days	14 Days	Overall
Sham	460±260	530±240	450±220	480±230
Injured (0.7J)	630±300	960±380	440±190	650±360
Increase	37%	81%	-2%	35%

The last row represents the percent increase in density in injured brains as compared to sham

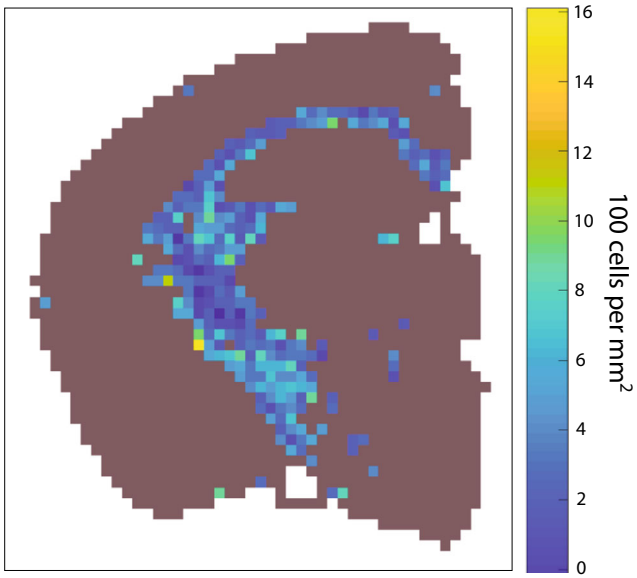


Fig. 10 A heat-map visualization of the resulting microglia densities. The color represents microglia count ranging from dark blue (0 cells per mm^2) to yellow (1600 cells per mm^2). Brown indicates grey matter regions. The labels on the legend indicate microglia count/maximum microglia count in the slide

into the mechanisms of recovery from diffuse axonal injury, some misclassification of the patches is not expected to change the interpreted TBI analysis.

As a further consideration, it is worth investigating the performance metrics that could be achieved by assessing additional classification methods. For instance, a CNN classification framework for white matter detection could learn better discriminating features in our image data such as underlying textures of tissue tracts, tissue homogeneity, etc. Furthermore, using different tissue staining, and validating the data using more pathologists and neuroscientists as a performance standard would be useful to further validate this technology.

With regards to microglia detection, we found that special considerations needed to be made. In particular, it was found that a machine learning approach needed to be employed in addition to traditional image segmentation techniques to achieve reasonable accuracies on our data sets. In this regard, the application of deep learning in the domain of microglia identification is a significant contribution. From our data sets, we found that this yields a boost in cell identification accuracy when applied in addition to other methods from literature. We believe that this final filtering step is critical to robustly applying these algorithms on slides subject to typical noise from iba1 staining. One limitation of our method is that we considered our ground truth to be the union of two labeller's sets. This is prone to introducing false positives to the ground truth set. Although we assume very few false positives exist (we assume labellers rarely select a non-cell), a consequence is that cell counts may be overstated.

Microglia morphology analysis was found to be a challenging problem, with results that have room for improvement. The difficulty largely falls in the fact that microglia are 3D bodies with branches that extend out of the imaging plane, leaving dim structures that sometimes blend into the background. Consequently, branches and therefore overall morphology can be often difficult to classify automatically. In our dataset, the multi-thresholding approach has demonstrated robustness in many cases where conventional thresholding fails. Using convolutional neural networks to detect subtle branch components is a future area we intend to explore which could potentially make this process even more robust.

Using automated means, investigations into injury response mechanisms and how they relate to the nature of injury using the CHIMERA model will be a priority in the future. Our preliminary Injured vs. Sham analysis shows that our pipeline can detect microglia densities that distinguish injured and non-injured brain specimens. The particular trend over the different time points also matches expectations, with the most significant change occurring at the 2 day mark, and with diminishing difference between the injured and sham specimens as the microgliosis resolves, which is evident in the results at the 14 day mark. The average density increase in the injured brains suggests the known correlation between microglial density and injury response, and such a marker should therefore assist in automatic detection of healthy or pathological areas in unseen whole-slide images, provided that image quality and contrast is sufficient. Further studies will be performed to identify the correlation between injury response and morphology activation. A limitation of our current pipeline is the ability to report microglia morphology activation in particular ROIs (eg. brachium of superior colliculus, olfactory nerve layer, optic tract, etc.), since in our experience, morphological changes are often more responsive in certain localized regions than others. This is, however, a clear next step in making this pipeline more powerful.

