Robust Linear Registration of CT images using Random Regression Forests

Ender Konukoglu[a]   Antonio Criminisi[a]   Sayan Pathak[b]   Duncan Robertson[a]
Steve White[b]   David Haynor[c]   Khan Siddiqui[b]

[a] Microsoft Research Cambridge, [b] Microsoft Corporation, [c]University of Washington

ABSTRACT

Global linear registration is a necessary first step for many different tasks in medical image analysis. Comparing longitudinal studies[1], cross-modality fusion[2], and many other applications depend heavily on the success of the automatic registration. The robustness and efficiency of this step is crucial as it affects all subsequent operations. Most common techniques cast the linear registration problem as the minimization of a global energy function based on the image intensities. Although these algorithms have proved useful, their robustness in fully automated scenarios is still an open question. In fact, the optimization step often gets caught in local minima yielding unsatisfactory results. Recent algorithms constrain the space of registration parameters by exploiting implicit or explicit organ segmentations, thus increasing robustness[4,5]. In this work we propose a novel robust algorithm for automatic global linear image registration. Our method uses random regression forests to estimate posterior probability distributions for the locations of anatomical structures – represented as axis aligned bounding boxes[6]. These posterior distributions are later integrated in a global linear registration algorithm. The biggest advantage of our algorithm is that it does not require pre-defined segmentations or regions. Yet it yields robust registration results. We compare the robustness of our algorithm with that of the state of the art Elastix toolbox[7]. Validation is performed via 1464 pair-wise registrations in a database of very diverse 3D CT images. We show that our method decreases the "failure" rate of the global linear registration from 12.5% (Elastix) to only 1.9%.

Keywords: Registration, CT images, Random Forests, Mutual Information, Kullback-Leibler, Regression

1. INTRODUCTION

Global linear registration of medical images is an important step in almost every analysis involving more than a single image. Due to its importance in many different applications, the automatic image registration problem has received much attention in the literature[5,12]. Guaranteeing high levels of robustness for diverse images is critical for registration's clinical use. This work focuses on improving the robustness of automatic registration by marrying modern object recognition algorithms with efficient linear registration techniques.

The biggest challenge in the linear registration problem is the existence of local minima in the energy functions. Depending on the initialization of the registration algorithm the optimization methods can get trapped in local minima yielding unsatisfactory alignment. This problem is usually avoided by manual interaction. Here we wish to construct a fully automatic yet robust algorithm.

Leveraging knowledge about labelled regions either via explicit/implicit segmentation or semantic labelling to aid registration is a new development[3,4,5,10,11]. The key motivation is to constrain the search space of the registration parameters and thus reduce the chances of local minima entrapment. Most of the works taking this approach use pre-defined (manual) segmentations of critical organs. These segmentations are then used for constructing the energy function whose optimum point corresponds to "good" alignment of images. Works taking this approach have successfully demonstrated increased robustness over other approaches using only the image intensity information for image registration[4,5]. However, these methods demand user assisted guidance. Recent algorithms[6] proposed by our group provide efficient and fully automated probabilistic localization of anatomical structures (organ posteriors). Here, we wish to use these probabilistic localizations to alleviate the need for user guidance and achieve robust registration.

Our contribution is two-fold: i) we integrate automatically estimated organ posteriors into an efficient linear registration framework, and ii) show significant improvement in terms of robustness in the alignment of CT images. The proposed method does not require pre-defined label maps and its execution speed is similar to that of algorithms based on image intensities alone, e.g. Elastix, MedINRIA. We validate the proposed method by comparing it to the state of the art registration method provided in the Elastix package[7]. Validation is performed via 1464 pair-wise registrations in a database of very diverse 3D CT scans. We show that the proposed methods "failure" rate, 1.9%, is substantially lower than the one of Elastix, 12.5%.

2. METHOD

Our method consists of two steps. The first one is the computation of the posterior distributions for the locations of bounding boxes around specific organs such as the heart, the liver, the kidneys, the spleen, the lungs and the pelvis[6]. The main contribution of this article is in the second step where we integrate these posteriors into a robust registration process. For completeness first we describe the organ recognition process. Then we discuss the registration energy model and its minimization.

