Multimodal Brain MRI Registration Using Generative Adversarial Networks

Norbert Vígh*
Supervised by: Vanda Benešová†

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

Abstract

Deep learning methods have recently found applications in several fields, including the processing of medical imaging data. We explore the application of convolutional neural networks (CNN) for automatically processing magnetic resonance imaging (MRI) scans in ceT1- and T2-weighted modalities to assist doctors with executing accurate and time-efficient tumor diagnostics.

The main challenge of the work is the multimodal registration of coronal and axial scans, which are perpendicular to each other and, therefore, cannot be registered directly. We use a generative adversarial network (GAN) architecture to convert between modalities, making it easier to register them. The resulting registered scans can be used for a wide variety of further tasks, utilizing the complementary information contained in different MRI modalities, i.e. image segmentation.

Keywords: Multimodal Image Registration, Deep Neural Networks, Generative Adversarial Networks, Medical Imaging

1 Introduction

In recent years, there has been a substantial rise in novel techniques used for diagnostics in medicine. The introduction of artificial intelligence (AI) can speed up most of the tasks that doctors perform daily.

Despite legal and ethical challenges, it can still greatly aid the doctor. The most influential advantage of using AI is the reduced time it takes to diagnose a patient since, in general, the AI can process more data quicker than even an experienced doctor could.

One of the main problems in the learning process of AI is the need for a lot of data. Many tasks require using annotated data that must first be created by a domain expert, which is very tedious. By creating an intelligent assistant based on a neural network, we could simplify the process of data annotation and thus help create larger datasets available for future usage.

In many cases, doctors need to work with multiple modalities, providing complementary information, sometimes even acquired by using several different diagnostic devices. Due to this, we see a new research possibility in intermodal conversion. Although intermodal conversion is not able to synthesize information that is not present in the original imaging data, it can be sometimes useful, for example for our presented co-registration method. By creating a tool that could convert images between modalities, we could reduce the time needed to scan the patient multiple times, thus enormously reducing the waiting times to get diagnosed.

To achieve this, we first need to be able to **register modalities** correctly in the same space. Afterward, we can add the provided labels from one modality to the second modality. This enables the creation of neural networks that could be trained on a compacted dataset. This could possibly create a tool that could classify into classes that are not that visible in the actual modality provided to the doctor, but the neural network could pick up on this.

Our contribution lies in creating a **multimodal 3D image registration tool** using a GAN network. The core problem we try to overcome is the perpendicularity of the available dataset, which makes this problem even harder. Unfortunately, most real-world MRI data uses an anisotropic voxel grid. The spacing of the scans is different on the third axis, which motivates the creation of a custom-made image registration algorithm. To make the problem more graspable, we reduce the complexity of registering two modalities by introducing a GAN network that converts images between MRI modalities, making the registration process more straightforward.

2 Background

Pituitary adenomas (PA) are a type of benign tumor affecting the pituitary gland, with a prevalence of 96 ± 20 cases per 100,000 population [8]. Despite being benign, PA can exert pressure on surrounding tissues, necessitating cau-

*xvighn@stuba.sk
†vanda_benesova@stuba.sk

tious surgical intervention or they can cause abnormal hormonal production [18]. A precise diagnosis and characterization of PA is crucial for treatment planning and patient care. The segmentation mask can be used for further statistical analysis of the tumor, such as classification based on its diameter [16, 4] and shape relative to surrounding tissues [12], or radiomics analysis, which extracts quantitative features from medical images to predict patient outcomes [9, 15].

Medical imaging plays an essential role in tumor diagnosis. Techniques such as X-rays, CT scans, MRI, and ultrasound provide detailed scans for analysis. MRI scans, in particular, offer high-resolution images acquired in different planes, allowing for better visualization of anatomical structures and abnormalities. Different MRI modalities, like T1-weighted and T2-weighted, offer varying tissue contrast and information about tissue composition and structure, aiding in the detection and characterization of tumors [16, 4, 3].

Computer vision techniques are increasingly employed in medical image processing for tasks such as image registration and segmentation. Image registration aligns images from different modalities or time points, enabling accurate comparison and analysis. Segmentation identifies and labels regions of interest within images, facilitating quantitative analysis and diagnosis. Evaluation metrics such as mutual information and the Dice similarity coefficient (DSC) assess the accuracy and performance of these techniques, guiding their optimization and application in clinical settings [14].