CHIMERA studies generate several hundred slides across the dimensions of treatment (i.e. sham, TBI, interventions) as well as time after impact, thereby posing a considerable burden for histopathology, in particular, image analysis. On average, approximately 5 hours are required for manual analysis of a single slide, which includes segmentation of five ROI, counting microglia in each ROI, and fractal analysis of a very limited sample of 3-5 microglia. By contrast, the methods described here require approximately 15 minutes to count and analyze microglia for the whole slide, which translates to a 20-fold increase in analytical throughput at comparable accuracy to a trained neuroscientist. Therefore, the integration of the proposed pipeline into our laboratory's research workflow greatly improves the processing throughput of the large amounts of histological data that has been amassed, and allows for more time to be invested into research on factors that modulate TBI recovery. Furthermore, automation provides the advantages of limiting biases to those that are consistent and predictable, and any further changes to the algorithm can easily be re-executed on any laboratory images resulting in repeatable and reproducible results. To our knowledge, a robust end-to-end pipeline covering the entire process of white matter ROI segmentation to microglia detection and analysis has not yet been integrated into a research environment in such a way.

We anticipate that the system's performance will improve as we increase the size of training data. As future work, we will evaluate the system against multiple annotators to study the performance through intra- and inter-observer variability. While the scope of this paper is limited to white matter regions, the system can be easily adapted to analyze grey matter regions as well. Initial qualitative results, as presented in Fig. 11, suggest that the pipeline can perform effectively on some regions of grey matter (this serves as another validation to the robustness of the classifier that was trained on white matter patches). However, in order

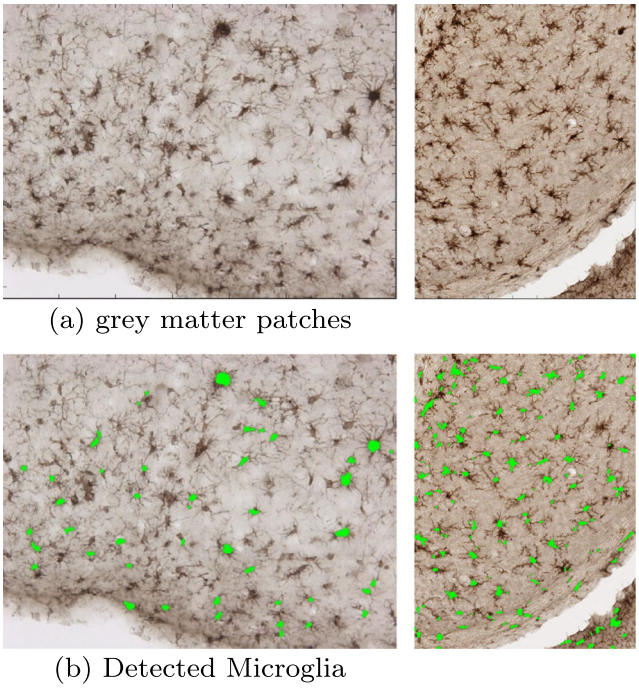


Fig. 11 Examples of applying the developed white matter microglia classifier to grey matter images. Although trained exclusively on white matter patches, the classifier was found to be robust enough to detect a diverse range of cells. However, dedicated training will be required to optimize the performance

to optimize its performance, the CNN should be adapted to such images by re-training on grey matter patches. This process might be expedited by applying transfer learning techniques. Analysis of grey matter has interesting applications across many neurological conditions in addition to TBI, including, for example, characterizing grey matter demyelination in multiple sclerosis. In this case, there are differences in inflammatory responses in grey matter when compared to white matter. Analyzing this with automated methods would be very useful. In addition to analyzing different tissue types, the proposed system can also be adapted to any staining, as long as the training and thresholding parameter choice is adjusted. Moreover, despite being only previously used for mouse models, we wish to validate CHIMERA and the automated pipeline across several species in the future.