2.1 Organ Localization with Random Regression Forests

The random regression forest (RRF) algorithm[6] automatically estimates locations of bounding boxes around organs in a given 3D CT scan. Each bounding box is defined by 6 faces and the problem is to estimate the real valued coordinates of these faces, i.e. 2-x, 2-y and 2-z values for each organ of interest. A regression forest is an ensemble of regression trees trained to predict such coordinate values for all desired anatomical structures simultaneously. Training is done on a pre-defined set of CT images with associated ground truth bounding boxes. The trees cluster voxels together based on their appearance, their spatial context and, above all, their confidence in bounding box predictions. During testing, all voxels of a previously unseen image are fed into the forest and posterior distributions on the locations of the bounding boxes around selected organs are estimated. Thanks to the algorithm's parallel structure, the testing process only takes a few hundred milliseconds on a 256^3 image. Please see [6] for details.

2.2 From Bounding Box Posteriors to Voxel Probabilities

As we have stated in the introduction the robustness of registration algorithms can be improved using other information sources than only the image itself. One of the most successful approaches is to use pre-defined segmentations of anatomical structures for aligning the images[3,4,5]. The methods proposed so far have used voxel-based segmentations that are defined either through manual or automatic segmentation algorithms. This is a source of problem since manual segmentations require extensive user interaction and automatic segmentation algorithms have robustness issues themselves. The RRF algorithm, in contrast to other approaches, provides low level semantic information on the image intensity. This type of low level information can be robustly and efficiently constructed[6], providing a suitable additional information source for a fully automatic and robust registration algorithm.

The RRF algorithm gives us the estimated coordinate values for the faces of the bounding boxes along with the associated uncertainties. In order to be able to couple these posterior probability distributions of face coordinates with the image information we need to construct a probabilistic map over the individual image voxels. This probabilistic map

assigns each voxel a discrete probability distribution describing which organ that voxel might belong to. We define this distribution by using sigmoid functions as follows:

$$P(B_i|\boldsymbol{x}) = \frac{1}{Z} \prod_{j=1}^{3} \left(1 + \exp\left(\frac{x_j - u_i^j}{\sigma_i^u}\right)\right)^{-1} \left(1 + \exp\left(\frac{l_i^j - x_j}{\sigma_i^l}\right)\right)^{-1}, \qquad (1)$$

where B_i represents the ith bounding box, $\boldsymbol{x}=(x_1,x_2,x_3)$ denotes voxel position, u_i^j and l_i^j are the upper and lower faces of B_i in the j direction and σ_i^u and σ_i^l are the standard deviations for these coordinates as estimated by the RRF algorithm. Using (1) we obtain for each voxel the probability of it being inside a given organ's bounding box. Figure 1 shows examples of these maps on an abdominal CT image for different structures. Having defined the probabilistic map over the image voxels next we can couple this with appearance information coming from the image intensities and formulate the registration energy.
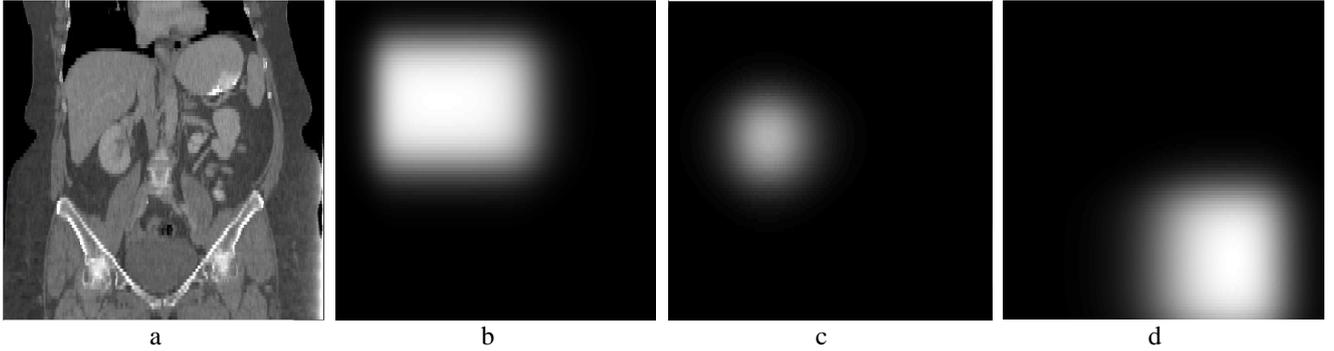

| a | b | c | d |

Figure 1. The input CT image (a) and the voxel-wise probability maps for the liver (b), right kidney (c) and the left pelvis (d). The probability maps are obtained using Equation 1. Bright corresponds to P=1 and dark is P=0. Images show 2D slices of 3D volumes.

