Automatic registration of the medical data for purpose of diagnosing the Alzheimer's disease

Bc. Ladislav Bari*
Supervised by: Ing. Lukáš Hudec†

Faculty of Informatics and Information Technologies
Slovak University of Technology
Bratislava / Slovak republic

Abstract

Image registration with deep learning methods has become an active field of research and an exciting area for a long-standing problems in medical imaging. However, supervised learning requires a large amount of accurately annotated corresponding control points, which can be difficult to obtain in medical imaging domain. Unsupervised learning provides us an option to get rid of manual annotation by exploiting unlabeled data without supervision. The goal is to optimize a neural network to map the appearance of input image pairs to parameters of a spatial transformation to align corresponding anatomical structures, in our case MRI brain scans of patients suffering from Alzheimer's disease. In this paper, we propose a method based on the usage of the Spatial transformer network which proved applicability for the registration task. We evaluate our method on 3D medical images from the TADPOLE dataset. We demonstrate that with the usage of affine transformations, our method outperforms the classical methods such as various registration methods provided in SimpleITK library.

Keywords: image registration, spatial transformer network, medical imaging, alzheimer's disease

1 Introduction

Medicine is an incredibly important part of the humankind. Its development, which is related to the technological progress of mankind, can be seen in better awareness of diseases, the development of new drugs and new various healing methods. Even though this progress is very pronounced, still, there are diseases that cause of origin and treatment remains a mystery to mankind. One such disease is also Alzheimer's disease.

Alzheimer's disease (AD) is a slowly developing neurodegenerative brain disease. It leads to gradual degeneration and extinction of nerve cells and associated nerve synapses between these nerve cells, especially in the parts of the brain responsible for memory and thinking. The brain shrinkage caused by this disease and the difference between the healthy patient and the patient with the AD can be seen in the figure 1.

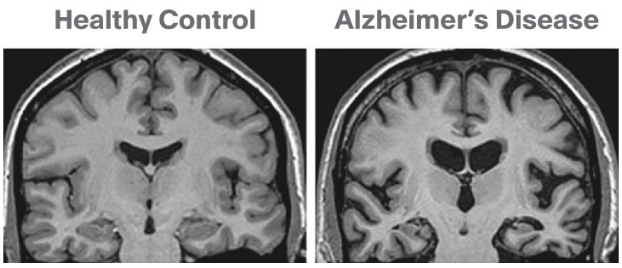

Figure 1: MRI scans of healthy patient and the patient with a Alzheimer's disease.

Significant help in medical processes is the usage of imaging devices such as MRI (magnetic resonance imaging) and CT (computed tomography). Monitoring the progress of the disease requires multiple measurements within a certain amount of time. However, this gives us the chance to gather the scans and put them under further analysis. But to make this analysis possible we need to handle the challenges such as the different positions of patients in the scanning device, the physiological changes such as brain shrinkage and also the different scanning devices. Therefore, the image registration has to be applied.

By the Zitova et. al [15], image registration, in general, is the process of overlaying/mapping two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. In the case of medical image registration, it can be used to aligning images of one subject taken at a different time (longitudinal studies) or also matching an image of one subject to some predefined coordinate system, such as anatomical atlas.

The most widely used approach in medical image registration by [14] is intensity-based registration formalized as an optimization problem seeking optimal transformation parameters Θ. The goal, by the [14], is to minimize a dissimilarity measure between fixed image and moving image undergoing a spatial transformation. Minimizing the dis-

*bari.ladislav.ml@gmail.com
†lukas.hudec@stuba.sk

similarity, or in this case, the cost function, between two images where iterative optimization strategies are used is the basic non-learning based approach.

On the other hand, deep learning-based approaches have become the active field of research. One of those approaches is learning the complex mapping from image appearance to spatial transformation. This leads us to focus on the usage of convolutional neural networks and more specifically spatial transformer network which can be used in the registration task.

2 Related works

Several existing works focused on verifying and the applicability of usage of the variations of convolutional neural network in the registration task.