There is also a great potential to expand the software to include detection of other biomarkers that are both relevant to mechanisms of injury, and visible in histology. Many new previously-unrealistic biomarkers could be realized with the aid of automated pathology software. Taken even further, this study focused exclusively on colorimetric immunohistochemistry to visualize microglial cells; however, the system could be easily adapted for fluorescent stained images, in analysis of other neural cells for other neurological or neurodegenerative diseases, and in regional specificity for immunohistological biomarkers of interest. Additionally, while the current system is based on an injury model that produces diffuse rather than focal injury, future development of the software could include a version that quantifies microglial morphology as a function of distance to the lesion epicenter. This would be useful for TBI models that generate focal injury, such as controlled cortical impact, and fluid-percussion, as well as stroke models.

An application that this system can be readily adapted to is the automated analysis of post-mortem human brain tissue. With enough training data, sophisticated software can be developed to include the detection and quantification of numerous biomarkers of interest from whole slide images of the human brain. Moreover, another application of this system is to extend it to the processing of organotypic brain slice culture models, which would help increase throughput in investigations such as the one in Healy et al. (2016).

We expect that software of this nature can amass large numbers of features from whole-slide images, and make predictive connections between histological data and clinical data serving as a powerful diagnostic tool that aids pathologists and neuroscientists to a significant extent. Additionally, through automation, after a more extensive validation procedure, this software pipeline could be made available for other research centers, e.g., by the means of a web application programming interface, to integrate into their analysis. Consequently, this would result in an increased number of pathologists and neuroscientists using the software and building an absolute consensus for the data with the possibility of integration into 3-dimensional pathology. Digital Pathology clearly has an exciting future in this imaging modality.

Acknowledgements This work was supported in part by funding from Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research (CIHR), the BC Innovation Council NRAS Program, and the Weston Brain Institute. Support from Professor Salcudean's C. A. Laszlo Chair is gratefully acknowledged. Dr. Nir is a recipient of a Prostate Cancer Canada Post-Doctoral Research Fellowship Award #PDF2016-1338.

Compliance with Ethical Standards

Ethical Approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted.

Conflict of interests The authors declare that they have no conflict of interest.

Information Sharing Statement The data and source code utilized in this work are available at https://www.frdr.ca/repo/handle/doi:10.20383/101.0120/ and https://github.com/rcldigitalpathology/Brain-Analyzer/, respectively.

References

Abdolhoseini, M., Walker, F., Johnson, S. (2016). Automated tracing of microglia using multilevel thresholding and minimum spanning trees. In *2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC)* (pp. 1208–1211). IEEE.

Bae, M.H., Pan, R., Wu, T., Badea, A. (2009). Automated segmentation of mouse brain images using extended mrf. *NeuroImage*, *46*(3), 717–725.

Balis, U., Hipp, J., Flotte, T., Monaco, J., Cheng, J., Madabhushi, A., Yagi, Y., Rodriguez-Canales, J., Emmert-Buck, M., Dugan, M., et al. (2011). Computer aided diagnostic tools aim to empower rather than replace pathologists: lessons learned from computational chess. *Journal of Pathology Informatics*, *2*(1), 25. https://doi.org/10.4103/2153-3539.82050.

Bishop, C.M. (2006). *Pattern recognition and machine learning (information science and statistics), 1st edn. 2006. corr. 2nd printing edn*. New York: Springer.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Chen, Z., Shin, D., Chen, S., Mikhail, K., Hadass, O., Tomlison, B.N., Korkin, D., Shyu, C.R., Cui, J., Anthony, D.C., et al. (2014). Histological quantitation of brain injury using whole slide imaging: a pilot validation study in mice. *PloS One*, *9*(3), e92,133.

Cheng, W.H., Stukas, S., Martens, K.M., Namjoshi, D.R., Button, E.B., Wilkinson, A., Bashir, A., Robert, J., Cripton, P.A., Wellington, C.L. (2018). Age at injury and genotype modify acute inflammatory and neurofilament-light responses to mild chimera traumatic brain injury in wild-type and app/ps1 mice. *Experimental Neurology*, *301*, 26–38.

