# Diffeomorphic Metric Learning and Template Optimization for Registration-Based Predictive Models

Ayagoz Mussabayeva[1], Maxim Pisov[1,3], Anvar Kurmukov[1,4], Alexey Kroshnin[1,4], Yulia Denisova[1], Li Shen[5], Shan Cong[6], Lei Wang[7], and Boris Gutman[1,2]

[1] The Institute for Information Transmission Problems, Moscow, Russia
[2] Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, USA
[3] Moscow Institute of Physics and Technology State University, Moscow, Russia
[4] Higher School of Economics, Moscow, Russia
[5] Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, USA
[6] Indiana University, Indianapolis, IN, USA
[7] Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

**Abstract.** We present a method for metric optimization and template construction in the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework. The construction treats the Riemannian metric on the space of diffeomorphisms as a data-embedding kernel in the context of predictive modeling, here Kernel Logistic Regression (KLR). The task is then to optimize kernel parameters, including the LDDMM metric parameters as well as the registration template, resulting in a parameterized argminimum optimization. In practice, this leads to a group-wise registration problem with the goal of improving predictive performance, for example by focusing the metric and template on discriminating patient and control populations. We validate our algorithm using two discriminative problems on a synthetic data set as well as 3D subcortical shapes from the SchizConnect cohort. Though secondary to the template and kernel optimization, accuracy of schizophrenia classification is improved by LDDMM-KLR compared to linear and RBF-KLR.

**Keywords:** image registration, machine learning, subcortical shape, metric learning, LDDMM

## 1 Introduction

The problem of image registration arises in a number of medical imaging contexts. The process of registering images underlies concurrent use of spatially distributed information from multiple data sets, possibly collected at different times or using different imaging modalities.

The practical goal of registration is spatial alignment. In addition, a subset of registration algorithms has the benefit of establishing a formal distance between images as a byproduct of spatial normalization. Of particular interest among this class of methods is the Large Deformation Diffeomorphic Metric Mapping (LDDMM) [1]. The distance between images defined by LDDMM derives from a Riemannian metric on velocity fields. The distance is then a metric on diffeomorphisms, or the deformations themselves. The Riemannian structure implies the possibility of using the metric as a kernel in a machine learning (ML) context. Such a distance can be derived in a straightforward manner, as in [2] or as the initial momentum "slope." The latter approach is taken here. Yet, to date, little effort in this direction has been undertaken, in part due to the computational complexity of the task. Further, the fact that the LDDMM distance is defined on diffeomorphisms means that in practice the metric in a group of images will be heavily dependent on the choice of the template image. The parameters of the metric itself also significantly influence the registration results [3]. In light of this, applying the LDDMM Riemannian metric as a kernel in a predictive model requires optimization of both the metric parameters and the template. This is the problem we address here. Kernel optimization is an established ML problem. Several canonical kernel forms exist [4], many of which have been studied in the context of metric learning [5], [6]. In line with this research, we extend previous work on diffeomorphic metric learning in [2].

A related, though conceptually different notion of learning the LDDMM metric focuses instead on spatially varying the kernel parameters [7] This method allows to find more flexible metric representation, while still satisfying the metric properties. However, the key difference lies in the cost function. In [7] and related work, the metric is optimized with respect to standard registration costs, with additional novel regularization on the LDDMM parameter distribution. Though far more flexible than LDDMM with global parameters, the framework is still conceptually similar to standard LDDMM in that it targets traditional pair-wise registration problems. The metric is not learned in a predictive modeling (group-wise) sense, and its inner product properties are not exploited in the sense of a machine learning "kernel trick."

Similarly to what we do here, the authors in [2] developed an automatic method to find LDDMM metric parameters for optimal classification accuracy. However, there are two important distinctions: (1) rather than using the LDDMM Riemannian product, a distance-derived image similarity measure was used instead. The previous method would be more accurately described as a similarity learning approach. (2) Because only pairwise image distances were used, no template optimization was performed. In fact, the improved template construction, along with optimal registration parameters, potentially has greater practical value to the imaging researcher than the classification tool itself. By optimizing the template for a specific biological question, we enable a more meaningful interpretation of traditional localized registration-based measures, e.g. Jacobian determinant or momentum operator maps, in neuroimaging studies. The focus is thus not on classification accuracy, though it is indeed improved

over simpler models, but on template and parameter optimization. In this respect, this work closely parallels the template construction and metric parameter optimization in [8].