In the work by Sloan et. al. [13], which is focused only on linear transformations (rigid registration), was presented as a novel method of medical image registration based on regressing the transformation parameters using a convolutional neural network. This work is also investigating the inverse consistency of the learned spatial transformations to impose additional constraints on the network during training and its possible accuracy improvement during detection.

The mono-modal experiment implements the archetype of a convolutional neural network to regress the transformation parameters. The model represents a typical structure of convolutional layers fed into a series of dense, fully connected layers, where the final layer is producing the regressed transformation parameters.

OASIS experiments	**Rotation**	**Translation X**	**Translation Y**
CNN	2.45 ± 2.78	1.66 ± 2.13	1.81 ± 2.78
FCN	$\mathbf{1.71} \pm 2.35$	$\mathbf{1.40} \pm 1.74$	$\mathbf{1.44} \pm 1.82$
FCN (ICE implicit)	2.21 ± 3.06	1.58 ± 2.09	1.70 ± 2.17
FCN (ICE explicit)	2.90 ± 3.80	1.52 ± 2.12	1.65 ± 2.20
SimpleITK	3.02 ± 5.04	18.97 ± 31.2	17.75 ± 30.26

IXI experiments	**Rotation**	**Translation X**	**Translation Y**
CNN	6.81 ± 7.85	$\mathbf{4.22} \pm 5.24$	4.66 ± 6.81
FCN	9.22 ± 11.06	4.92 ± 6.08	4.67 ± 5.61
FCN (ICE implicit)	8.80 ± 10.86	5.80 ± 7.20	4.56 ± 5.71
FCN (ICE explicit)	8.94 ± 10.66	4.80 ± 6.25	$\mathbf{4.26} \pm 5.35$
SimpleITK	$\mathbf{1.59} \pm 2.88$	21.33 ± 34.78	23.90 ± 38.91

Table 1: Mono-modal results from proposed solution. Mean absolute error and standard deviation between the measured and known transform parameters for the multi-scale iterative registration and the CNN regression methods. Rotation error is measured in degrees and translation errors are measured in pixels. Taken from [13].

The evaluation was made by comparing results of registration by proposed method and ground-truth solution, on 2 datasets to see how well are models able to generalize to different datasets. Results can be seen at the table 1. The proposed method performs well comparing to the registration method from SimpleITK library but the major failure of the method is very poor rotation regression on the first dataset, and it looks like it's prediction does not correlate with the known rotation at all. That means that the proposed method does not generalize well to subjects scanned by another scanner or scanning protocol.

Paper by Jingfan Fan et. al. [4] presents the approach for medical image registration by predicting deformation from image appearance using the fully convolutional network (FCN) that is subject to dual-guidance, first coarse guidance using deformation fields obtained by an existing registration method and second fine guidance using image similarity. The solution is based on overlapping the image patches. As a basic architecture in this paper is used U-Net architecture, with several strategies which are meant to be the addition of a better registration - hierarchical dual-supervision, gap filling, usage of multi-channel.

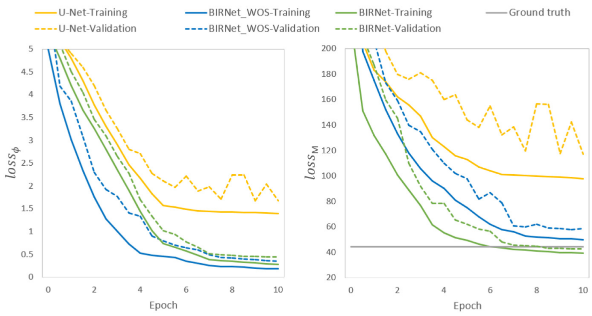

Figure 2: The training and validation curves for $loss_\Phi$ on the left and $loss_M$ on the right. The value of $loss_\Phi$ is shown in mean square error of displacement, meanwhile the value of $loss_M$ is shown in mean square error of intensity. Taken from [4].

The evaluation was performed on the LPBA40 dataset. As you can see at the table 2, U-Net saturates as training progresses, BIRNet_WOS improves the performance, but the best performance is given by the BIRNet. It also has reached better accuracy than the ground truth. Even the fact that this approach shows us, that it is fast and sufficiently accurate, it is struggling with two problems. One is that the simply transferred model needs to be refined for a new template image. This problem is caused by the recent model, which registers the subject image to a fixed template image. The second problem is that the smoothness of the predicted deformation fields is supervised by the diffeomorphic training samples. The solution for this is to add additional diffeomorphic constraints for the learning model.

