Improving the accuracy of intraoperative neuroimaging by spatially-regularized semiparametric regression

Nishant Kumar
Supervisor: Nico Hoffmann

Abstract

The development of intraoperative imaging-guided neurosurgery represents a substantial improvement in the microsurgical treatment of malignant tissues in human brain. However, changes in regional cerebral blood flow greatly alter the emitted heat radiation of the cortex leading to a non-linear random behavior. Semi-parametric regression model adds the deterministic or parametric components of state-of-the-art Generalized Linear Models (GLM) with non-parametric components such as P-Splines to combat the non-linearity. However, in order to model spatial-temporal interactions in the thermographic brain imaging data, the semi-parametric regression using penalized splines has to be extended by a Markov random field (MRF) component. The MRF requires fast inference schemes such as Belief Propagation (BP), Tree-reweighted message passing (TRWS) etc in order to fulfill intra-operative performance requirements. Therefore, OpenGM framework is used to achieve this goal. The advancements proposed in this work will aid the neurosurgeons to achieve accurate removal of malignant tissues with minimal disruption of surrounding healthy neuronal matter.

Overview & Purpose

The Generalized Linear Model (GLM) is a way of unifying various interactions, such that each outcome Y of the dependent variables are generated from a particular probability distribution that includes the normal, binomial, Poisson, Bernoulli and gamma distributions, among others. A common shortcoming with this approach is that it is unable to catch certain random effects in the data which may portray non-linear relationships originating from background noise. Hence, the GLM can be extended by non-parametric components which lead to semi-parametric regression or partially linear model. These models integrate deterministic components with non-parametric components such as penalized splines or P-Splines [8]. P-Splines are able to approximate certain signals by a low degree of freedom (so called knots) using a penalty function which is applied to the P-Spline estimate's coefficients.

The prior derivation can be applied to the time-series of a single pixel. However, the spline coefficients can be interpreted as labels which are to be assigned to each pixel. Furthermore, the coefficients are assumed to behave as a spatially correlated stochastic or random process. The latter can be modeled by a Markov random field (often abbreviated as MRF) which is a set of random variables indexed by spatial positions. In MRF, each pixel (often termed as a site) of an image act as a random variable having a Markov property described by an undirected graph. This basically means that the label of a single pixel is only affected by the labels of its nearest neighbors. In other words, a random field is said to be Markov random field if it satisfies the Markovianity. Markovianity can be termed as the property of a random variable, if the conditional probability distribution of the random variable depends only upon its neighboring variables instead of all the random variables present in the system. Hence, the inter-relationship between the sites is maintained by the so-called neighborhood system or clique potentials. A label set with discrete labels is known as discrete Markov random fields. The energy function of a discrete Markov random field is a sum of clique potentials over all possible cliques. The value of clique potential depends on the local configuration on the clique. An important special case is when only cliques of size up to two i.e first order are considered. The problem of minimization of this energy function in the form of factor graphs to spatially smooth the data and detect the foreground and background pixels will help in improving the accuracy of intraoperative neuroimaging. The inference algorithms used to minimize the energy function can be roughly grouped into three classes. Linear programming based methods provide a lower bound for the optimum. Move-making based methods iteratively improve the labeling, and message passing methods are simple to implement and can be parallelized easily – often motivated by linear programming and variational optimization.

Overall, in order to model spatial-temporal interactions in the thermographic brain imaging data, the semiparametric regression using penalized splines in the generalized additive model approach has to be extended by a Markov random field component. The latter requires an efficient model formulation as well as fast inference schemes in order to fulfill intraoperative performance requirements.

Focus

1. Literature research on spatially-regularized Spline regression models
    a. Basis functions (Wavelet, B-Spline, Truncated Polynomial)
    b. Regularizers (local (L1, L2), spatial (MRF))

2. Provide a mathematical model of the chosen framework and derive the inference procedure

3. Implement the chosen framework in Python by using existing frameworks (OpenGM, SciPy, numpy, keras etc.)
    a. Discretization of the model’s coefficients
    b. Hypothesis testing to recognize the deterministic component
    c. Discuss potential performance issues and provide potential solutions

4. Evaluate performance (accuracy, speed) with respect to univariate semiparametric regression
    a. Semi-synthetic data:
        i. Quantify the effect of certain spatial regularizers (Potts potential, Squared Distance, Truncated Squared Distance).
        ii. Quantify the discretization error.
        iii. Do a performance analysis of the univariate and spatially-regularized regression framework and unveil potential bottlenecks (such as inference schemes (TRW-S, LBP, AD3), parameter estimation of the univariate model)
    b. Intraoperative data:
        i. Compare your method qualitatively to the univariate method with respect to the accuracy in detecting neural activity.