2.3 Robust Registration with Organ Location Posteriors

In the previous section we defined the voxel-wise posterior probability maps. Among different ways of combining this with the image intensity information we choose to use a statistical energy term and formulate the registration problem as an energy minimization task as follows

$$T^* = \arg\min_T E(T), \qquad E(T) = -MI\left(I(\boldsymbol{x}),J(T(\boldsymbol{x}))\right) + KL\left(I(\boldsymbol{x}),J(T(\boldsymbol{x}))\right), \qquad (2)$$

where I and J are the two images, T is the spatial transformation (rigid-similitude-affine), MI is the mutual information defining the appearance term and KL is the Kullback-Leibler divergence acting on the anatomical location probabilities. Next we detail these terms.

2.3.1 Appearance Term.

The first term in (2) is quite conventional and it focuses on maximizing the mutual information[8] between the intensities of the two images. The mutual information between an image I and the transformed image J is:

$$MI\left(I(\boldsymbol{x}),J(T(\boldsymbol{x}))\right) = H\left(I(\boldsymbol{x})\right) + H(J(T(\boldsymbol{x}))) - H(I(\boldsymbol{x}),J(T(\boldsymbol{x}))), \qquad (3)$$

where H is the entropy. The joint entropy $H(u,v)$ is written as

$$H(u,v) = \int p(u,v)\ln(p(u,v))dudv, \qquad (4)$$

where p is the probability density function of image intensities (not to be confused with organ probabilities). In contrast to other intensity-based similarities, such as sum of square distances, this formulation assumes statistical dependency

between the values of image voxels. This relaxed hypothesis has allowed mutual information to be successfully applied in registering both multi-modal images and images with very large appearance differences[12], such as CT images with different levels of injected contrast agent.

2.3.2 Anatomical Location Posterior Term

The second term in Equation (2) is a function of the anatomical location posterior term, which is based on the voxel-based discrete probability distribution derived from the probabilistic output of regression forests as described in Section 2.2. We use the probability distributions defined in Equation (1) as

$$KL\left(I(x), J(T(x))\right) = \int_{\Omega} p(x) \sum_i P(B_i^I|x) \ln\left(\frac{P(B_i^I|x)}{P\left(B_i^I\middle|T(x)\right)}\right) dx, \tag{5}$$

where the point probability distribution $p(x)$ is assumed uniform over the image domain Ω. We note that this formulation is similar to the one used in [4] but without the requirement of pre-defined segmentations. The combination of the appearance and anatomical location prior terms gives us a better behaved minimization surface with fewer or no local minima. In Figure 2 we show cross sections of an example minimization surface demonstrating the improvement obtained by integrating the organ location posteriors. For the registration problem given in Figure 4 we computed the energy defined in Equation 2 and the negative mutual information for different translations in the x-direction (Fig. 2.a), in the y-direction (Fig. 2.b.) and in the z-direction (Fig. 2.c.). We observe that the minimization surface is much better behaved and easier to optimize. Next we provide details on the optimization scheme we use to solve Equation (2).
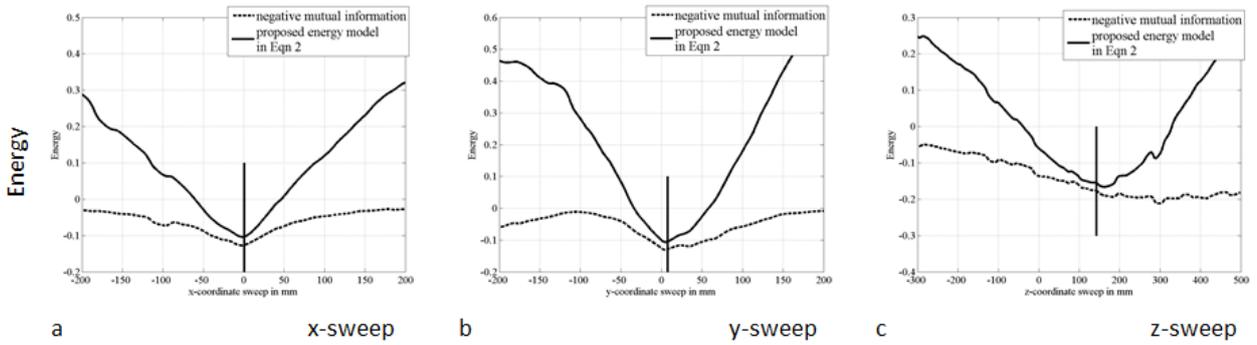