In this paper by Christodoulidis et. al. [2], authors are proposing a novel convolutional neural network architecture that couples linear and deformable registration within a unified architecture endowed with a near real-time performance. Proposed framework is modular with respect to the global transformation component, as well as with respect to the similarity function while it guarantees smooth displacement fields.

Architecture is based on unsupervised CNN for the registration of pairs of medical images, in this specific case, MRI scans. CNN has two inputs source image and reference image and outputs are deformation along with the registered source image. One of the main components of the proposed CNN is the 3D transformer layer. This layer, which is part of the CNN is used to warp its input under a deformation. The whole architecture of the CNN is based on an encoder-decoder framework.

Method	Inhale-Exhale	All Combinations	Time/subjects (s)
Unregistered	75.62 ± 10.89	57.22 ± 12.90	-
Deformable with NCC	84.25 ± 6.89	76.10 ± 7.92	~ 1 (GPU)
Deformable with DWM	88.63 ± 4.67	75.92 ± 8.81	~ 2 (GPU)
Deformable with MI	88.86 ± 5.13	76.33 ± 8.74	~ 2 (GPU)
Deformable with all above	88.81 ± 5.85	78.71 ± 8.56	~ 2 (GPU)
SyN	83.86 ± 6.04	-	~ 2500 (GPU)
Proposed w/o Affine	91.28 ± 2.47	81.75 ± 7.88	~ 0.5 (GPU)
Proposed	91.48 ± 2.33	82.34 ± 7.68	~ 0.5 (GPU)

Table 2: Dice coefficient scores (%) calculated over the deformed lung masks and the ground truth. Taken from [2].

Evaluation shows, that proposed method generates deformations with no self-crossings due to the way the deformations layer is defined. Network was tested and evaluated on the dataset of lung's MRI scans, but by the authors, this approach can be used also for the registration of other organs. Final score can be seen at the table 2.

Experiment	Method	Jac	d_H	Running Time
-	No registration	0.491	2.270	-
E_p	SimpleITK [6, 7]	0.561	1.876	166.537±89.844
	AIRNet-$S_{1,2,c}$	**0.613**	**1.431**	0.785±0.027
	AIRNet-$S_{1,2}$	0.586	1.546	0.778±0.036
	AIRNet-$D_{1,2}$	0.599	1.512	0738±0.016
$E_{p,m}$	SimpleITK [6, 7]	0.524	1.978	216.848±114.369
	AIRNet-$S_{1,2,c}$	**0.615**	**1.459**	0.793±0.034
	AIRNet-$S_{1,2}$	0.577	1.574	0.772±0.028
	AIRNet-$D_{1,2}$	0.601	1.475	0.755±0.022

Table 3: Performance of the conventional methods and the proposed framework when evaluated on the temporal lobe. The metrics are 1) mean Jaccard index across all subjects, 2) mean modified Hausdorff distance across all subjects, and 3) average and standard deviation of running time per registration. For the two respective experiments, the rows list results before registration, with registration using SimpleITK, and results obtained using our three trained models. The best results are highlighted in bold. Taken from [1].

In this paper by Evelyn Chee et. al. [1] is proposed a self-supervised learning method for affine image registration on 3D medical images. Affine image registration network (AIRnet) used in this work, is unlike the optimization-based methods, designed to directly estimate the transformation parameters between input images without using any metric. AIRnet enables to learn the discriminative features of the images which are useful for registration purpose. Architecture is based on the two main components which are encoder and decoder.

The evaluation shows very good performance compared to a conventional image registration method such as SimpleITK and also shorter execution time, which can be seen in figure 3. The architecture of the AirNet enables to find a good feature representation of the images which provides a summarizing description of the complex morphological patterns of the images regardless of the modality of the input images.