Cherry, J.D., Tripodis, Y., Alvarez, V.E., Huber, B., Kiernan, P.T., Daneshvar, D.H., Mez, J., Montenigro, P.H., Solomon, T.M., Alosco, M.L., et al. (2016). Microglial neuroinflammation contributes to tau accumulation in chronic traumatic encephalopathy. *Acta Neuropathologica Communications*, *4*(1), 112.

Cireşan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems* (pp. 2843–2851).

Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J. (2013). Mitosis detection in breast cancer histology images with deep neural networks. In *International conference on medical image computing and computer-assisted intervention* (pp. 411–418). Springer.

Costa, A. (2013). Hausdorff (box-counting) fractal dimension. https://www.mathworks.com/matlabcentral/fileexchange/30329-hausdorff-box-counting-fractal-dimension/, [Online; accessed March 2017].

Coughlin, J.M., Wang, Y., Minn, I., Bienko, N., Ambinder, E.B., Xu, X., Peters, M.E., Dougherty, J.W., Vranesic, M., Koo, S.M., et al. (2017). Imaging of glial cell activation and white matter integrity in brains of active and recently retired national football league players. *JAMA Neurology*, *74*(1), 67–74.

Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *Medical imaging 2014: digital pathology, international society for optics and photonics* (Vol. 9041, p. 904103).

Ding, Y., Pardon, M.C., Agostini, A., Faas, H., Duan, J., Ward, W.O., Easton, F., Auer, D., Bai, L. (2017). Novel methods for microglia segmentation, feature extraction, and classification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *14*(6), 1366–1377.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31*(4-5), 198–211. https://doi.org/10.1016/j.compmedimag.2007.02.002.

Duda, R.O., Hart, P.E., Stork, D.G. (2012). *Pattern classification*. New York: Wiley.

Ertosun, M.G., & Rubin, D.L. (2015). Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. In *AMIA annual symposium proceedings, American medical informatics association* (Vol. 2015, pp. 1899–1908).

Farahani, N., Parwani, A.V., Pantanowitz, L. (2015). Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Journal of Pathology and Laboratory Medicine International*, *7*, 23–33.

Finkelstein, E., Corso, P.S., Miller, T.R. (2006). *The incidence and economic burden of injuries in the United States*. Oxford: Oxford University Press.

Freund, Y., Schapire, R., Abe, N. (1999). A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, *14*(5), 771–780.

Ghaznavi, F., Evans, A.H., Madabhushi, A., Feldman, M.D. (2013). Digital imaging in pathology: whole-slide imaging and beyond. *Annual Review of Pathology*, *8*(1), 331–59.

Gonzalez, R.C., & Woods, R.E. (2008). *Digital image processing*. Upper Saddle River: Pearson-Prentice-Hall.

Healy, S., McMahon, J., Owens, P., FitzGerald, U. (2016). Significant glial alterations in response to iron loading in a novel organotypic hippocampal slice culture model. *Scientific Reports*, *6*, 36,410.

Holmin, S., Söderlund, J., Biberfeld, P., Mathiesen, T. (1998). Intracerebral inflammation after human brain contusion. *Neurosurgery*, *42*(2), 291–298.

Hosmer, D.W.Jr., Lemeshow, S., Sturdivant, R.X. (2013). *Applied logistic regression,* Vol. 398. New York: Wiley.

Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, *7*(29). https://doi.org/10.4103/2153-3539.186902.

Johnson, S.J., & Walker, F.R. (2015). Strategies to improve quantitative assessment of immunohistochemical and immunofluorescent labelling. *Scientific Reports*, *5*, 10,607.

Johnson, V.E., Stewart, J.E., Begbie, F.D., Trojanowski, J.Q., Smith, D.H., Stewart, W. (2013). Inflammation and white matter degeneration persist for years after a single traumatic brain injury. *Brain: A Journal of Neurology*, *136*(1), 28–42.

Kaplan, K.J., & Rao, L.K. (2015). *Digital pathology: historical perspectives, current concepts & future applications*. Cham: Springer.

Karperien, A., Ahammer, H., Jelinek, H.F. (2013). Quantitating the subtleties of microglial morphology with fractal analysis. *Frontiers in Cellular Neuroscience*, *7*(3), 1–18. https://doi.org/10.3389/fncel.2013.00003.

