Automated Feature Extraction from Brain Lesion Masks using Deep Learning for Predicting Secondary Progressive Multiple Sclerosis Progression

Marco Law\textsuperscript{1} \quad \texttt{MARCOLAW@MSMRI.MEDICINE.UBC.CA}

\textbf{1} SBME, UBC

Anthony Traboulsee\textsuperscript{2} \quad \texttt{ANTHONY.TRABOULSEE@MD.UBC.CA}

\textbf{2} Department of Neurology, Faculty of Medicine, UBC

David K.B. Li\textsuperscript{3} \quad \texttt{D.K.LI@MEDICINE.UBC.CA}

\textbf{3} Department of Radiology, Faculty of Medicine, UBC

Robert Carruthers\textsuperscript{2} \quad \texttt{ROBERT.CARRUTHERS@MD.UBC.CA}

Mark S. Freedman\textsuperscript{4} \quad \texttt{MARK.FREEDMAN@OHRI.UBC.CA}

\textbf{4} Division of Neurology, Department of Medicine, OHRI

Shannon Kolind\textsuperscript{3} \quad \texttt{SHANNON.KOLLIND@MD.UBC.CA}

Roger Tam\textsuperscript{3} \quad \texttt{ROGER.TAM@MD.UBC.CA}

Editors: Under Review for MIDL 2019

Abstract

Predicting disability progression in multiple sclerosis (MS) has been largely unsuccessful using traditional image analysis and statistical approaches. Both clinical features and imaging biomarkers such as brain atrophy measurements and lesion volume have only weak correlations with disability scores. Identification of patients at high risk of disability progression could accelerate clinical trial development, inform choice of disease modifying therapy, and lead to improved understanding of MS pathogenesis. Deep learning networks have been shown to automatically extract features from brain lesion masks for predicting conversion from one MS disease course to the next. We propose a deep learning network for automatic extraction of brain lesion features from baseline binary lesion masks for short-term prediction of disability worsening and compare it to a multivariate logistic regression model using only baseline user-defined clinical, demographical, and MRI features.

\textbf{Keywords:} deep learning, machine learning, disease progression, prediction, secondary progressive multiple sclerosis

1. Introduction

Multiple sclerosis (MS) is a chronic autoimmune demyelinating disease of the central nervous system (CNS), characterized by the catabolism of the myelin sheath that surrounds and insulates axons of nerve cells. Predicting disability progression in MS has largely been unsuccessful using traditional analysis and statistical approaches due to the high degree of variability in disability outcomes (Kister et al., 2013). Many clinical features such as age, disease duration, and sex have moderate-at-best correlations with disability scores on the Expanded Disability Status Scale (EDSS) (Mowry, 2011; Amato and Ponziani, 2000;...
Predicting Secondary Progressive Multiple Sclerosis Progression with Deep Learning

Kurtzke, 1983). For monitoring disease progression, imaging biomarkers used such as brain atrophy and white matter lesion count and volume also only weakly correlate with current or future disability (Barkhof, 2002). The ability to predict patients at risk of disability progression could better allow informed choice of disease modifying therapy, accelerate clinical trials and improve MS pathogenesis understanding.

Deep learning (DL) is a class of machine learning algorithms capable of extracting features from multi-dimensional data shown to automatically learn features from existing binary brain lesion masks that are better correlated with MS disability scores than lesion volume (Brosch et al., 2014). Additionally, DL has been used to learn features for distinguishing between healthy controls and MS patients, differentially diagnose Neuromyelitis Optica and MS, as well as predict conversion from clinically-isolated syndrome to definite-MS (Yoo et al., 2017a, 2018, 2017b). We hypothesize that deep learning networks may be capable of extracting features from lesion masks with prognostic value for secondary progressive MS, a disease course characterized by a period of acute disease worsening followed by periods of plateau. We use a deep learning network (DLN) for extracting features from binary lesion masks for predicting short-term disability progression in SPMS compared to logistic regression.

2. Methods

SPMS patients with complete clinical visits and baseline clinical data (n=485) enrolled in a 2-year negative placebo-controlled trial assessing the efficacy of MBP8298 were categorized as either progressors, defined by a 6-month sustained-increase in Expanded Disability Status Scale (EDSS) (≥1.0 and ≥0.5 for baseline of ≤5.5 and ≥6.0 respectively), or not. Features were extracted from brain lesion masks using a deep learning network (Figure ??), used with (coDLN) and without (lmDLN) user-defined demographic (age, sex), clinical (disease duration, EDSS, Multiple Sclerosis Functional Composite components), and MRI (T2 lesion volume, T2LV, and brain parenchymal fraction, BPF) features at baseline to predict progression, and compared to logistic regression (LR) with user-defined features on area under the receiver-operator characteristic curve (AUC), precision, negative predictive value (NPV), sensitivity, and specificity using 10-fold stratified cross validation. Random under-sampling was used to correct for class-imbalance during network training.

MS lesions are typically very dispersed, resulting in highly sparse volumes which may lead to noise patterns being learned (Yoo et al., 2017b). To increase information density, the signed-Euclidean distance transform (EDT) was applied to each binarized brain lesion mask. EDT assigns voxel intensity as the Euclidean distance from each voxel to the nearest lesion voxel. An sample slice of a typical brain lesion mask and its EDT is shown in Figure 1.

The deep learning network used is comprised of three convolutional layers using leaky-rectified linear activation units (LeakyRELU) for nonlinearity (Xu et al., 2015). The convolutional layers consisted of 12, 24, and 48 filters of sizes 7x7x7, 5x5x5, and 3x3x3 respectively. Max-pooling layers of size 2x2x2 were used after each convolutional layer to reduce dimensionality. The output of the final max-pooling layer was then flattened into a one-dimensional feature vector and used as input to a fully-connected layer with 256 nodes. In lmDLN, the outputs are then passed into a logistic regression layer, whereas in coDLN,
Predicting Secondary Progressive Multiple Sclerosis Progression with Deep Learning

Figure 1: Axial (transverse) slice of a brain lesion mask (left) and the Euclidean distance transformed lesion mask (right)

the user-defined features were concatenated with the image features before being used for classification by a logistic regression layer.

3. Results

There were 115 progressors and 370 non-progressors. The DLN performed better than LR in AUC (55.0% vs. 45.0%), precision (27.0% vs 22.0%) and NPV (79.1% vs. 74.9%) at \( p < 0.05 \) using paired t-test. The inclusion of user-defined features did not alter DLN performance. No differences in sensitivity and specificity between deep learning networks and logistic regression.

4. Discussion and Conclusions

Using only lesion masks, DL was able to learn features more predictive of disability progression in SPMS than LR using only user-defined features. Observed improvements in PPV and NPV compared to the naïve implementation of multivariate logistic regression of user-defined features when using deep-learned features from binary lesion masks may be due to the ability for the DLN to consider spatial information (which is lost in conventional MRI metrics such as BPF and T2LV) in addition to volumetric information from the masks. A major limitation of this experiment is the limited sample of 485 participants. This sample size is unlikely to capture enough variation of lesion distributions or user-defined features of the SPMS population for modelling with either logistic regression or deep learning. Future work would look at increasing the sample size, particularly that of progressors, and the inclusion of longitudinal data.

Acknowledgments

We thank the Multiple Sclerosis Society of Canada for their support and interest in our work on predict disability progression in MS
References


Youngjin Yoo, Lisa Y.W. Tang, Tom Brosch, David K.B. Li, Shannon Kolind, Irene Vavasour, Alexander Rauscher, Alex L. MacKay, Anthony Traboulsee, and Roger C. Tam.