The remainder of the paper is organized as follows. Section 2 briefly describes LDDMM. Section 3 builds the model and optimization specifics of LDDMM kernel and template learning in the context of kernel logistic regression. Section 4 describes the synthetic and real brain MRI-based experiments, and 5 concludes the paper.

## 2 Image Registration in the LDDMM Framework

Large Deformation Diffeomorphic Metric Mapping (LDDMM) was first introduced and implemented in [1]. Since then, LDDMM has been extended into a several methods. The main idea remains the same: to compute a matching diffeomorphism $\phi\colon [0,1]\times\Omega\to\Omega$, where $\Omega$ is the image domain and $[0,1]$ is the time interval. The diffeomorphism $\phi$ can be seen from different points of view. In its original definition, $\phi$ belonged to a large class of functions. For the purposes of explicit computation, diffeomorphism $\phi$ can be considered as the flow generated by the evolution equation:

$$\begin{aligned}\frac{\partial\phi(t,x)}{\partial t} &= v(t,\phi(t,x)),\\ \phi(0,x) &= x,\end{aligned} \tag{1}$$

where $v\in V$ is an element in the space of smooth vector fields.

Let $I,J\in\Omega$ be the source and target images, respectively. The optimization problem for registering $I,J$ is generally formulated as a cost function comprised of a similarity measure between warped source and target images and a regularization term:

$$E(v)=\frac{1}{\sigma^2}d(I\circ\phi,J)+R(v) \tag{2}$$

where $\sigma^2$ is the normalization coefficient.

Here, we define similarity as the sum of squared differences (SSD) $d(I\circ\phi,J)=\|I\circ\phi-J\|^2_{L_2}$, where $I\circ\phi$ is the warped image.

The diffeomorphism space is a Lie Group with respect to (1) and it suggests a right-invariance property, allowing to recall the adjoint representation and the scalar product on the tangent space. Smoothness of the warp can be enforced by defining an appropriate regularization term. As long as the tangent space $V$ has a well-defined inner product, regularization is defined by $\|\cdot\|_V$ through a Riemannian metric $L$. The operator $L$ should be naturally defined by the geometric structure of the domain and implies smoothness. As it was in [1] $L$ is defined as an adjoint differential operator (3). Finally, regularization is defined as the inner product integrated along the path of the diffeomorphism $\phi^v$ and shown in equation (3). Note that here the superscript of $\phi$ defines the dependence on $v$; we omit this below:

$$R(v) = \int_0^1 \|v_t\|_V^2 \, dt = \int_0^1 < Lv_t, v_t > \, dt$$
$$L = (\alpha\Delta + \gamma E)^2, \tag{3}$$

where $\Delta$ is the laplacian operator and $E$ is identity operator.

The regularization term can also be thought of as a distance between images, i.e. the minimal length of diffeomorphic path required to transform the appearance of $I$ to be as similar as possible to $J$. Path length is defined as an integration of the velocity norm based on the corresponding Riemannian metric along the path. Importantly, the velocity field is defined on vector space with all the implicit properties.

The choice of the operator $L$ is a trade-off between simplicity and expressivity. We choose operator $L$ as in the original paper [1]. It consists of two scalar parameters, providing minimal flexibility for metric tuning. Parameters of the operator $(\alpha, \gamma)$ correspond to convexity and normalization terms, respectively. These parameters significantly affect the quality of the registration, as shown in [3].

### 2.1 Kernel Construction

As was mentioned above, the LDDMM registration produces a properly defined structure of interrelation between two images. One of several possible ways to construct kernel matrix is to use the length of the path in the diffeomorphic space. But since we consider registration to on a particular template, the inner product of the vector space, where velocity fields defined, is a natural choice. This accounts for diffeomorphic space geometry and employs the possible simplification in the case of template-based registration.