Figure 2. The optimization of the registration problem is prone to local extrema. These graphs are obtained for the registration problem given in Figure 4. In the graphs we plot the 1D cross sections of the energy function defined in Eqn 2 (in solid line) and the negative mutual information alone (in dashed line) on x-,y-, and z-translations. We see that the full energy model is less prone to local extrema. The minimum found for this problem is marked by solid vertical lines. The different registration results obtained for this problem are given in Figure 4 c (Elastix tool using only the mutual information) and 4 d (proposed method).

2.4 Optimization

The optimization of the energy function is an essential part of any registration algorithm. Considering the existence of local extrema in the optimization surface the choice of the algorithm is crucial. However, introducing the posterior distributions of organ locations in the registration algorithm makes the optimization surface well behaved. Therefore, the specific choice of the minimization algorithm becomes less critical. In order to demonstrate this point we implemented two versions of the proposed method: i) one using gradient-free global optimization scheme BOBYQA as detailed in [9] and ii) another one using a stochastic gradient descent scheme as explained in [13]. The BOBYQA algorithm is a gradient free optimization scheme which constructs successive quadratic approximations to the energy function. The optimum point is achieved through minimizing the quadratic approximations rather than the energy function itself. The second optimization scheme we have implemented is a stochastic gradient descent scheme which

uses 2nd order statistics to obtain a better descent direction at each iteration. These two optimization schemes have different characteristics and usually work well in different sets of problems.

So far we have described the proposed registration algorithm by defining the energy function and the optimization schemes. In the next section we perform thorough evaluation of the proposed algorithm and compare its registration accuracy with that of a state-of-the-art method.

3. RESULTS

The proposed algorithm has been evaluated quantitatively by running a large number of pair-wise image registrations and comparing the results with those obtained for the same image pairs via the publicly available Elastix toolbox[*7]. The MedINRIA registration tools[†] were also tested. However, they produced worse results than Elastix and in the interest of space those results are not reported.

We applied the two tools to a database of varying CT images. The database includes 180 thorax, abdominal, pelvis and full-body CT and CTA images acquired in different institutions and with different acquisition parameters. The high variability of this database provides a challenging test-bed (see Figure 4 for an example). In each image, bounding boxes around 9 structures (heart-liver-spleen-kidneys (l/r)-lungs (l/r)-pelvis (l\r)) were drawn by an expert and these boxes were used to train the RRF algorithm (not in testing though) and also to compute the registration errors. The RRF algorithm was trained on 130 images and the remaining 50 images were used for testing both registration methods. Each of these 50 test images was registered to the remaining 49 via a similarity transformation (7-DOF: 3 rotations, 3 translations and 1 scaling) using the proposed method (with both optimization schemes explained in Section 2.4) and also the Elastix tool. The parameter settings for the Elastix tool were optimized to obtain the best possible registration results. Image pairs which did not have at least one organ in common were discarded from the experiment. The computation times for both algorithms were similar despite the fact that our technique also incorporates semantic information. It took 30 seconds in average to register two 256^3 images for both methods in a multi core Intel Xeon ® machine with 6 GB of RAM.

