Generative Augmentation for Brain Adenoma MRI Segmentation

Bc. Richard Szarka*
Supervised by: prof. Ing. Vanda Benešová, PhD.†

Faculty of Informatics and Information Technoology
Slovak University of Technology
Bratislava / Slovakia

Abstract

Deep learning has shown great promise in fields such as radiology, but training robust segmentation networks is difficult in small healthcare systems due to the limited availability of annotated MRI datasets. To overcome this challenge, we propose a generative data augmentation approach that enhances training efficiency by synthesizing both MRI scans and their corresponding segmentation masks.

Our method consists of two key components: the Key-Slice Extractor Network and the Conditional Diffusion Projected Generative Adversarial Network (CDP-GAN). The Key-Slice Extractor Network is designed to identify the most relevant MRI slices, reducing redundancy and improving the quality of the training dataset. These selected slices serve as inputs for our CDP-GAN, which utilizes slice index conditioning and diffusion-based augmentation to generate realistic 3D MRI volumes of pituitary adenoma (PA). Our model produces anatomically consistent synthetic MRI scans and generates corresponding segmentation masks, enabling a direct expansion of labelled training data.

To assess the impact of our approach, we evaluate the performance of a 3D U-Net segmentation model trained on different combinations of real and synthetic MRI data. Our findings suggest that incorporating CDP-GAN-generated samples, complete with segmentation masks, should enhance segmentation accuracy, demonstrating the effectiveness of generative augmentation in data-scarce settings. This framework offers a scalable solution for improving AI-driven medical imaging applications, particularly in regions with limited access to large annotated datasets.

Keywords: Generative Neural Networks, Segmentation, Pituitary Adenoma, Projected GAN

1 Introduction

In recent years, advancements in neural networks have facilitated their application across numerous domains, including medicine. These computational models now serve essential functions in classification, segmentation, and synthetic data generation tasks within healthcare settings [13].

Pituitary adenoma predominantly affects individuals between 30 and 35 years of age. Despite their generally benign nature, these tumors represent a significant health and economic burden globally. Current diagnostic protocols rely heavily on subjective assessment by clinical specialists, introducing potential for human error [15]. Pituitary adenomas exhibit dimensional variability, with smaller formations classified as microadenomas and larger ones as macroadenomas. The latter can induce serious symptoms including partial vision loss, headaches, and hyperpituitarism [19]. Given the relatively high incidence of pituitary adenomas, early diagnosis and treatment can substantially improve patient outcomes [4].

Developing effective segmentation networks offers the potential to enhance clinical workflow efficiency, enabling more rapid diagnosis and improved patient care. However, medical imaging datasets in specialized domains are characteristically limited, often resulting in suboptimal model performance due to insufficient training examples. This paper presents a data augmentation approach utilizing synthetic generation of three-dimensional pituitary adenoma MRI volumes with corresponding annotations. Our pipeline is specifically designed for implementation in resource-constrained environments.

The primary contribution of our work is:

- We have implemented a synthetic medical data generation framework, comprising a preprocessing pipeline, key-slice extraction mechanism, and conditional diffusion-projected GAN architecture optimized for volumetric consistency on a tiny dataset (31 volumes).

Due to our small sample size, we avoided using Fréchet Inception Distance (FID) and similar distribution-based metrics, which require larger datasets to produce reliable estimates. Instead, we evaluated quality through clinical expert assessment (with only 30% of synthetic volumes correctly identified as artificial) and through the consistent improvement in CNN segmentation performance when using our synthetic data. Additionally, we quantitatively assessed our key-slice extractor using standard classification metrics (accuracy, precision, recall, F1) and evaluated

*xszarkar@stuba.sk
†vanda_benesova@stuba.sk

segmentation performance with the Dice similarity coefficient.

