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

**Nishant Kumar**

**Supervisor: Nico Hoffmann**

**Overview & Purpose**

Thermography studies the heat radiations in different structures or regions. One of the unique example is its use in detecting tumors in human brain. It analyses the heat radiated from brain tissues to determine whether it is a benign or a malign tissue. This greatly helps medical professionals to locate the bad tissues in brain. However, the radiations obtained from the surface is highly dynamic and non-stationary with continuous cerebral blood flow changes along with high environmental interference. Therefore, it is of utmost priority to analyze the multidimensional data in such a way that it captures spatial as well as temporal interactions.

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 leads 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 MRFs, 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 Markov property. The inter-relationship between the sites is maintained by a so-called neighborhood system or pairwise potentials. A label set may be categorized as being continuous or discrete. MRFs with discrete labels are known as discrete Markov random fields whereas MRFs with continuous labels are known as Gaussian Markov random fields.

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
   1. Basis functions (Wavelet, B-Spline, Truncated Polynomial)
   2. 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, …)
   1. Discretization of the model’s coefficients
   2. Hypothesis testing to recognize the deterministic component
   3. Discuss potential performance issues and provide potential solutions
4. Evaluate performance (accuracy, speed) with respect to univariate semiparametric regression
   1. Semi-synthetic data:
      1. Quantify the effect of certain spatial regularizers (Potts potential, Squared Distance, Truncated Squared Distance).
      2. Quantify the discretization error.
      3. 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)
   2. Intraoperative data:
      1. Compare your method qualitatively to the univariate method with respect to the accuracy in detecting neural activity.