The rise of deep neural networks, particularly convolutional neural networks (CNNs), has revolutionized medical image analysis by enabling automated feature extraction and classification. Architectures like U-net [17] excel in segmentation tasks, accurately delineating tumor boundaries and aiding in treatment planning. Generative models like variational autoencoders (VAEs) [11] and generative adversarial networks (GANs) [10] offer novel approaches for data generation and augmentation, enhancing the availability and diversity of training data for deep learning models.

2.1 Related Work

The fundamental problem with multimodal registration is that most models try to find a good mapping between the different intensities in the fixed and moving images. Depending on the actual modalities, it might be problematic to find such mappings in some cases. However, with the recent advancements in CNNs, there have been attempts to **facilitate the registration process by applying segmentation to both modalities** and subsequent image registration in the space of the segmented images.

One such work was published by Blendowski et al. [2] The authors suggest using a **convolutional autoencoder** architecture to extract shape features of both modalities. Using the encoder part on the input volume, they extract a 1584-dimensional shape space describing the objects in the volume. They propose that by applying linear interpolations between the moving and fixed image encodings, they can achieve iterative guidance of the image registration process. This should help the registration process by eliminating large non-linear deformations that could occur when the algorithm tries to register the moving image directly onto the fixed image. Their results confirm this, since when applying direct registration, they only achieved a DSC of 0.526, while when using the approach with iterative guidance, they achieved a DSC of 0.653 [2].

Another work by Cao et al. [5] is focused on the registration of CT and MRI using models that can **synthesize images of either modality** from the other image. The subsequent image registration is then much more manageable. However, especially synthesizing MRI from CT is a complicated, non-linear task. The authors do not use a neural network but rather a Multi-Target Regression Forest since, this way, they arguably needed less training data. The forest synthesizes both modalities, so the registration algorithm has two images (one original and one synthesized) in both modalities. The calculations can then be done in both modalities, where one way of transformation is inverse to the other way. The paper by Cao et al. [5] was used for MRI and CT pelvic image registration. However, there have been other papers about synthesizing other modalities for brain-related data, even by using neural networks. In a paper by Li et al. [13], the authors try synthesizing MRI from CT of brains using deep learning methods. They compare approaches based on a CycleGAN, Pix2Pix (conditional GAN) model, and the U-net architecture. Unexpectedly, the best results were achieved using the U-net architecture with L1 and L2 regularizations [13].

In a paper by Zheng et al. [20], the authors propose a method for multimodal image registration using a GAN network. Compared to the CycleGAN model used for image synthesis in the previous section, the authors propose a **Symmetric registration GAN model**, which also creates a cycle-consistent mapping between the two modalities. However, a transformation is also applied to the images, which transforms them into the same space as an image from the other modality. This way, even though the images are not aligned, they can be directly compared, and the loss can be calculated more precisely [20]. The transformation is done using Affine Transformation Regressors, which try to predict the affine transformation parameters between the two images they receive as input. There are also non-linear transformation regressors, which try to predict the non-linear transformation parameters using a VoxelMorph model [1]. These regressors are pre-trained on a dataset of artificial transformations with known parameters. Ultimately, the proposed model generates two symmetric transformations, which can be applied to register the images in either way [20].

3 Our solution

The main contribution of our work lies in the way we use data from multiple modalities by creating a registration algorithm that could enable the usage of labels of classes from multiple modalities, which are not greatly visible on only a scan from one modality. A neural network can still detect such cases, meaning full diagnostics could be possible by providing only one modality.
We have the following goals:

- to help the process of creation of new more robust datasets,
- to register MRI scans in multiple modalities, which could generate a dataset with labels from both modalities,
- to prove the possibility of precise segmentation of classes even if provided with an input consisting of only one modality,
- facilitate further research in segmentation using only one modality by providing a tool to create a dataset with labels from both modalities.

In summary, we aim to aid the diagnostics of pituitary adenoma by providing a tool that should enable the creation of more extensive datasets and incite more research in this field.

To make the next sections more followable, we introduce a notation of different kinds of scans. Letters C and A show whether the scans are coronal or axial respectively. These letters can be followed by an apostrophe $'$, indicating that the image was crated by inference through the GAN network.

3.1 Dataset

We obtained the dataset based on cooperation with a doctor from the Military University Hospital in Prague, It consists of Axial T2-weighted (A_{T2}) and Coronal contrast-enhanced T1-weighted (ceT1, C_{ceT1}) MRI scans in the Nifti format. The dataset is anonymized and contains no personal information about the patients.