Karperien, A.L., & Jelinek, H.F. (2015). Fractal, multifractal, and lacunarity analysis of microglia in tissue engineering. *Frontiers in Bioengineering and Biotechnology*, *3*, 51.

Kettenmann, H., Hanisch, U.K., Noda, M., Verkhratsky, A. (2011). Physiology of microglia. *Physiological Reviews*, *91*(2), 461–553.

Kollmannsberger, P. (2009). Skeleton3d. https://www.mathworks.com/matlabcentral/fileexchange/43400-skeleton3d/, [Online; accessed March 2017].

Kozlowski, C., & Weimer, R.M. (2012). An automated method to quantify microglia morphology and application to monitor activation state longitudinally in vivo. *PloS One*, *7*(2), e31,814.

Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).

Li, J., Najmi, A., Gray, R.M. (2000). Image classification by a two-dimensional hidden markov model. *IEEE Transactions on Signal Processing*, *48*(2), 517–533.

Li, W., Liu, H.D., You, W.C., Zhou, M.L., Ling, H.P., Shen, W., Zhu, L., Hang, C.H. (2013). Enhanced cortical expression of myeloid differentiation primary response protein 88 (myd88) in patients with traumatic brain injury. *Journal of Surgical Research*, *180*(1), 133–139.

Lieh-Lai, M.W., Theodorou, A.A., Sarnaik, A.P., Meert, K.L., Moylan, P.M., Canady, A.I. (1992). Limitations of the glasgow coma scale in predicting outcome in children with traumatic brain injury. *The Journal of Pediatrics*, *120*(2), 195–199.

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

Madabhushi, A., & Lee, G. (2016). Image analysis and machine learning in digital pathology: challenges and opportunities.

McKee, A.C., Cantu, R.C., Nowinski, C.J., Hedley-Whyte, E.T., Gavett, B.E., Budson, A.E., Santini, V.E., Lee, H.S., Kubilus, C.A., Stern, R.A. (2009). Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *Journal of Neuropathology & Experimental Neurology*, *68*(7), 709–735.

McKee, A.C., Stein, T.D., Nowinski, C.J., Stern, R.A., Daneshvar, D.H., Alvarez, V.E., Lee, H.S., Hall, G., Wojtowicz, S.M., Baugh, C.M., et al. (2013). The spectrum of disease in chronic traumatic encephalopathy. *Brain: A Journal of Neurology*, *136*(1), 43–64.

Namjoshi, D.R., Cheng, W.H., McInnes, K.A., Martens, K.M., Carr, M., Wilkinson, A., Fan, J., Robert, J., Hayat, A., Cripton, P.A., et al. (2014). Merging pathology with biomechanics using chimera (closed-head impact model of engineered rotational acceleration): a novel, surgery-free model of traumatic brain injury. *Molecular Neurodegeneration*, *9*(1), 55.

Namjoshi, D.R., Cheng, W.H., Bashir, A., Wilkinson, A., Stukas, S., Martens, K.M., Whyte, T., Abebe, Z.A., McInnes, K.A., Cripton, P.A., et al. (2017). Defining the biomechanical and biological threshold of murine mild traumatic brain injury using chimera (closed head impact model of engineered rotational acceleration). *Experimental Neurology*, *292*, 80–91.

Nir, G., Hor, S., Karimi, D., Fazli, L., Skinnider, B.F., Tavassoli, P., Turbin, D., Villamil, C.F., Wang, G., Wilson, R.S., et al. (2018). Automatic grading of prostate cancer in digitized histopathology images: learning from multiple experts. *Medical Image Analysis*, *50*(12), 167–180.

Petushi, S., Garcia, F.U., Haber, M.M., Katsinis, C., Tozeren, A. (2006). Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC Medical Imaging*, *6*(1), 14–24. https://doi.org/10.1186/1471-2342-6-14.

Popescu, C., Anghelescu, A., Daia, C., Onose, G. (2015). Actual data on epidemiological evolution and prevention endeavours regarding traumatic brain injury. *Journal of Medicine and Life*, *8*(3), 272–277.