3 Method

Our main goal is to propose a novel approach to improve the registration process of the sequential MRI scans, created within a certain amount of time. As the neural networks have proven themselves profitable in multiple application domains, we want to examine their capabilities in the field of medical image registration. The Neural networks are robust and with sufficient training, they should overcome the results of the classical methods. We want to register mainly the raw data and then evaluate how the results differ in comparing the classical methods and our method. We also want to evaluate what are the advantages and disadvantages of our approach and its performance.

In figure 3, there is a simplified workflow of the proposed registration method. In the beginning, there are two MRI scans, which should be registered. One of them is a fixed image and the other one is a moving image, which should be registered to the fixed image. The next step is the pre-processing part, where the images are being prepared to move to the next step which is the registration network. In our case, the registration network contains the spatial transformer network. After successful training, we have the trained model, which is used to warp (register) the moving image.

Spatial Transformer Network [5] is the building block of the registration network. It is consisting of a few convolutional and fully connected layers that can learn a mapping from input images to parameters Θ of a spatial transformation T_Θ. In this work, we decided to use the mean squared error (MSE) [8] as the loss function.

Before the images from the dataset can be processed in the neural network, there is a need for pre-processing, which can help during the training of the network. Preprocessing in our case consists of 3 very simple steps which are:

1. **Normalization** - min-max normalization
2. **Resampling** - downsampling to $(170x170x170)$
3. **Initial alignment** - Centred Transform Initializer - we are using it to align the centers of the two volumes and set the center of rotation to the center of the fixed image. This alignment is executed by the method from the SimpleITK library.

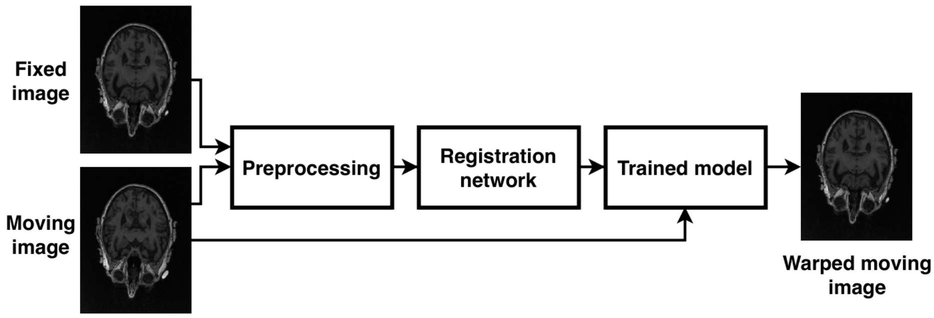

Figure 3: Workflow of the proposed registration approach

As was mentioned above, the resampling especially downsampling of the input images was applied. The main reason to do that is to handle one of the challenges of medical registration - usage of different scanning devices. The performed analysis of the dataset showed that the scanning device was changed during the time in which the images were scanned. It means that the images from different times have different resolutions, from 170 to 256. This brought us to perform downsample operation leading to the unification of the resolution of the images. The resampling itself was performed by setting the needed spacing, which was multiplied with the current sizes in each dimension and then divided by the desired size in each dimension to get the resampling spacing needed. The value of 170 was chosen as the starting value with a good ratio of the level of details and image size.

Another pre-processing operation was initial alignment. This operation was used to prevent the situation where the centers of the images are not the same. As was mentioned before, we used the method from the Simple ITK.

3.1 Network architecture

Let F, M be two image volumes defined over a $n-D$ spatial domain $\Omega \subset R^n$. In the case of 3D registration, we focus on the case $n = 3$. We assume that F and M contain single-channel, grayscale data and also assume, that they were successfully pre-processed as was said in the previous section.

Our network takes M and F as an input and computes the registration field based on a set of parameters Θ, the kernels of the convolutional layers. We warp the moving image using a spatial transformation function, enabling the model to evaluate the similarity of this warped image and fixed image. We use batch gradient descent, minimizing the expected loss over a training set. As a loss function, we decided to use the Mean Square Error (MSE) loss function. To train our network, we do not require supervised information such as ground truth registration fields or anatomical landmarks. We learn appropriate values for Θ by training to align a dataset of volume pairs from a population.