2 Related Work

Pituitary adenoma (PA) segmentation research has advanced considerably through deep learning innovations. W. Rui et al. pioneered MTMAU-Net, an integrated framework that concurrently addresses macroadenoma segmentation while determining cavernous sinus invasion status, with performance surpassing clinical experts particularly for Knosp grade 3 classifications [16]. Q. Zhang's et al. contribution, MSR-Net, represents a sophisticated architecture leveraging dual decoding pathways that implement channel-specific attention to prioritize informative features while simultaneously employing spatial attention mechanisms to establish global contextual relationships [21]. M. Černý et al. advanced the field with an automated system that intelligently identifies pertinent MRI slices using a dedicated selection network before applying segmentation algorithms [2]. X. Shu et al. explored the adaptability of nnU-Net across diverse pituitary adenoma presentations, demonstrating comparable efficacy between general adenoma segmentation and non-functioning pituitary adenoma-specific models, offering an accessible solution requiring minimal technical expertise [18]. Further architectural refinements include R. Mutten's et al. implementation of 2D U-Net variants for comparative pre-surgical and post-surgical volume assessment [3], while X. Jiang et al. enhanced traditional U-Net designs through innovative connectivity patterns—introducing cross-layer pathways, inception-dense blocks, and comprehensive skip connections that establish direct links between encoding layers and multiple subsequent upsampling stages [7].

Generative Adversarial Networks (GANs) have transformed medical imaging synthesis capabilities. A. Elazab et al. conceptualized GP-GAN, an architecture featuring interconnected generator networks with 3D U-Net foundations that enable temporal prediction of tumor evolution across multiple time points, incorporating tumor-specific feature representations as guiding inputs [6]. D. Mukherkjee's et al. AGGrGAN introduced a novel approach combining multiple generative models with strategic style transfer techniques, implementing pixel-level fusion weighted according to edge prominence to preserve critical structural details [10]. E. Jung et al. focused on disease trajectory modeling through conditional GANs incorporating specialized 3D discriminators, producing temporally coherent Alzheimer's progression sequences through dedicated transition modules and specialized loss functions, including attention and variational components [8]. A. Mahboubisarighieh et al. established 3D Dual-CycleGAN methodology for cross-modal translation of volumetric T1-weighted acquisitions to alternative contrast sequences, employing a comprehensive loss framework encompassing voxel relationships, gradient preservation, perceptual similarity, and structural similarity [9]. H. Alrashedy et al. developed BrainGAN as an end-to-end solution combining Deep Convolutional GAN synthesis with subsequent CNN-based analytical capabilities, creating an integrated pipeline that generates realistic brain MRI data which then supports automated pathology classification [1].

Diffusion Probabilistic Models (DDPM) and Latent Diffusion Models (LDM) represent emerging paradigms with distinct advantages over GAN frameworks. W. Peng et al. demonstrated 2D conditional Diffusion Probabilistic Models capable of generating comprehensive MRI subvolumes based on contextual slice information, employing strategic slice combinations and self-attention mechanisms to capture long-range anatomical relationships while maintaining computational efficiency [12]. Pinaya et al. explored the capabilities of combined latent diffusion and denoising diffusion implicit models for producing high-fidelity three-dimensional neuroimaging volumes, successfully conditioning generation on clinically relevant parameters including ventricular dimensions and normalized brain volume metrics, though noting substantial computational and data requirements [14]. S. Pan et al. introduced Cycle-guided DDPM architecture featuring bidirectionally conditioned diffusion models that simultaneously synthesize complementary MRI contrasts through coordinated latent noise exchange during reverse diffusion processes, integrating Swin-transformer components to enhance spatial coherence across generated T1, FLAIR, and T2 sequences [11]. L. Zhu's et al. Make-A-Volume framework addressed cross-modality synthesis challenges through latent diffusion models operating in compressed representational spaces, incorporating specialized volumetric architectural elements within primarily slice-oriented processing pipelines to ensure three-dimensional consistency [22]. Z. Dorjsembe et al. developed conditional diffusion models specifically engineered for semantic 3D MRI synthesis, employing strategic residual blocks and attention mechanisms at critical sampling transitions while accepting tumor mask inputs as conditioning signals [5].