Registration errors were computed (on the test set only) as the Euclidean distances between the ground-truth bounding boxes of the target image and those (appropriately transformed) of the transformed image. If the distance exceeded a fixed threshold that case is declared to be a "fail" for the registration. Figure 3 shows the percentage of failed cases for different thresholds. Table 1 provides numerical values for four selected thresholds (marked in figure).
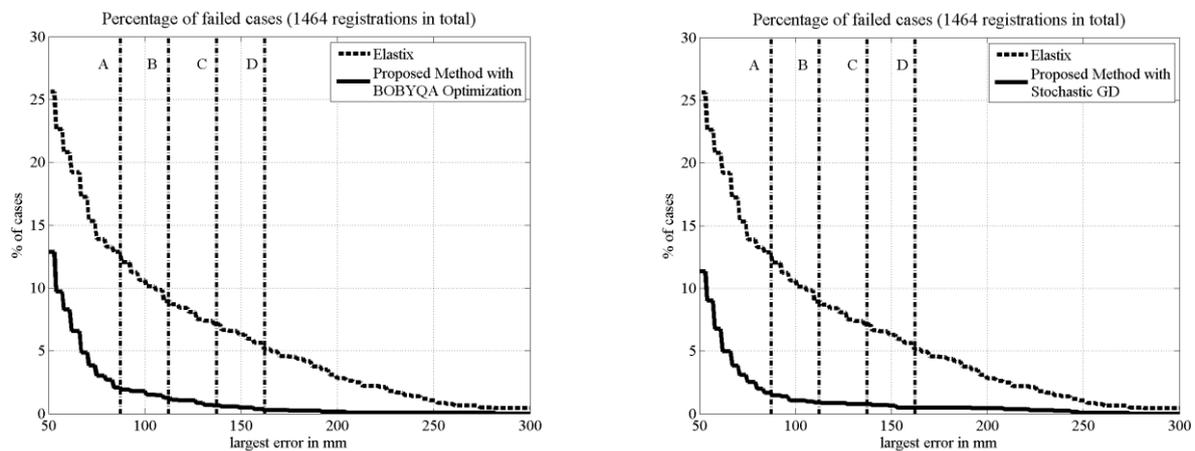

Figure 3. Percentage of failed pair-wise image registrations for different error thresholds. Left: Comparison of Elastix with the proposed method implemented using the BOBYQA optimization method, Right: Comparison of Elastix with the proposed method using Stochastic Gradient Descent.

% (#) failed cases	Elastix	Proposed Method with BOBYQA	Proposed Method with Stochastic Gradient
A	12.5% (183)	**1.9% (28)**	**1.4% (20)**
B	8.8% (129)	**1.2% (18)**	**0.9% (13)**
C	6.9% (101)	**0.5% (7)**	**0.8% (12)**
D	5.3% (78)	**0.3% (4)**	**0.5% (7)**

Table 1. Specific values from the plots in Figure 4. The thresholds are drawn in the images as vertical lines..

The results given in Figure 3 and Table 1 show the improved robustness of the proposed method. We see that the percentage of failed cases is substantially reduced, from 12.5% to 1.9 – 1.4 %. We also notice that the proposed method whether implemented using a global optimization scheme or a local gradient descent algorithm produces good results and remains robust. This shows that the optimization problem given in Equation 2 is well behaved and has very few or no local extrema. The improved robustness of the proposed method arises from the integration of the organ location posteriors in the energy.

A. MOVING IMAGE

B. TARGET IMAGE

C. REGISTRATION RESULT WITH ELASTIX

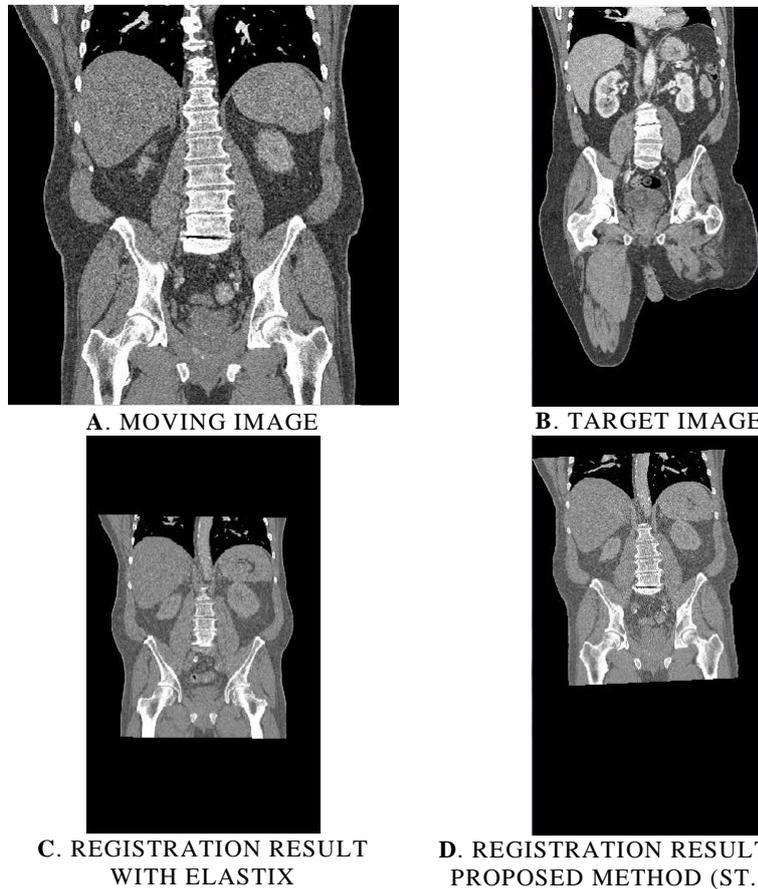**D**. REGISTRATION RESULTS WITH PROPOSED METHOD (ST. GRAD)