In total, we have MRIs of 928 patients. However, the label masks in both modalities are only available for 330 patients, which makes the base of our dataset. The other 340 patients scanned in both modalities have annotations on neither (or just one) of the modalities. Annotating many MRI scans is very time-consuming, so even in the annotated samples, only a handful of 2D scans are annotated, and others are left untouched.

Axial scans are scans in the horizontal plane. They consist of 22 to 52 cross-sections with resolutions between a quarter and a half millimeter. On the contrary, coronal images are cross-sectional and consist of 12 to 24 cross-sections with a resolution of half to one millimeter.

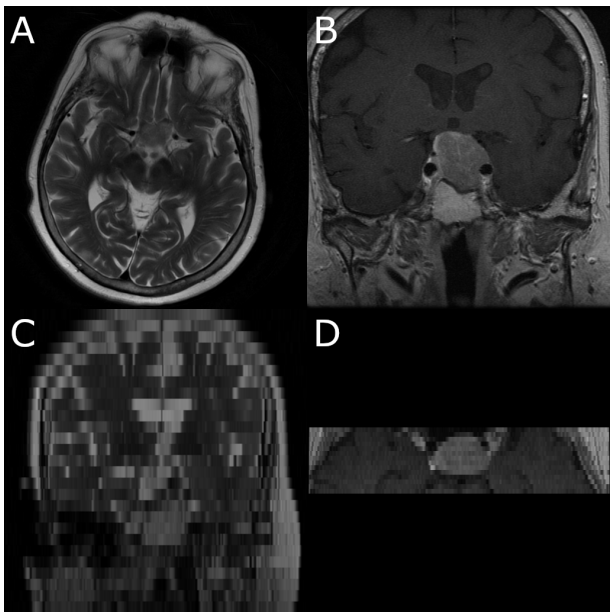

Figure 1: Comparison of axial (A) and coronal (B) scans. Axial scans shown in coronal plane (C) and coronal scans shown in axial plane (D)

There are several crucial problems with the dataset that we need to solve:

- The third axis of the resolution is always significantly worse and can be around 2 to 5 millimeters. Doctors only create a few slices, showing key brain areas that must be examined. This makes the MRI sampling faster, more convenient for the patient, and more economically feasible. Minimizing the examination time can also decrease the spatial shift between subsequent scans caused by patient movement.
- Another problem is that the axial and coronal slides may not be precisely perpendicular. Therefore, no clear transformation is available that could align these two modalities, and thus, it has to be calculated using image registration methods.

Label masks are available for both modalities; however, the classes are not equivalent. It is impossible to distinguish some tissue types with certainty in the respective modalities. The modality will be selected based on the tissue type the doctors want to examine. For example, arteries marked on ceT1-weighted scans may be poorly distinguishable from the surrounding tissue on T2-weighted scans.

In addition, we use another dataset of registered coronal MRI scans. This dataset consists of scans of 1157 patients, all of which were scanned in the coronal plane, half of which are contrast-enhanced T1-weighted and the other half T2-weighted slices. In total, this represents 10802 slices of each modality. Each sample is paired with a corresponding sample from the other modality, meaning that each patient was scanned in both modalities.

3.2 Challenges

There are several problems that we need to face:

- The moving image are just approximately perpendicular to the fixed image (can differ up to 5 degrees).
- The spacing of A_{T2} is too large, so we cannot interpolate them to an isotropic space, instead we can interpolate C_{ceT1} scans with a smaller spacing.
- C_{ceT1} scans stick out of the space of the A_{T2} scans, necessitating the need to find the correct subset of C_{ceT1}.

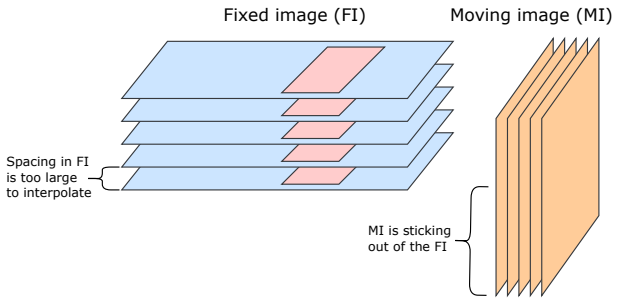

Figure 2: The core problems of the used dataset. (blue) axial slices, (orange) coronal slices, (pink) expected transformation of coronal slices into the space of axial slices.

3.3 Proposed pipeline

Figure 3 shows an overview of the proposed pipeline for processing the data.