Addressing limited training data challenges, A. Sauer's et al. Projected GAN approach combines FASTGAN generation with multiple discriminator networks leveraging EfficientNetLite1 representations, introducing cross-scale and cross-channel mixing techniques that accelerate convergence and enhance output quality, particularly in data-constrained environments [17]. Z. Wang et al. formulated Diffusion-GAN methodology that integrates diffusion-based augmentation strategies for both synthetic and authentic images, demonstrating empirical performance improvements without necessitating strict noise scheduling dependencies [20].

Diverse dataset sizes in the analyzed related work span from large collections exceeding 1,000 samples [21, 14, 10, 5, 9, 11, 1], intermediate-sized datasets of 500-1,000 examples [16, 8, 2], small collections of 100-499 instances

[3, 18, 22], to very limited datasets below 100 samples [7, 22, 6]. Some of the analyzed studies were excluded due to methodological constraints (for example, using 2D MRI slices).

3 Dataset

As highlighted in our related work review, small dataset sizes consistently challenge the application of deep neural networks in brain MRI analysis or generation. Our study faces this same limitation, which directly motivated our approach. Our dataset contains 31 brain MRI volumes with pituitary adenoma, each with corresponding ground truth segmentation masks. We allocated 25 volumes for training and 6 for validation throughout all experiments. While this validation set is small, it is representative of the anatomical variability in our dataset and sufficient for our comparative evaluation purposes, as we primarily assess relative performance improvements between different training configurations rather than claiming absolute generalization to all possible cases. These scans were acquired on different MRI machines, resulting in volumetric data with variable dimensions and characteristics. Our solution effectively addresses these inconsistencies through appropriate preprocessing techniques. Figure 1 presents a representative example from our dataset.

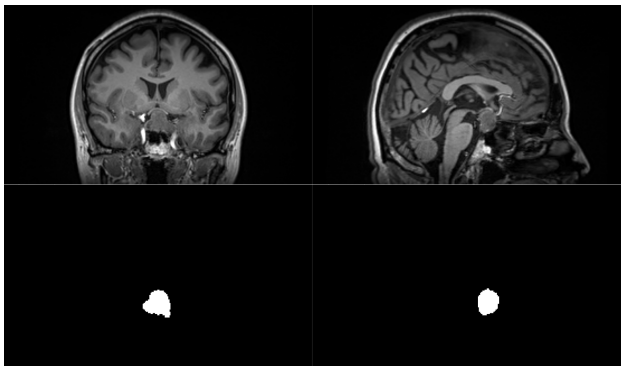

Figure 1: A sagittal slice (left) and slice coronal (right) sample from our dataset (top) with corresponding ground truth masks (bottom)

4 Our approach

In this section, we present the core components of our approach. Our framework addresses the challenge of limited data availability, a common constraint when applying deep neural networks to specialized domains. We developed this solution specifically to tackle the small sample size problem that affects many research areas, particularly in medical imaging. While our framework offers effective strategies for working with limited datasets, it is important to acknowledge that increasing the quantity of training data remains one of the most powerful methods for improving deep neural network performance, especially for generative tasks.

4.1 Pipeline of our approach

Our pipeline consists of 4 phases:

In the first preprocessing phase, we used our preprocessing pipeline to extract the volume of the adenoma region. Taking advantage of the fact that the pituitary adenoma has a stable position in the anatomy of the brain is advantageous when it comes to neglecting unnecessary information from the other MRI parts and also decreases drastically the memory requirements.

The second phase, key-slice extraction, employs a 2D CNN classification network that identifies whether each slice qualifies as a key-slice. This classification enhances the generative network's learning by focusing on the most informative regions of the volume. In the key-slice areas, conditional information about pituitary adenomas provides maximum impact, creating optimal learning conditions for the generative model. Our implementation classifies slices through the sagittal plane, allowing the network to visualize the pituitary adenoma, cavernous sinus, and internal carotid artery simultaneously. It's important to note that alternative applications might benefit from different plane orientations depending on the specific anatomical structures of interest. The first 12 slices in each volume were used as an input to the next step. Larger indexes of slices were underrepresented due to the different sizes of pituitary adenoma.