The matrix of pairwise distances is considered as a kernel for the predictive model and constructed as follows. Let $x_i \in D$ be an image, which after registration to on a particular template $J$ produces a corresponding velocity field and warping. The velocity field is constructed over time steps, denoted as $v_t^i = v^i(t, \phi(t, x))$ for each image $x_i, i = 1, \ldots, n$. Every vector field $v_1^i, i = 1, \ldots, n$ at time 1 is defined on the same tangent space at the identity transformation i.e., at the template $J$. Since this tangent space admits an inner product, the kernel is defined as $k(x_i, x_j) =< Lv_1^i, v_1^j >$. From here on, the time subscript will be omitted. As was mentioned above the meaning of such distance is a "slope" from source image to target image.

## 3 Metric Learning

In our work, we use Kernel Logistic Regression (KLR) as the base classifier. KLR is a flexible and tractable method that can be naturally used in the proposed method. Though a variety of classification and regression models admit kernels,

KLR is among the simplest and most robust, making minimal assumptions about data distribution.

Let $D = \{(x_i, y_i)\}_{i=1}^n$ be a sample, where the target variable $y \in \{1, 0\}$ is a class label. KLR leads to the following optimization problem:

$$Loss(\beta) = -\sum_{i=0}^{n} \big(y_i \cdot \ln p_i + (1 - y_i) \cdot \ln(1 - p_i)\big) + C \cdot \|\beta\|_{L^2}^2$$
$$p_i = \frac{1}{1 + e^{-\beta^T k(x_i, x)}}, \tag{4}$$

where $k(x_i, x)$ is a vector $(k(x_i, x_1), \ldots, k(x_i, x_n))^T$ and $C$ is a regularization coefficient.

### 3.1 Optimization Strategy

To address the registration parameter selection problem, we provide a gradient descent method as the strategy for optimizing the loss function (4) with respect to registration parameters $(\alpha, \gamma, J)$. The pipeline is shown in (fig. 1) and an outline of the algorithm is given in section 3.2. At the back propagation stage, we consider the solution of LDDMM and the classifier to be fixed, denoting both with a superscript $(\beta^*, v^*)$.

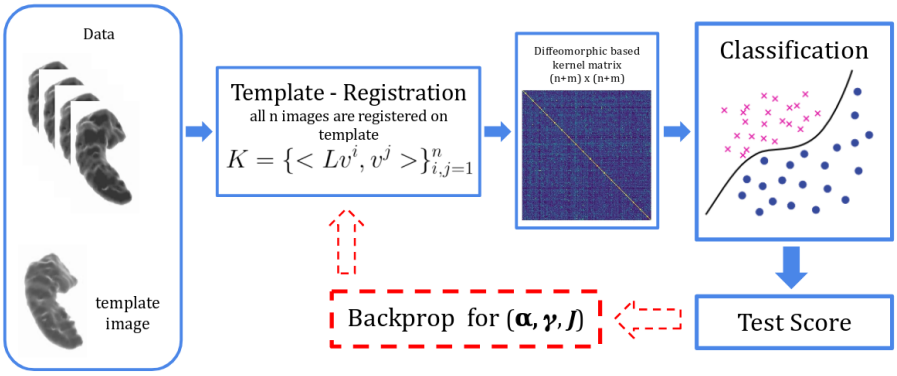


Fig. 1: **Schematic visualization of registration-based optimization algorithm ($v^i$ corresponds to optimal solution of LDDMM for subject $i$).**

Briefly, the KLR and LDDMM optimization pipeline can be describe as follows. We initialize $\alpha, \gamma, J$, compute the kernel, and find optimal KLR coefficients . Next, we fix $\beta^*, v^*$ and update $\alpha, \gamma, J$ with the gradient step of the loss function (4). The parameters $\beta^*$ and $v^*$ are the solutions of the optimization problem and depend on $\alpha, \gamma, J$. Therefore the derivative of loss function (4) is a bi-level optimization problem. The solution of a general bi-level optimization problem

is given in [9]. Below we provide step-by-step derivatives with respect to $\alpha$, and explain the main differences in derivatives with respect to $J$.