In our experiments, input is of size $(170 \times 170 \times 170 \times 2)$. As the size of the input shows, the input contains 2 scans, which should be registered. We apply 3D convolutions with a kernel size of $(5 \times 5 \times 5)$. The convolutional layers capture hierarchical features of the input image pair necessary to estimate the correspondence. At least we apply transformations to the incoming data.

We created two architectures of the network which differs in the way of work with the transformation parameters. The first network which can be seen in the figure 4, can be considered as a basic STN [5], where the localization network takes the input feature map and through several hidden layers outputs the parameters of the spatial transformation that should be applied to the feature map. In this case, the affine transformation matrix is completely created by the network itself.

However, in the second architecture which can be seen in the figure 5, the network outputs the parameters for each transformation separately. This allows us to have more control above the transformation matrixes and avoid the cases when applied spatial transformation transforms the image to one point. Also, we can decide which transformations will be applied to the image. Available transformations are translation, rotation, scaling, and shearing. Although this approach gives us freedom and bigger control over the transformation, it also brings one disadvantage that arises from the fact that the order of the matrixes is not interchangeable. Analysis of the results of the different combinations showed us that the ideal order of the transformations is shear, scale, rotation, translation.

We implemented our network using Pytorch. As an optimizer, we used the ADAM optimizer [6] with a starting learning rate of 0,001. We used learning rate decay, which decreases the learning rate after 40 epochs by 10%. The size of the training batch was set to 4 and the number of epochs was set to 100. Images were resampled to size of $(170 \times 170 \times 170)$ with spacing set to $(1.5 \times 1.5 \times 1.5)$.

4 Experiments

Testing of the developed model had different iterations to find out the ability to learn the different types of transformations on different derivates of the dataset. For the experiment purposes, we decided to use the dataset from The Alzheimer's Disease Prediction Of Longitudinal Evo-