The third phase, generation, utilizes a Conditional Diffusion Projected GAN framework. Projected GANs with diffusion augmentation have demonstrated superior stability, addressing many common challenges in GAN training. Given our limited dataset size, mode collapse represented a significant concern, but our approach effectively mitigated this risk while maintaining healthy training dynamics. The network is conditioned on the key-slice index identified during the previous extraction phase. For volume generation, we employ a consistent latent vector Z while systematically iterating through the condition index, enabling coherent three-dimensional synthesis across the entire volume.

The fourth phase, segmentation, serves as an evaluation mechanism to assess how our synthetically generated data enhances performance on our already limited dataset. We implemented a 3D U-Net architecture for this segmentation task. Additionally, we sought to determine whether our generative approach could significantly improve segmentation capabilities in scenarios with even more severely constrained real training data availability.

A visual representation of our pipeline is given in the figure 2.

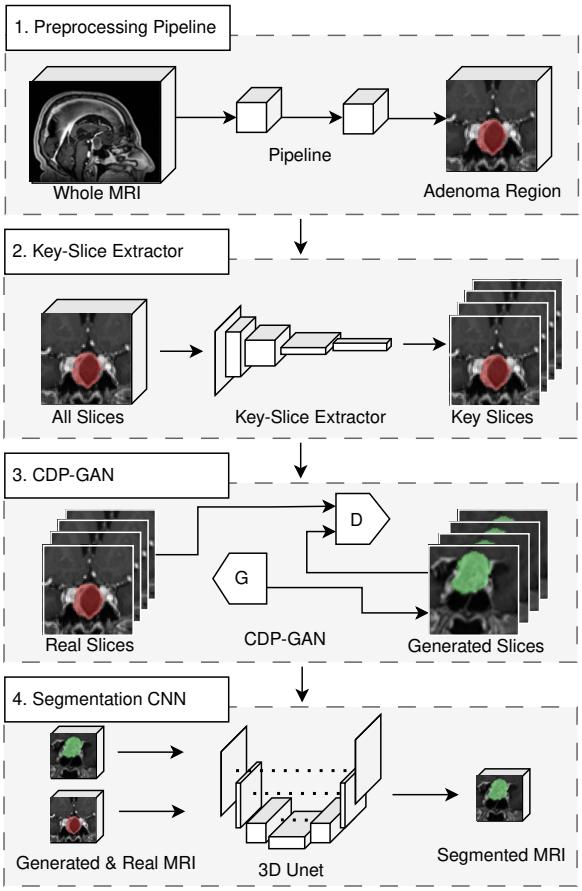

Figure 2: Our proposed pipeline. It consists of 4 phases: preprocessing, key-slice extraction, synthetic data generation and segmentation training. All red annotated MRIs are real and green annotated are generated.

4.2 Data Preprocessing

Our data preprocessing pipeline, illustrated in Figure 3, addresses several critical challenges inherent to medical imaging datasets. The workflow begins with false annotation cleaning to remove inaccurate segmentations that could potentially compromise model training. This step is particularly crucial since false annotation could mislead our adenoma region extraction process. Furthermore, it erases human mistakes, which would mislead the evaluation process of our study.

Following annotation refinement, we implement rigid registration to standardize the spatial orientation of all volumes. The criterion is mutual information. This transformation ensures anatomical consistency across the dataset, facilitating more effective learning of spatial relationships. It unifies the dimensionality of all volumes; therefore, MRI from multiple scanning machines can be utilized.

The registered images then undergo rotational augmentation, artificially expanding our limited dataset while enhancing the model's robustness to orientation variations that naturally occur in clinical settings (+-3 degrees on every axis).

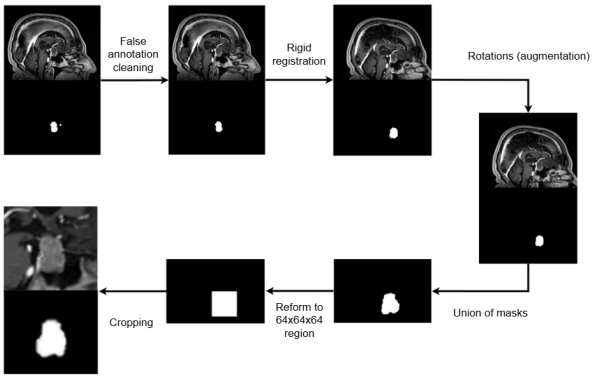

Figure 3: Our preprocessing pipeline designed for extraction of relevant 64x64x64 volume around pituitary adenoma.