Ramlackhansingh, A.F., Brooks, D.J., Greenwood, R.J., Bose, S.K., Turkheimer, F.E., Kinnunen, K.M., Gentleman, S., Heckemann, R.A., Gunanayagam, K., Gelosa, G., et al. (2011). Inflammation after trauma: microglial activation and traumatic brain injury. *Annals of Neurology*, *70*(3), 374–383.

Rashid, S., Nir, G., Fazli, L., Boag, A.H., Siemens, D.R., Goldenberg, S.L., Abolmaesumi, P., Salcudean, S.E. (2018). Automatic pathology of prostate cancer in whole mount slides incorporating individual gland classification. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. https://doi.org/10.1080/21681163.2018.1514280.

Rouchdy, Y., & Cohen, L.D. (2013). Geodesic voting for the automatic extraction of tree structures. methods and applications. *Computer Vision and Image Understanding*, *117*(10), 1453–1467.

Sirinukunwattana, K., Raza, S.E.A., Tsang, Y.W., Snead, D.R., Cree, I.A., Rajpoot, N.M. (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging*, *35*(5), 1196–1206.

Stein, T.D., Alvarez, V.E., McKee, A.C. (2014). Chronic traumatic encephalopathy: a spectrum of neuropathological changes following repetitive brain trauma in athletes and military personnel. *Alzheimer'S Research & Therapy*, *6*(1), 4.

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Berlin: Springer Science & Business Media.

Strekalovskiy, E., & Cremers, D. (2014). Real-time minimization of the piecewise smooth mumford-shah functional. In *European conference on computer vision* (pp. 127–141). Springer.

Taylor, C., Bell, J., Breiding, M., Xu, L. (2017). Traumatic brain injury–related emergency department visits, hospitalizations, and deaths—United States, 2007 and 2013. *MMWR Surveillance Summaries*, *66*(9), 1–16.

Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., Murray, G. (2014). The glasgow coma scale at 40 years: standing the test of time. *The Lancet Neurology*, *13*(8), 844–854.

The MathWorks (2017a). Matlab: Deep learning. https://www.mathworks.com/help/nnet/examples/create-simple-deep-learning-network-for-classification.html/ [Online; accessed 1 December 2017].

The MathWorks (2017b). Matlab: Regional maxima. https://www.mathworks.com/help/images/ref/imregionalmax.html/ [Online; accessed 1 December 2017].

Vadakkan, T.J. (2009). Lacunarity of a binary image. https://www.mathworks.com/matlabcentral/fileexchange/25261-lacunarity-of-a-binary-image/, [Online; accessed August 2018].

Valous, N.A., Lahrmann, B., Zhou, W., Veltkamp, R., Grabe, N. (2013). Multistage histopathological image segmentation of iba1-stained murine microglias in a focal ischemia model: methodological workflow and expert validation. *Journal of Neuroscience Methods*, *213*(2), 250–262.

Van Gool, L., Szekely, G., Ferrari, V. (2014). Computer Vision.

Wu, T., Bae, M.H., Zhang, M., Pan, R., Badea, A. (2012). A prior feature svm-mrf based method for mouse brain segmentation. *NeuroImage*, *59*(3), 2298–2306.

Affiliations

Alexander D. Kyriazis[1] · Shahriar Noroozizadeh[1] · Amir Refaee[1] · Woongcheol Choi[1] · Lap-Tak Chu[1] · Asma Bashir[2] · Wai Hang Cheng[2] · Rachel Zhao[2] · Dhananjay R. Namjoshi[2] · Septimiu E. Salcudean[3] · Cheryl L. Wellington[2] · Guy Nir[4]

Septimiu E. Salcudean
tims@ece.ubc.ca

Cheryl L. Wellington
wcheryl@mail.ubc.ca

[1] Engineering Physics, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

[2] Department of Pathology and Laboratory Medicine, University of British Columbia, 2215 Wesbrook Mall, Vancouver, BC V6T 1Z3, Canada

[3] Department of Electrical and Computer Engineering, University of British Columbia, 2332 Main Mall, Vancouver, BC V6T 1Z4, Canada

[4] Department of Urologic Sciences, University of British Columbia, 2775 Laurel Street, Vancouver, BC V5Z 1M9, Canada