Denoting the train and validation subsets of the data as $X_{train}, X_{val}, y_{train}, y_{val}$, we compute $\beta^*$ as the argminimum of the loss function (4) on $\{X_{train}, y_{train}\}$. (Note that all unnecessary indices are again omitted.) The resulting gradients are computed on the validation set based on the following equations ($y_i, x_i \in \{X_{val}, y_{val}\}$):

$$\frac{dLoss(\beta^*)}{d\alpha} = -\sum_{i=1}^{n_{val}} \left( y_i \cdot \frac{d \ln p_i}{d\alpha} + (1 - y_i) \cdot \frac{d \ln(1 - p_i)}{d\alpha} \right) + 2C\beta^* \frac{d\beta^*}{d\alpha} \tag{5}$$

Expanding the mixed partial derivatives, we write:

$$\begin{gathered} \frac{\partial \ln p_i}{\partial \alpha} = -\frac{\partial \ln(1 + e^{-<\beta^*, k(x_i, x_{train})>})}{\partial \alpha} = \\ = \frac{e^{-<\beta^*, k(x_i, x_{train})>}}{1 + e^{-<\beta^*, k(x_i, x_{train})>}} \cdot \left( (\beta^*)^T \frac{\partial x_i}{\partial \alpha} + \frac{\partial (\beta^*)^T}{\partial \alpha} x_i \right) = \\ = (1 - p_i) \left( (\beta^*)^T \frac{\partial x_i}{\partial \alpha} + \frac{\partial (\beta^*)^T}{\partial \alpha} x_i \right); \\ \frac{\partial \ln(1 - p_i)}{\partial \alpha} = -p_i \left( (\beta^*)^T \frac{\partial x_i}{\partial \alpha} + \frac{\partial (\beta^*)^T}{\partial \alpha} x_i \right) \end{gathered} \tag{6}$$

$$\begin{gathered} \frac{\partial k(x_i, x_{train})}{\partial \alpha} = [k(x_i, x_{train}) =< Lv_i, v_{train} >] = \\ =< \frac{\partial L}{\partial \alpha} v_i, v_{train} > + < L \frac{\partial v_i}{\partial \alpha}, v_{train} > + < Lv_i, \frac{\partial v_{train}}{\partial \alpha} > \end{gathered} \tag{7}$$

The gradients below are also computed exclusively on $\{X_{train}, y_{train}\}$:

$$\begin{gathered} \frac{\partial \beta^*}{\partial \alpha} = \left[ \beta^* = \underset{\beta}{\operatorname{argmin}} \, Loss(\beta), (4), \textit{use formulae from ([9])} \right] \\ \frac{\partial^2 Loss}{\partial \beta^2} = X^T B X + 2E \\ \frac{\partial}{\partial \alpha} \frac{\partial Loss}{\partial \beta} = \frac{\partial}{\partial \alpha} (X^T (p - y)) = \frac{\partial X^T}{\partial \alpha} (p - y) + X^T \frac{\partial p}{\partial \alpha}, \end{gathered} \tag{8}$$

where $p$ and $y$ correspond to the KLR probability vector and training labels, and the matrix $B$ is diagonal with entries $p_i(1 - p_i)$.

We can see that all derivatives above include $\frac{dv^*}{d\alpha}$. To differentiate $v^*$ w.r.t. the scalar parameters, we use the same bi-level optimization problem solution. We use the finite difference approach rather than calculating $\frac{\partial^2 E(v)}{\partial v^2}$ for scalar parameters such as $\alpha$. In this work, we use the second order central finite difference. The same strategy can be used for $\gamma$. Also of note is that one of these parameters in 2 is redundant. It is the ratio of $\alpha, \gamma, \sigma^2$ that is of import.

One of the possible ways to find the derivative of the template image $J$ is to treat it like a scalar intensity function $J(x, y)$ defined on $\Omega$. Each voxel's intensity is treated as independent from the rest. We suggest a combined approach to obtain the derivative $\frac{dv^*}{dJ}$ as in the bi-level problem [9]. The derivative $\frac{\partial^2 E(v)}{dJdv}$ is derived from the closed form of the full derivative [1]. As only the second term of $\frac{\partial E(v)}{dv}$ depends on the template image $J$ and each voxel in $J$ is independent, the final form of the derivative is substantially simplifed. Below, we describe the optimization strategy for 2-dimensional images. However, it can be easily expanded to higher dimensions. To obtain $\frac{\partial^2 E(v)}{\partial v^2}$ we use a forth order central finite difference scheme for unmixed second derivative and for mixed second derivatives. For computational reasons, this approximation uses two key assumptions: first, only the voxel's immediate neighbors are correlated; second, the entry $v_y^{ij}$ does not interact with the entry $v_x^{kl}$ if $i \neq k$ or $j \neq l$, where the superscript refers to the axis, $i, j, k, l$ are the corresponding voxel coordinates, and $v_x$, $v_y$ denote components of the velocity field.