The pipeline continues with a union of masks operation that consolidates segmentation information when multiple annotations are available. This step is followed by reformatting all volumes to a standardized 64×64×64 voxel dimension, ensuring computational consistency during network training. The final preprocessing step involves focused cropping around the region of interest, which not only reduces computational requirements but also concentrates the network's learning capacity on the anatomically relevant structures surrounding the pituitary gland.

Finally, as a last step, the volumes' values are normalized to a range between 0 and 1.

4.3 Key-Slice extractor

We define a key slice as one containing at least 50% of the maximum pituitary adenoma area observed across any slice in the volume. For a slice s_i in volume $S = \{s_1, s_2, ..., s_n\}$, let $A(s_i)$ represent the area of pituitary adenoma in slice s_i. We define the maximum adenoma area as $A_{max} = \max_{s_i \in S} A(s_i)$. The set of key slices K is then defined as:

$$K = \{s_i \in S \mid A(s_i) \geq 0.5 \cdot A_{max}\}$$

We found that using all slices for the generative network was suboptimal, as peripheral (edge) slices contain limited relevant information and introduced noise to the learning process. The generative network struggled to capture the continuous character of MRI volumes when trained on all slices. By focusing only on information-dense key-slices, we achieved more efficient training and better anatomical consistency in the generated volumes, which proved particularly valuable given our small dataset size.

Our key-slice extractor architecture consists of a standard convolutional feature extractor followed by a classification head with fully connected layers. The network comprises 6 layers that produce feature maps with dimensions 1024×4^3. Since spatial information is not critical for

this classification task, we apply global average pooling to reduce dimensionality, creating a condensed feature vector of length 1024. This vector is then processed through two fully connected layers, culminating in a 64-value classification output that identifies each slice as either a key-slice or non-key-slice. We trained the network using binary cross entropy (BCE) loss function. Additionally, we conducted experiments with architectural variations, including inception blocks and skip connections to evaluate their impact on classification performance. This binary classification guides the subsequent generation process by highlighting the most information-rich regions within each volume.

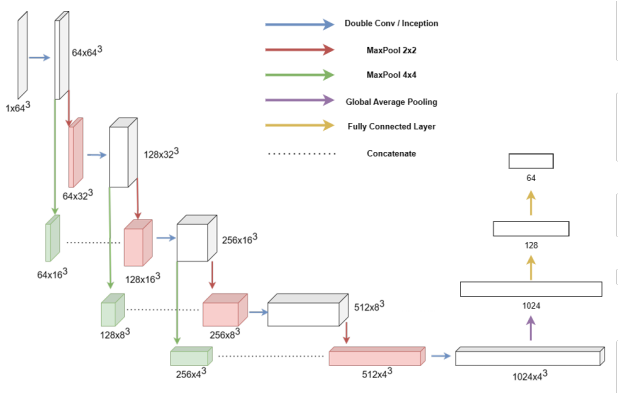

Figure 4: The architecture of Key-Slice Extractor

4.4 Conditional Diffusion Projected GAN

In this section, we discuss our generative neural network implementation based on ProjectedGAN, which demonstrates exceptional performance on low-sample datasets - precisely, our operating scenario. Our generator utilizes the Fast-GAN architecture, while feature extraction occurs through EfficientNet lite-1, which feeds into four distinct discriminators. Fast-GAN in our scenario produces a two-channel output from random latent vector Z, conditioned on the MRI slice index to be produced. Projected GAN utilizes random projection, with cross-scale and cross-channel mixing applied to the features from EfficientNet lite-1, before feeding it to the discriminator.