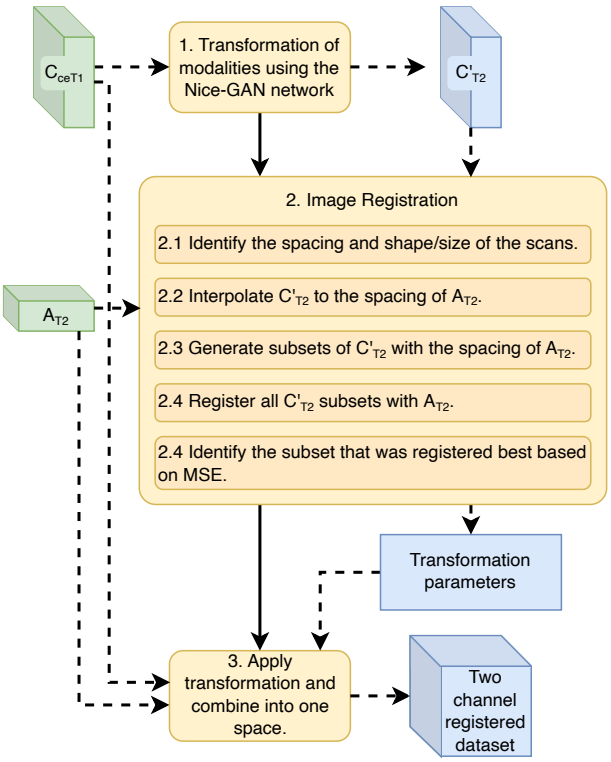

Figure 3: The data-flow chart of the processing pipeline. Green blocks show data that is used as input, blue blocks are data generated by the pipeline.

3.4 GAN network for intermodal conversion

The first step of our pipeline is to convert one of the modalities into the other modality.

We use the Nice-GAN architecture [7, 6], an improved version of the CycleGAN architecture. Compared to the CycleGAN, this architecture reuses the results from the Encoder part of the generator, which are then further processed in the Discriminator network. This makes the training process more stable and allows the generator to learn more complex mappings [7].

Moreover, this architecture tries to get to the same hidden vector (latent space) from both modalities, which makes the mapping more consistent. This is achieved by using a shared latent space, which is then used to reconstruct the original image since they can switch the decoders to generate the other modality. So, there are two loss functions. The first one is the reconstruction loss, which is calculated by comparing the original and reconstructed images of the same modality. The other loss is the cycle loss, calculated by comparing the original image and the image obtained by converting the original image to the other modality and then back to the original modality [7].

We train this network on the datatest mentioned in the last paragraph of Section 3.1.

The training was done on a machine with an NVIDIA RTX4090 desktop GPU with 24GB of VRAM. It ran for 40,000 iterations and took about 10 hours to complete. The results were saved after every 10,000 iterations, and visualization was generated for the intermediate results after every 1000 iterations. The learning rate was set to $1 * 10^{-4}$ with a batch size 1 and Adam optimizer.

We used the results of our GAN to convert C_{ceT1} into C'_{T2} in a slide-by-slide manner.

3.5 Our registration algorithm

The steps of the registration algorithm are indicated in Fig. 3. In this section we describe these steps in detail.

3.5.1 Transformation of coronal to axial slices

The images were first loaded to read the Nifti files, including their metadata. The metadata includes information about the spacing of the scans, which is crucial for the registration process. Based on the spacing, we were primarily interested in the third axis, which had the largest spacing.

The C'_{T2} scans were then interpolated to the same spacing as the A_{T2} scans. The interpolation was done by calculating the expected slice dimensions based on the spacing of A_{T2} scans and then interpolating C'_{T2} scans to the same spacing.

We have created a helper function to rotate C'_{T2} scans by a given angle.

3.5.2 Slice selection

From the interpolated C'_{T2} slices, we can extract subsets of the slices with the same spacing as A_{T2} scans. We calculate the number of slices that must be skipped between each slice to achieve the desired spacing. Then, starting from the first slice, we extract subsets of slices separated by the calculated number of slices. This process is repeated until we reach the end of the C'_{T2} scans. We are left with about 500 subsets of C'_{T2} slices, which are not disjoint.

To optimize the registration process, we can remove about 50% of the subsets since the C'_{T2} scans overflow from A_{T2} scans as depicted in Fig. 2.

3.5.3 Registration

The registration algorithm itself is based on finding the best transformation of the C'_{T2} scans into the space of the A_{T2} scans. We use the Mean Squared Error metric. This algorithm is run on each subset of C'_{T2} slices separately, and the MSEs and transformations used to achieve them are saved for each subset.