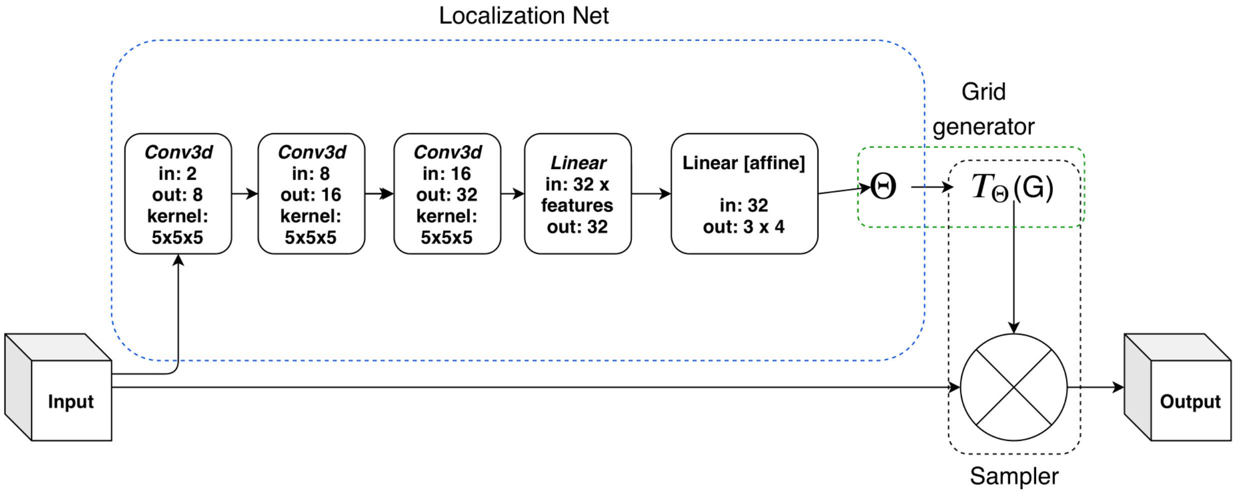

Figure 4: Architecture of the network with the one affine transformation matrix

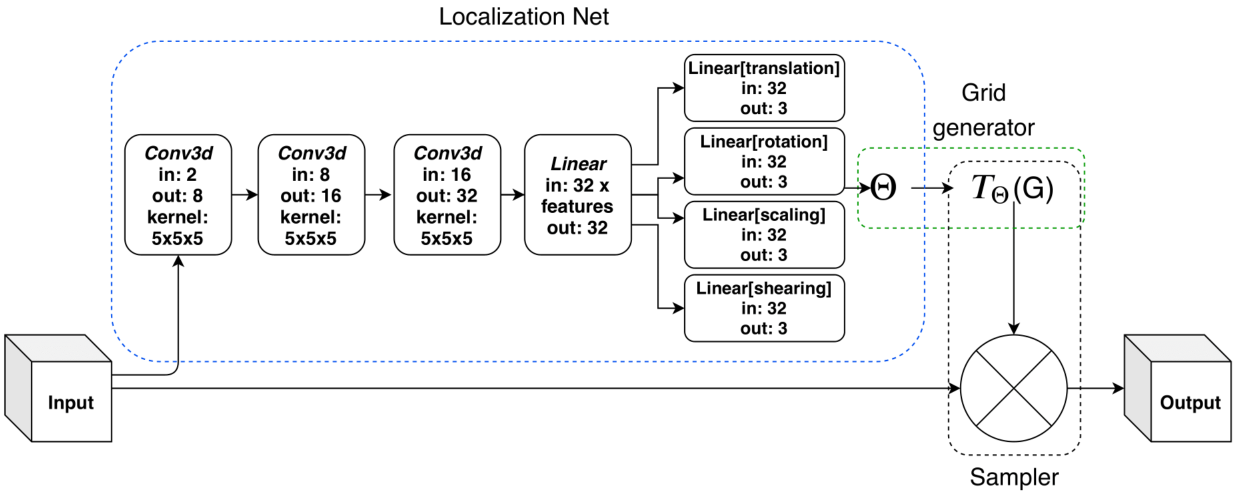

Figure 5: Architecture of the network with the transformation matrix for each transformation

lution (TADPOLE) Challenge. Dataset used in this challenge is a standard derived dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI). This dataset contains over 7 thousand MRI scans obtained from several hundred patients with Alzheimer's disease. All scans are T1-weighted. The main advantage of this dataset is that all scans were captured over a long period of time which brings challenges such as different positions, physiological changes, and different scanning devices.

Initial testing started with a synthetic 3D dataset [9] represented by the pairs of cubes. Every image has a size of $(128 \times 128 \times 128)$ and is randomly translated, rotated and also scaled. Every cube has structural impairment represented by cracks and holes. Testing shows not even the ability to register the images but also to show that the network is capable of dealing with the structural impairments. Data was represented by 1000 pairs for training, 100 pairs for validation and 100 pairs for testing. For training purpose images were downsampled to the size of $(64 \times 64 \times 64)$.

To better understand how can network handle the different transformations, we created the dataset with 3 different transformations:

- **Rotated dataset** - only rotated 3D images
- **Translated dataset** - only translated 3D images
- **Rotated-translated dataset** - a combination of two previous, translated and rotated 3D images

In all 3 datasets, we are using a 3D MRI scan from TADPOLE dataset [11]. To focus only on transformation, the dataset is created by transforming only one MRI scan, which shows to be the cleanest, without the need for further pre-processing like for example denoising. For every dataset were created 1000 pairs for training, 500 pairs for validation and 1000 pairs for testing.

To show how the network is capable of registering real-world dataset, we created two different datasets:

- **TADPOLE random** - random selection of images
- **TADPOLE longitudinal** - pair of images from one patient

In the first dataset - TADPOLE random, as the name implies we randomly chose 1000 pairs of scans for training, 500 pairs for validation and 1000 pairs for testing purposes. The goal is to show how the can model handles the registration of scans belonging to a different patient so every pair is composed of scans from two different patients. The goal of the usage of the second dataset - TADPOLE longitudinal, is to show the use case of the model, where we want to register several scans of one patient to one of this patient's scans. Every pair is composed of scans from one patient. Also as in the previous dataset, 1000 pairs of scans for training, 500 pairs for validation and 1000 pairs for testing purposes were chosen.