To enable conditional generation based on key-slice indices, we integrated conditional information into both the generator and all four discriminators. This conditioning is implemented through one-hot encoding, is concatenated with latent vector Z. Besides the conditioning is propagated through the initial layers of the generator. For the discriminators, we reshape the one-hot vectors and concatenate them with features extracted by EfficientNet lite-1. The complete architectural design and conditioning mechanism are illustrated in Figure 5.

Following Wang et al.'s approach [20], we enhance training stability by introducing Gaussian noise to the images based on the ratio between generator and discrim-

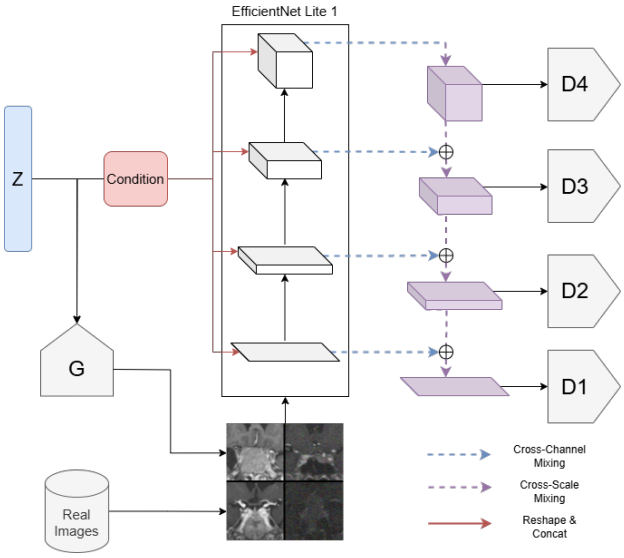

Figure 5: The architecture of proposed Conditional Diffusion Projected GAN (G stands for generator and D stands for discriminator).

inator performance. The intuition behind this technique is straightforward: discriminators typically face a simpler challenge than generators and can easily identify synthetic images. By applying Gaussian noise with varying timestamp parameters t, we shift the distributions of generated and real images closer together, creating a more balanced adversarial relationship and preventing mode collapse. Additionally, we employ differentiable augmentation to further address the limited data challenge, apply hinge loss for adversarial training to improve convergence stability, and implement Exponential Moving Average (EMA) for the generator's weights to produce more consistent and higher quality outputs during inference.

We addressed an important architectural challenge regarding input channels: while EfficientNet lite-1 expects 3-channel inputs, MRI images are typically single-channel. Rather than simply duplicating the MRI channel three times, we developed a more informative approach. We assigned the MRI image to the first channel (red channel), the segmentation mask to the second channel (blue), and filled the third channel with zeros (green). This configuration yielded superior results, particularly for simultaneous generation of both images and annotations. The cross-channel mixing mechanism in ProjectedGAN, which applies random 1×1 convolutions (hence "projected"), effectively processes this multi-channel information and distributes it across all discriminators.

4.5 3D CNN segmentation network

The CNN segmentator consists of a standard 3D U-Net architecture with 4 downsampling blocks, a bottleneck layer, and 4 upsampling blocks. The bottleneck layer expands feature maps to 1024×4^3 dimensions. Each block con-

tains a simple double convolution with ReLU activation and batch normalization. We deliberately kept this architecture basic since the network's primary purpose is to evaluate our synthetically generated dataset rather than to advance segmentation techniques.