The registration is initialized by creating a transformation where the C'_{T2} slides are placed in the middle of the image. Since we are working with pituitary adenoma, which is located in the sellar region of the brain, we can assume that the tumor is approximately located in the middle of the image. This is a good starting point for the registration process. We calculate the MSE of the C'_{T2} scans nudged by a few pixels in each direction as well as rotated by one degree in each way. We choose the direction with the lowest MSE, creating a simple gradient descent algorithm. The transformation is then updated in the chosen direction by a number of pixels dependent on a learning rate defined as a hyperparameter of the algorithm. However, the learning rate decays during the registration process to prevent the algorithm from overshooting the optimal transformation.

We apply a function that calculates a new learning rate based on the index of the current iteration and the total number of iterations. The function is defined as seen in Eq. 1, where i is the index of the iteration, n is the total number of iterations, lr_0 is a hyperparameter of the training process.

$$lr_i = \frac{lr_0}{\frac{i}{n} + lr_0} \tag{1}$$

The best transformation is identified by comparing the achieved MSEs of all the subsets. The subset with the lowest MSE is chosen as the best one, and the transformation used to achieve it is saved.

Upon successfully identifying the best subset, we can apply the transformation to the *original* C_{ceT1} scans, and the result can be saved as a 3D tensor with multiple channels, each representing one of the modalities.

4 Results and evaluation

In this section we present the results of our registration pipeline. To fairly evaluate the results, we used a baseline method to compare the results with.

4.1 Baseline multimodal registration

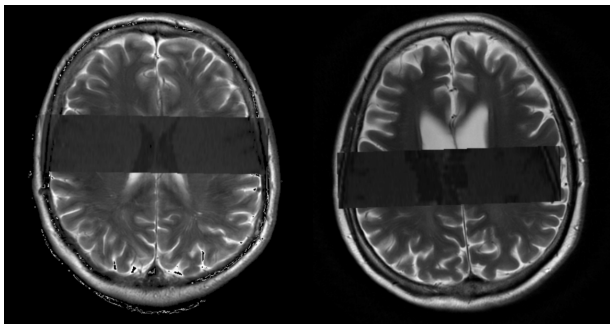

Figure 4: Results of image registration using the BSpline interpolation method. The first image shows a successful registration, while the second image was unsuccessful.

The registration was performed by an accommodated version of the sample code from the SimpleITK image-analysis notebooks collection [19] adjusted for 3D image registration. First, it was necessary to interpolate the C_{ceT1} slices to achieve isotropic voxels via BSpline or linear interpolation. Subsequently, we registered C_{ceT1} scans using the Mutual Information metric since we were dealing with multimodal images.

Regardless of the interpolation method, the registration was unsuccessful in many cases. The results of the registration using BSpline interpolation are shown in Fig. 4. Changing the interpolation method to linear interpolation did not improve the results significantly. To visualize the registration results, we overlapped the original A_{T2} scans with the C_{ceT1} scans (a dark horizontal rectangle).

4.2 GAN evaluation

The GAN network was trained to convert the input images between the two domains of MRI scans. The training setup is described in Sec. 3.4.

Since the GAN network is trained on a paired dataset, we can directly compare the original and generated images. Fig. 5 compares the input and output images. The results are of high quality, especially for the generated ceT1-weighted scans. For generated images of both modalities, the overall position of the whole head and all the structures inside the head are preserved. This is crucially important for the registration method to succeed. The contrast and lightness of the images are also very similar to the original images. There seems to be more difference in the lower part of the scans, but this is not a problem since the registration method only uses the upper part of the scans.