5 Evaluation

For the evaluation purposes, we used three different metrics. First and the most simple metric is the Dice coefficient score. The value represents the overlap between images, whereas 0 means no overlap and 1 means perfect overlap. However, this metric is very simple and also only binary what is its biggest disadvantage. Nevertheless, it was used mainly as a comparison to the other methods and works which widely use this metric. A second used metric is Hauffsdorf distance [12], which measures how far two subsets of a metric space are from each other. The last and the most sophisticated metric which is suitable also for the non-rigid registration is Mutual information (MI) [3]. MI is a measure of the mutual dependence between the two variables. It quantifies the amount of information obtained about one random variable by observing the other random variable.

We have evaluated our registration approach and compared it with the registration method from the Simple ITK library [10].

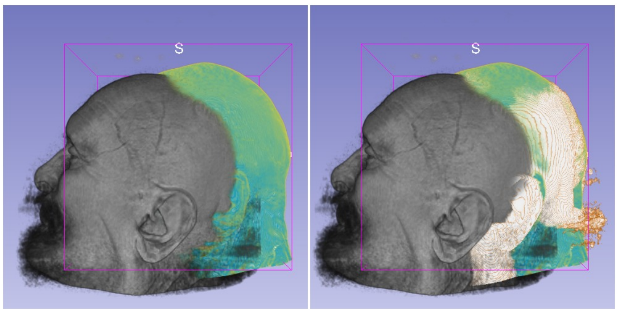

Figure 6: Example of images from tadpole dataset. On the left overlap of moving (gray) and target image (blue). On the right overlap of moving (gray), target (blue) and registered image (orange) via 3D STN with affine transformation model.

Resulting values of the dice coefficient score can be seen at the table 4. We also have introduced the dice coefficient for datasets with no registration applied, to see how the registration method improved the overlap between images in datasets. To have a proper evaluation, two registration models were used. First is the model with the separate matrixes for each transformation and the other is the model where we have only one affine matrix. Finally, in comparison to the classic registration method, we decided to evaluate the registration method from Simple ITK library.

Resulting values of another metric - Hauffsdorf distance can be seen at the table 5. As you can see results confirmed the statement, that our proposed method can especially in the real world TADPOLE dataset outperform the classical registration method, in this case, the registration method from Simple ITK library.

Resulting values of Mutual Information metric can be seen at the table 6 which is showing that the usage of our 3D STN affine model with the separate matrixes for each transformation outperforms the other registration methods in both real-world datasets.

Results from all 3 metrics are showing us that our proposed models can outperform the existing registration method from the Simple ITK library. The most important results are those from real-world datasets - TADPOLE random and TADPOLE longitudinal. If we look at all metrics, we can see that our models are better than the registration method from Simple ITK. However, an interesting thing is that while in the dice score model with separate matrixes in better in the random dataset and the model with one matrix is better in the longitudinal dataset. But in the case of Hauffsdorf distance, the situation is vice versa and in the case of Mutual Information, best results are provided by the model with separate matrixes.

6 Future work

Many different tests and experiments have been left for the future due to lack of time (i.e. the experiments with real data are usually very time consuming). Future work concerns a deeper analysis of particular mechanisms and new proposals to try different methods.

Since the proposed image registration method is focused on rigid registration, it could be interesting to consider the usage of the non-rigid registration. This mechanism can bring us interesting results that can be compared to the existing results gained by rigid registration. However, it is advisable to consider when to use this method appropriately due to the deformation of the images. In future work, we want to focus on the non-rigid registration using B-Spline [7] which is popular due to its general applicability, transparency and also computational efficiency. The main disadvantage of this approach is that special measures are required to prevent the folding of the deformational field which is more difficult to enforce at a finer resolution.

Method/Dataset	Synthetic	Rotated	Translated	Transl.-rotated	TADPOLE random	TADPOLE longitudinal
No registration	0,764	0,880	0,787	0.821	0,674	0,751
3D STN affine separate matrixes	**0,937**	**0,960**	0,843	0,878	**0,791**	0,825
3D STN affine one matrix	0,889	0,891	0,814	0,855	0,758	**0,837**
Simple ITK	0,820	0,893	**0,860**	**0,887**	0,772	0,816

Table 4: Dice coefficient scores calculated over the datasets using different registration methods.