### 3.2 Optimization algorithm

To summarize the previous section, We have two loss functions $E(v)$ and $Loss(\beta)$, whose argminima $v^*, \beta^*$ parameterize the optimization of the registration parameters $(\alpha, \gamma, J)$. In practice, we find that (1) the LDDMM kernel multiplication constant does not affect the solution of optimization task and only the ratio $\alpha/\gamma$ has effect on the loss, we do omit the optimization of $\gamma$. The complete algorithm consists of two main parts: the "forward pass" and "the backward pass".

**Result:** $(\alpha^*, J^*)$
$\alpha = \alpha_0, J = J_0$;
**for** $i \leftarrow 0$ **to** $N$ **do**
  with fixed $\alpha_i, J_i$ forward pass:
  {
  calculate $\mathrm{v}_i^* = \operatorname{argmin} E(v)$;
  construct Kernels for train-val-test;
  fit KLR on train;
  evaluate metrics (ROC AUC) on val-test;
  }
  backward pass $\alpha, J$:
  {
  calculate gradients $(\frac{dLoss}{d\alpha_i}, \frac{dLoss}{dJ_i})$ on val;
  do gradient step:
  $\alpha_{i+1} = \alpha_i - lr \cdot \frac{dLoss}{d\alpha}$
  $J_{i+1} = J_i - lr \cdot \frac{dLoss}{dJ}$
  }
**end**

**Algorithm 1:** Pipeline of the proposed algorithm

## 4 Experiments

In our initial experiments, we used synthetic data to sanity-check our approach and hippocampal shape data from the ShizConnect dataset [10]. In all experiments, we fixed $\sigma^2 = 10^3$ and $\gamma = 1$. As well as the step for finite differences $h = 0.1$.

### 4.1 Synthetic Data

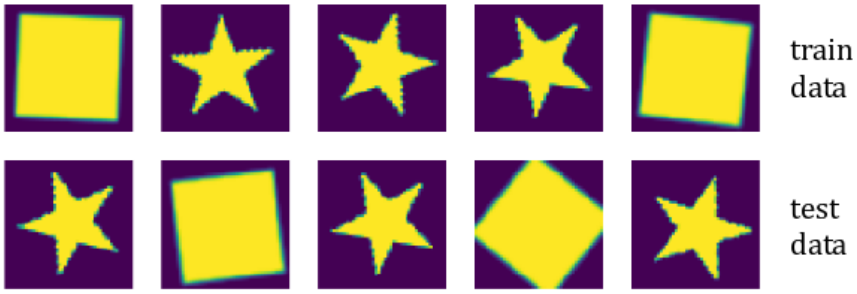


Fig. 2: Example of synthetic data: first row: train data, second row: test data (rectangle is labeled as 1 and star - 0.)

Our initial experiments were based on 200 synthetic images with two balanced classes: randomly rotated stars and rectangles (fig. 2). This task is very simple from the classification perspective. The primary interest here is to assess the validity of our simplifying assumptions in computing the derivatives, especially w.r.t to the template image $J$. Here, we used 10-fold cross-validation, fitting the L2 regularization term in KLR independently.

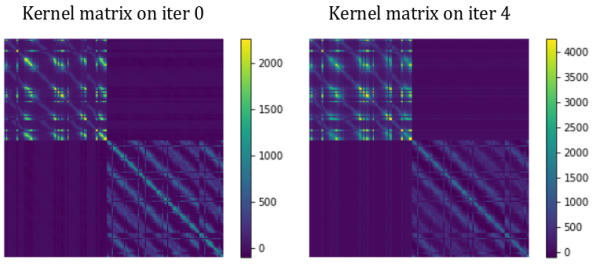


Fig. 3: Kernel matrix for the synthetic data set. The first half of the elements corresponds to data points labelled as stars, and the rest to rectangles

The result of template optimization is shown in (fig. 4). First, we initialize template $J$ as the mean image of the training set. Next, we make a gradient step on template $J$ and update its initial value. We initialize $\alpha = 0.4$. As we can see, the template begins to look like a star. In this task, the classification-optimal template appears to be closer to one of the classes. After iteration step 4, the

optimal template fluctuate around the template on 4 step. The distribution of distances is plotted in (fig. 3). We note that the ROC area under the curve (AUC) for this experiment is 1, achieving perfect classification.