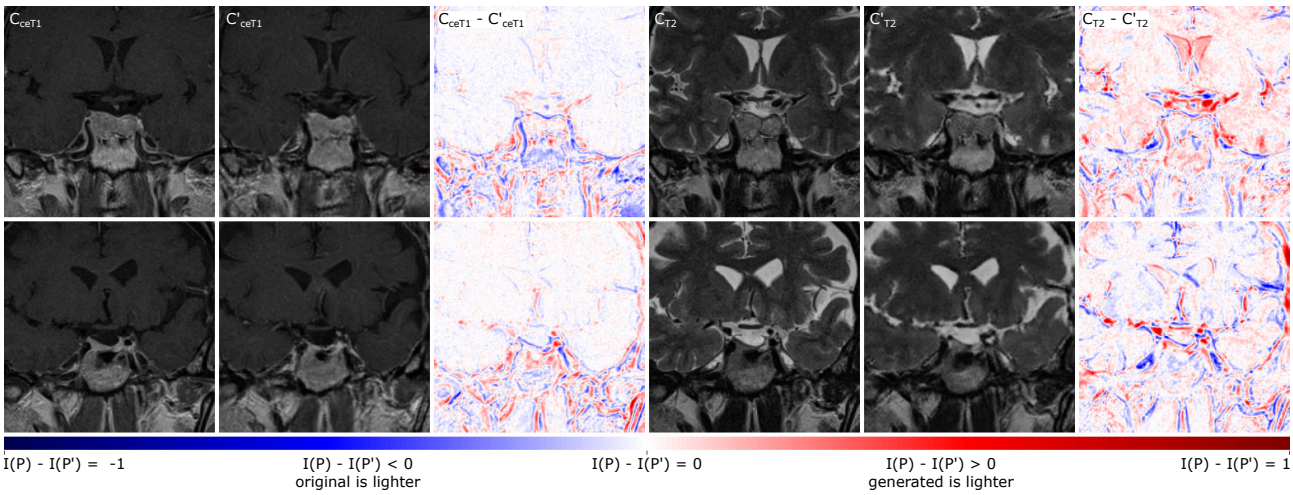

Figure 5: Comparison of original and GAN-generated coronal slices. Column 1 - C_{ceT1}. Column 2 - C'_{ceT1} generated from C_{T2} in column 4. Column 3 - difference between C_{ceT1} and C'_{ceT1} scans using the depicted colorscale. Columns 4 to 6 show the same for the T2 scans. $I(P)$ and $I(P')$ are the intensities of pixels in original and generated images respectively.

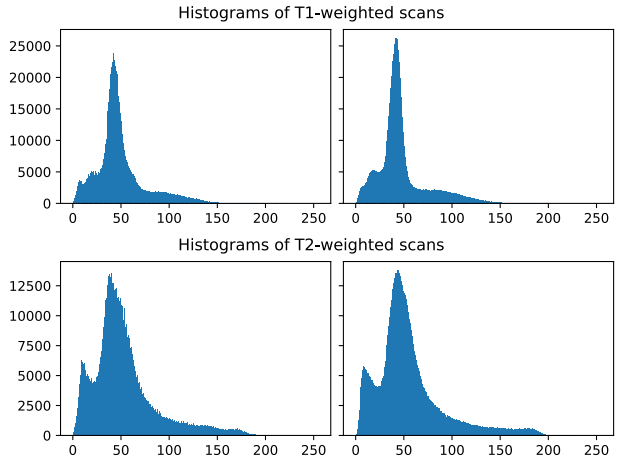

Figure 6: Comparison of histograms of the original and generated scans. Left column - original scans C_{ceT1} and C_{T2}, Right column - generated scans C'_{ceT1} and C'_{T2}.

	MSE of C'_{ceT1}	MSE of C'_{T2}
Mean	547.00	856.25
St. Dev.	181.11	292.48
Min	244.16	434.10
Lower quartile	437.74	724.79
Median	520.95	823.96
Upper quartile	612.72	931.70
Max	2969.93	4578.64

Table 1: Overview of the distribution of achieved MSEs by the GAN network over the paired dataset.

Based on the histograms displayed in Fig. 6, we can see that the intensity distribution of the generated images is very similar to the intensity distribution of the original images. This is a good sign since it means that the GAN network can preserve the intensity distribution of the input images as the generator network is trained to produce images that are indistinguishable from the original images.

Additionally, we compared the original and generated slices using the MSE metric. The distribution of these values is shown in Table 1. Similarly to our previous observations, we can see that the differences are more significant when generating T2 scans.

4.3 Registration method evaluation

The registration method described in Sec. 3.5 uses coronal slices converted from C_{ceT1} slices to C'_{T2} slices. We use the MSE of the C'_{T2} and the original A_{T2} slices to identify the correct subset of C'_{T2} slices.

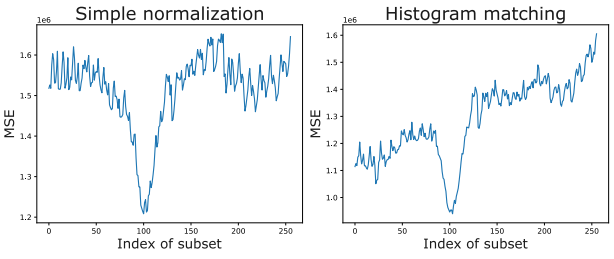

Figure 7: Cumulative MSEs over the subsets: Simple normalization (left) vs. Histogram matching (right). We can use this to identify the correct subset of matching images.