Method/Dataset	Synthetic	Rotated	Translated	Transl.-rotated	TADPOLE random	TADPOLE longitudinal
3D STN affine separate matrixes	**17,68**	**33,15**	**37,15**	40,13	70,98	**63,27**
3D STN affine one matrix	18,13	39,41	39,75	41,61	**65,23**	63,73
Simple ITK	19,43	64,40	43,14	**34,95**	74,20	67,39

Table 5: Hauffsdorf distance calculated over datasets using different registration methods.

Method/Dataset	TADPOLE random	TADPOLE longitudinal
3D STN affine separate matrixes	**0,344**	**0,547**
3D STN affine one matrix	0,319	0,538
Simple ITK	0,331	0,535

Table 6: Mutual information calculated over datasets using different registration methods.

7 Conclusion

In this paper, we propose a novel method which is using a convolutional neural network with the spatial transformer network trained in an unsupervised manner for 3D medical image registration. After designing suitable architectures, one that uses adjusted STN giving the parameters for all the defined transformations separately and the other one which uses the one transformation matrix obtained from the network, we implemented our method and subsequently tested it on five datasets. The results show that our models have very balanced results compared to the registration method from Simple ITK library, but it surpasses it in both real-world datasets It should be noted, however, that the network has not yet been optimized, which could increase its performance. Also, the next goal of future work is to create and implement a model based on a non-rigid registration approach using B-Spline techniques.

References

[1] Evelyn Chee and Zhenzhou Wu. Airnet: Self-supervised affine registration for 3d medical images using neural networks. *CoRR*, abs/1810.02583, 2018.

[2] Stergios Christodoulidis, Mihir Sahasrabudhe, Maria Vakalopoulou, Guillaume Chassagnon, Marie-Pierre Revel, Stavroula G. Mougiakakou, and Nikos Paragios. Linear and deformable image registration with 3d convolutional neural networks. *CoRR*, abs/1809.06226, 2018.

[3] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006.

[4] Jingfan Fan, Xiaohuan Cao, Pew-Thian Yap, and Dinggang Shen. Birnet: Brain image registration using dual-supervised fully convolutional networks. *CoRR*, abs/1802.04692, 2018.

[5] Max Jaderberg, Karen Simonyan, Andrew Zisserman, and koray kavukcuoglu. Spatial transformer networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 2017–2025. Curran Associates, Inc., 2015.

[6] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations*, 12 2014.

[7] G.D. Knott. *Interpolating Cubic Splines*. Interpolating Cubic Splines. Birkhauser, 2000.

[8] Erich L. Lehmann and George Casella. *Theory of Point Estimation*. Springer-Verlag, New York, NY, USA, second edition, 1998.

[9] Imperial College London. The research computing service's data repository.

[10] Bradley Lowekamp, David Chen, Luis Ibanez, and Daniel Blezek. The design of simpleitk. *Frontiers in neuroinformatics*, 7:45, 12 2013.

[11] Razvan V. Marinescu, Neil P. Oxtoby, Alexandra L. Young, Esther E. Bron, Arthur W. Toga, Michael W. Weiner, Frederik Barkhof, Nick C. Fox, Stefan Klein, Daniel C. Alexand er, the EuroPOND Consortium, and for the Alzheimer's Disease Neuroimaging Initiative. TADPOLE Challenge: Prediction of Longitudinal Evolution in Alzheimer's Disease. *arXiv e-prints*, May 2018.

[12] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*. Springer Verlag, Heidelberg, Berlin, New York, 1998.

[13] J. M. Sloan., K. A. Goatman., and J. P. Siebert. Learning rigid image registration - utilizing convolutional neural networks for medical image registration. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: BIOIMAGING,*, pages 89–99. INSTICC, SciTePress, 2018.

[14] Guoli Song, Jianda Han, Yiwen Zhao, Zheng Wang, and Huibin Du. A review on medical image registration as an optimization problem. *Current Medical Imaging Reviews*, 13, 07 2017.

[15] Barbara Zitová and Jan Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21:977–1000, 10 2003.