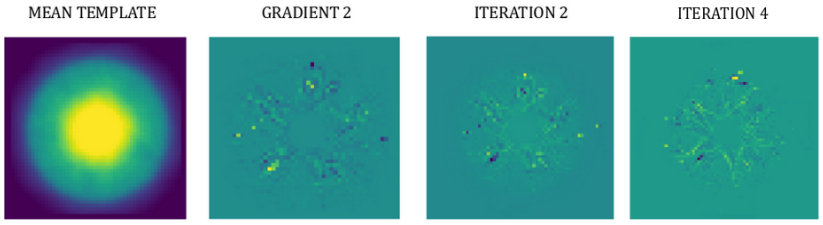


Fig. 4: 'MEAN TEMPLATE' image correspond to initialization of template as mean image of training set, 'GRADIENT 2' image is an example of derivative $\nabla_J Loss(\beta^*)$ on iteration step 2, 'ITERATION i' image is the template image $J_i$ on at the i'th iteration

### 4.2 Subcortical Shapes

Our 3D hippocampal shape sample was derived from the SchizConnect brain MRI data set. We used right&left hippocampal segmentations extracted with FreeSurfer [11] from 227 Schizophrenia (SCZ) patients and 496 controls (CTL). All shapes were affinely registered to the ENIGMA hippocampal shape atlas [12], and their binary masks were computed from the transformed mesh model and used as the input data. Classes were balanced to reduce bias in classification and template reconstruction. We used down sampled images in 4 time steps to reduce computational time.

It's worth to notice that the same pipeline can be applied to other subcortical structures, such as amygdala or ventricles as well as the whole brain. Out method produce individual optimal parameters for each structure since it depends on geometry of image domain.

Our initial hippocampal experiments followed those for synthetic data, optimizing only the template. We modified slightly the full optimization as described above for $(\alpha, \gamma, J)$ as follows. First, we initialized all parameters. At each subsequent step, we optimized $(\alpha, \gamma)$ and every 5 steps we updated the template $J$. Initialization was analogous to what was done in synthetic experiments. We show the initial and final template in figure 5. To compare the results, we used a simple Logistic Regression on vectorized images and KLR with Radial Basis Function (RBF) kernel. RBF-KLR appears to overfit substantially, likely because the RBF embedding space is infinite-dimensional and not in regularized in any domain-specific sense. Though LDDMM embedding is also infinite-dimensional, the fact that it is derived from image-based regularization likely explains the improved classification accuracy. Linear LR performs much better than RBF-KLR, but is still inferior to LDDMM-KLR. ROC AUC scores of the models' performance are shown in table 1.

Table 1: All results are rounded to 2 significant digits. The result is presented by the best ROCAUC score, counted on the same test set. LR was performed with a grid search to optimize the L2 regularization term weight. Registration-based model was derived as described in the experiments section, for 11 iterations.

| | left hippocampal | right hippocampal |
|---|---|---|
| Logistic Regression (LR) | 0.77 | 0.74 |
| Registration-based Optimized | 0.81 | 0.83 |

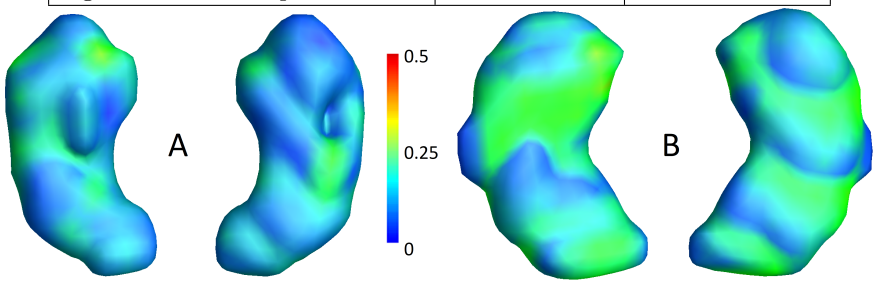


Fig. 5: Hippocampal templates. (a) Mean initial template; (2) LDDMM-KLR optimized template.