In Fig. 7, we can see the MSEs of all the subsets. We omitted the MSEs of the subsets that were removed from the registration process due to sticking out of the fixed image, as shown in Fig. 2. These subsets would show an even higher MSE than the ones shown in the chart.

The chart shows the cumulative MSEs of the subsets, which are calculated as the mean of the MSEs of the indi-

vidual slices in the subset. As seen in Fig. 7, we can effortlessly identify the correct subset of C'_{T2} slices by looking at the global minima of the cumulative MSEs.

The correct subset has an index of about 100. The registration results are shown in Fig. 8 over several slices of the selected subset.

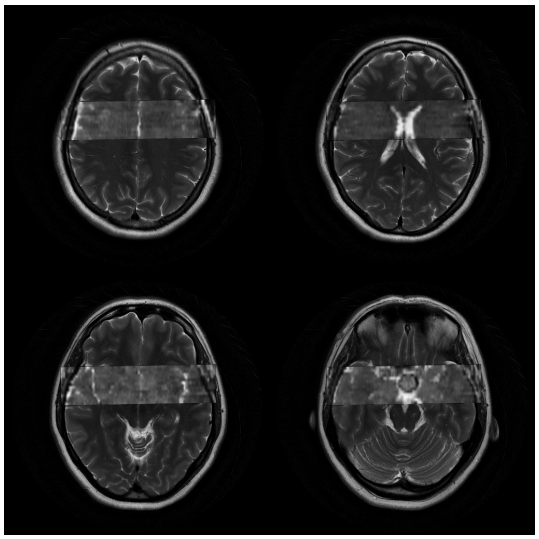

Figure 8: Registration results. Four A_{T2} slices of the same patient overlaid with the generated C'_{T2} slices.

The registration method was able to align the two images very well. The structures of the brain are aligned almost perfectly. The only visible differences are the graininess and lightness of the images. The interpolation of C'_{T2} slices to the same spacing causes graininess, which could be solved by applying some denoising techniques.

We opted for using just a simple normalization to the images, which converts the intensities of the produced images to the same range of intensities as the original images. Additionaly, we tried using a more sophisticated normalization technique, such as histogram matching, which matched the intensity distributions of the generated and original images over the area where we have both modalities. However, as depicted in Fig. 7, this approach didn't yield results that would make the process of the global minima identification clearer, and the registration algorithm took significantly longer to complete. Due to these reasons, we see no real benefit in a better normalization process.

Finally, we evaluate the achieved MSEs over a set of 77 patients. The registration process yielded in most of the cases relatively low MSEs as seen in Table 2 and Fig. 9. There have been some results that didn't converge to the correct results. Upon inspecting results of all 77 patients, we have found only 4 results that were noticeably misaligned, all of which were in the top 6 results with highest MSEs. We therefore evaluated, that the results with an MSE higher than 6×10^6 should be discarded as non-successful. This leaves us with 73 successful registrations meaning a $94,81\%$ success rate.

	MSE of the registration
Mean	2.811766×10^6
St. Dev.	1.980013×10^6
Min	5.884174×10^4
Lower quartile	1.464993×10^6
Median	2.275179×10^6
Upper quartile	3.716739×10^6
Max	1.037747×10^7

Table 2: Overview of the distribution of achieved MSEs over the registered dataset.

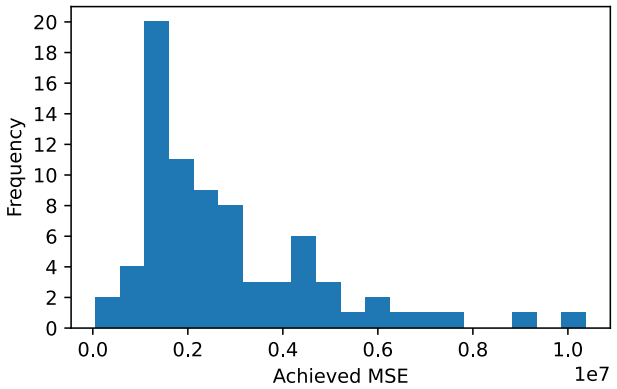

Figure 9: Histogram of achieved MSEs over the registered dataset.