Figure 4. Example showing the robustness introduced using the posterior distributions for organ locations in the registration process. For this problem we registered image A to image B using similitude transformation (7 degrees of freedom). C: Methods using only the appearance information might not find the "right" alignment in patient images. D: The proposed method improves the robustness of the alignment process. Notice that image D is closer to image B compared to C.

In Figure 4 we show an example registration problem from our validation study to provide insights on why models based on only intensity information might fail while the proposed method succeeds in finding the "correct" alignment. Figure 4.A shows the moving image and B shows the target image, which is an extreme example but clinically possible. We see in Figure 4.C. that the Elastix tool fails in aligning these images using the similitude transformation while the proposed method can overcome the local extrema introduced by the amputated leg, (Figure 4.D).

4. CONCLUSIONS

We have presented a robust and fully automatic linear registration algorithm which combines semantic organ recognition and global registration. The results show that including automatically computed posterior distributions for organ locations significantly improves the robustness of linear registration while maintaining its computational efficiency. The key advantages of the proposed method are: i) it does not require pre-defined labels (segmentation maps) and ii) the execution time remains the same (30 sec for 256^3 image volume) as the registration algorithms which only uses intensity information[7] and iii) it is more robust than any automatic state-of-the-art algorithm. This makes for a practical and reliable tool.

Our future work is focused on two different directions. Firstly we would like to validate the presented results on multi-modal registration problems. Second, we would like to integrate the organ recognition within a non-linear registration framework. Such an integration would allow us to define organ specific priors on the deformation field yielding realistic-"organ preserving"- deformations for intra and inter-subject registration.

5. REFERENCES

[1] Hill, D., Batchelor, P.,et al., "Medical image registration," Phys. Med. Biol. 46, (2001).

[2] Wein, W., Brunke,et al., "Automatic CT-Ultrasound registration for diagnostic imaging and image-guided intervention," MedIA. 12, (2008)

[3] Commowick, O., Arsigny, N., et al., "An efficient locally affine framework for the smooth registration of anatomical structures," MedIA. 12, (2008).

[4] Chan, H.M., Chung, A., et al., "Multi-modal image registration by minimizing Kullback-Leibler distance between expected and observed joint class histograms." Proc. CVPR (2003).

[5] D'Agostino, E., Maes. F., et al., "An information theoretic approach for non-rigid image registration using voxel class probabilities," MedIA. 10, (2006).

[6] Criminisi, A., Shotton, J., et al., "Regression Forests for efficient anatomy detection and localization in CT studies," MICCAI-MCV Workshop, (2010).

[7] Klein, S., Staring., M., et al., "Elastix: a toolbox for intensity based medical image registration", IEEE TMI, 29, (2010).

[8] Wells, W., Viola, P., et al., "Multi-modal volume registration by maximization of mutual information," MedIA. 1 (1996).

[9] M.J.D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives," Dep. App. Math. and Th. Physics, Cambridge, England, technical report NA2009/06, (2009).

[10] Ou, Y., Davatzikos, C.: Dramms, "Deformable registration via attributed matching and mutual-saliency weighting," Proc. IPMI (2009)

[11] Pohl, K., Fisher, J., Grimson, W., Kikinis, R., Wells, W., "A Bayesian model for joint segmentation and registration," Neuroimage, 31, (2006).

[12] Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., "Mutual-information-based registration of medical images: a survey," IEEE TMI, 22,(2003).

[13] Bordes, A., Bottour, L., Gallinari,. P., "SGD-QN: Careful Quasi-Newton Stochastic Gradient Descent," JMLR, 10, (2009).