## 5 Conclusion

We have presented a novel method for diffeomorphic metric learning for image registration and predictive modeling. The unique contribution of the paper is to exploit the much-celebrated Riemannian structure of the LDDMM framework in the essence concise representation of the geometry of objects. To show the comprehensive view of our approach, we used a predictive task of practical utility. Though we have focused here on 2-class classification problems, the framework is general both with respect to the predictive modelling approach and the type of target variable. For example, rather than using diagnosis and logistic regression, we may as well have predicted the age of onset of Schizophrenia using kernel Support Vector Regression. Our only requirement is that the ML model admit kernels, and that the kernel optimization remains tractable for the given model.

A further improvement on previous work is the template optimization. Unlike previous work on template optimization [8], which focused on intrinsic statistical properties, we optimize a template with a specific predictive task in mind. In this way, we learn a metric that highlights the most essential deformations with respect to a specific biological question.

## 6 Acknowledgements

The research was conducted in the IITP RAS and solely supported by the Russian Science Foundation grant (project 17-11-01390).

## References


1. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comput. Vision **61**(2) (February 2005) 139–157
2. Mussabayeva, A., Kroshnin, A., Kurmukov, A., Denisova, Y., Shen, L., Cong, S., Wang, L., Gutman, B.A.: Image registration and predictive modeling: Learning the metric on the space of diffeomorphisms. In: International Workshop on Shape in Medical Imaging, Springer (2018) 160–168
3. Ceritoglu, C., Tang, X., Chow, M., Hadjiabadi, D., Shah, D., Brown, T., Burhanullah, M.H., Trinh, H., Hsu, J., Ament, K.A., et al.: Computational analysis of lddmm for brain mapping. Frontiers in neuroscience **7** (2013) 151
4. Cho, Y., Saul, L.K.: Kernel methods for deep learning. In: Advances in neural information processing systems. (2009) 342–350
5. Ayat, N.E., Cheriet, M., Suen, C.Y.: Optimization of the svm kernels using an empirical error minimization scheme. In: International Workshop on Support Vector Machines, Springer (2002) 354–369
6. Huang, J., Yuen, P.C., Chen, W.S., Lai, J.H.: Kernel subspace lda with optimized kernel parameters on face recognition. In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings., IEEE (2004) 327–332
7. Niethammer, M., Kwitt, R., Vialard, F.X.: Metric learning for image registration. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (June 2019)
8. Zhang, M., Singh, N., Fletcher, P.T.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In Gee, J.C., Joshi, S., Pohl, K.M., Wells, W.M., Zöllei, L., eds.: Information Processing in Medical Imaging, Berlin, Heidelberg, Springer Berlin Heidelberg (2013) 37–48
9. Gould, S., Fernando, B., Cherian, A., Anderson, P., Cruz, R.S., Guo, E.: On differentiating parameterized argmin and argmax problems with application to bi-level optimization. arXiv preprint arXiv:1607.05447 (2016)
10. Wang, L., Alpert, K.I., Calhoun, V.D., Cobia, D.J., Keator, D.B., King, M.D., Kogan, A., Landis, D., Tallis, M., Turner, M.D., Potkin, S.G., Turner, J.A., Ambite, J.L.: Schizconnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. NeuroImage **124** (2016) 1155 – 1167 Sharing the wealth: Brain Imaging Repositories in 2015.
11. Fischl, B.: Freesurfer. Neuroimage **62**(2) (2012) 774–781
12. Roshchupkin*, G.V., Gutman*, B.A., Vernooij, M.W., Jahanshad, N., Martin, N.G., Hofman, A., McMahon, K.L., van der Lee, S.J., van Duijn, C.M., de Zubicaray, G.I., Uitterlinden, A.G., Wright, M.J., Niessen, W.J., Thompson, P.M., Ikram**, M.A., Adams**, H.H.H.: Heritability of the shape of subcortical brain structures in the general population. Nature Communications **7** (2016) 13738