5 Conclusions

We managed to create a comprehensive review of the main challenges of this field and gain an overview of the related state-of-the-art works. We ran several experiments that directed our research toward solving the most pressing issues. Moreover, we have employed a GAN network to convert images between MRI modalities. This helped to create a more robust image registration algorithm that yields respectable results of an acceptable accuracy.

Creating an excellent multimodal image registration is an important step that enables the fusion of the masks from the axial and coronal slides and, thus, the training of a segmentation network to segment tissue classes only marked for the other type of modality. As part of future work, we aim to create a segmentation CNN trained on the created registered dataset as a proof-of-concept of segmentation capabilities beyond what is possible when using scans of multiple modalities separately.

Acknowledgment

We want to thank MUDr. Martin Černý (Military University Hospital Prague) for providing the extensive PA dataset with annotations, the registered multimodal dataset of coronal MRI scans, for his valuable expertise in radiology, and his feedback on this paper.

References

[1] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, Adrian V. Dalca, and John Guttag. An unsupervised learning model for deformable medical image registration. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9252–9260, 2018.

[2] Max Blendowski, Nassim Bouteldja, and Mattias P Heinrich. Multimodal 3D medical image registration guided by shape encoder–decoder networks. *International journal of computer assisted radiology and surgery*, 15:269–276, 2020.

[3] Jean-François Bonneville, Fabrice Bonneville, and Francoise Cattin. Magnetic resonance imaging of pituitary adenomas. *European radiology*, 15:543–8, 04 2005.

[4] Michael Buchfelder and Sven Schlaffer. Imaging of pituitary pathology. *Handbook of clinical neurology*, 124:151–166, 2014.

[5] Xiaohuan Cao, Jianhua Yang, Yaozong Gao, Qian Wang, and Dinggang Shen. Region-adaptive deformable registration of CT/MRI pelvic images via learning-based image synthesis. *IEEE Transactions on Image Processing*, 27(7):3500–3512, 2018.

[6] Runfa Chen. NICE-GAN-pytorch: Official PyTorch implementation of NICE-GAN: Reusing discriminators for encoding: Towards unsupervised image-to-image translation. `https://github.com/alpc91/NICE-GAN-pytorch`, 2020, Accessed on 05.03.2024.

[7] Runfa Chen, Wenbing Huang, Binghui Huang, Fuchun Sun, and Bin Fang. Reusing discriminators for encoding: Towards unsupervised image-to-image translation. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8165–8174, 2020.

[8] Adrian F. Daly and Albert Beckers. The epidemiology of pituitary adenomas. *Endocrinology and Metabolism Clinics of North America*, 49(3):347–355, 2020.

[9] Robert J Gillies, Paul E Kinahan, and Hedvig Hricak. Radiomics: Images are more than pictures, they are data. *Radiology*, 278(2):563–577, 2016.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

[11] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[12] Engelbert Knosp, Erich Steiner, Klaus Kitz, and Christian Matula. Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings. *Neurosurgery*, 33(4):610–618, 1993.

[13] Wen Li, Yafen Li, Wenjian Qin, Xiaokun Liang, Jianyang Xu, Jing Xiong, and Yaoqin Xie. Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy. *Quantitative imaging in medicine and surgery*, 10(6):1223, 2020.

[14] F. Maes, D. Vandermeulen, and P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10):1699–1722, 2003.

[15] Marius E. Mayerhoefer, Andrzej Materka, Georg Langs, Ida Häggström, Piotr Szczypiński, Peter Gibbs, and Gary Cook. Introduction to radiomics. *Journal of Nuclear Medicine*, 61(4):488–495, 2020.

[16] Gerald Raverot, Emmanuel Jouanneau, and Jacqueline Trouillas. MANAGEMENT OF ENDOCRINE DISEASE: Clinicopathological classification and molecular markers of pituitary tumours for personalized therapeutic strategies. *European Journal of Endocrinology*, 170(4):R121–R132, 04 2014.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[18] Walavan Sivakumar, Roukoz Chamoun, Vinh Nguyen, William T Couldwell, et al. Incidental pituitary adenomas. *Neurosurgical focus*, 31(6):E18, 2011.

[19] Ziv Yaniv, Bradley C Lowekamp, Hans J Johnson, and Richard Beare. SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. *Journal of digital imaging*, 31(3):290–303, 2018.

[20] Yuanjie Zheng, Xiaodan Sui, Yanyun Jiang, Tongtong Che, Shaoting Zhang, Jie Yang, and Hongsheng Li. SymReg-GAN: Symmetric image registration with generative adversarial networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):5631–5646, 2022.