5 Results

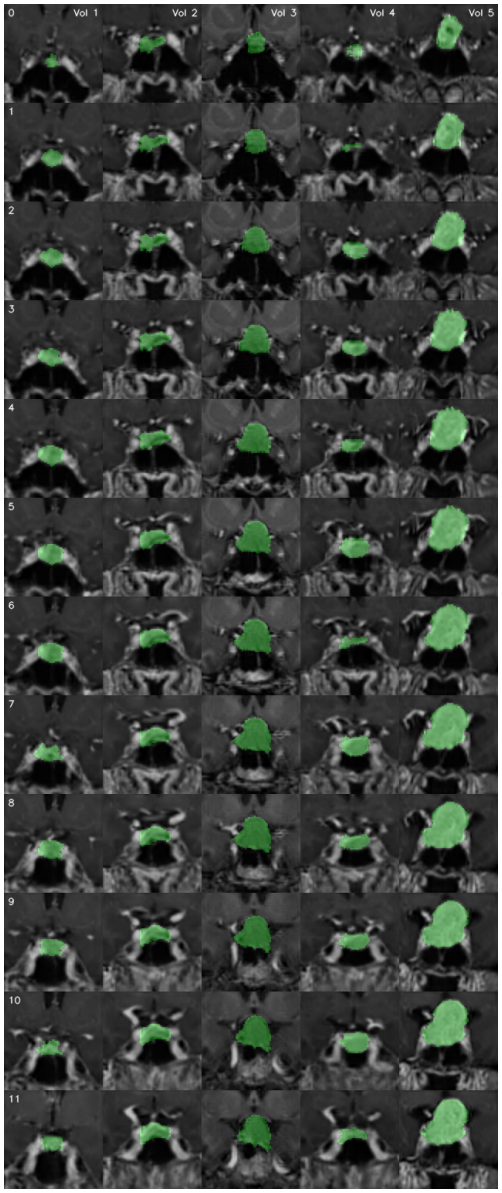

Figure 6: Generated annotations for the corresponding MRI slice are highlighted with a transparent green color. A grid visualizing slices of 5 generated MRI volumes. Each volume is in a separate column, while rows represent the slice index conditioning.

We tested several architectural variations for the Key-Slice extractor to determine the optimal design. Importantly, we excluded segmentation masks from the classification inputs, as using masks would defeat the purpose of creating an automated extractor - the main goal is to identify key slices without requiring time-consuming expert annotations.

Our experiments, shown in Table 1, revealed that the base model without modifications performed best for this task. Both inception blocks and cross-layer connections actually decreased performance across all configurations. We implemented early stopping based on a "patience" parameter to optimize training duration.

Model	Accuracy	Precision	Recall	F1
Base	**0.9479**	**0.9683**	**0.8141**	**0.8845**
Cross-Layer	0.9401	0.9610	0.7951	0.8703
Inception	0.9297	0.8687	0.8031	0.8346
Both	0.9167	0.8990	0.7584	0.8228

Table 1: Performance comparison of different architectural variations of Key-Slice Extractor.

The validation results confirmed that the model successfully extracts continuous sequences of relevant slices without gaps, which is essential to maintain anatomical coherence in the generation phase.

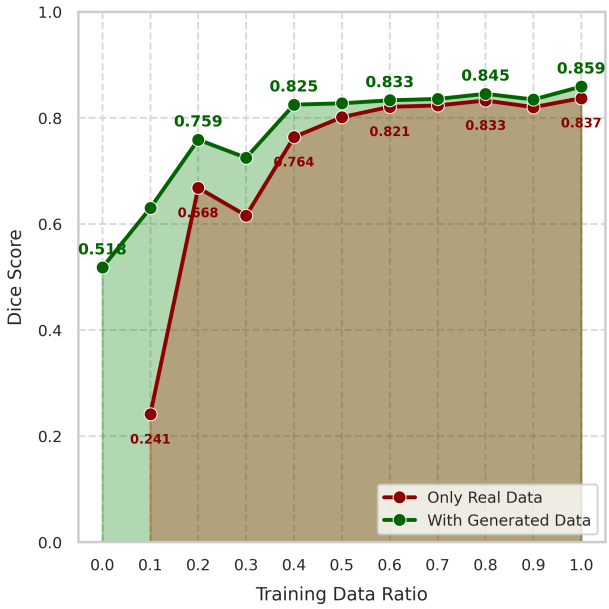

Figure 7: Dice similarity scores across varying real training data ratio (ratio of how much of the real dataset was used, zero meaning non real slices and 1.0 meaning all of the real slices from MRI volumes were used), showing consistent improvements when augmenting with synthetic data, particularly in low-resource scenarios

For the CDP-GAN, we conducted training for over 310 hours using an NVIDIA TITAN V GPU. The model exhibited stable training dynamics with generator and discriminator loss functions gradually converging toward similar values, indicating a healthy adversarial balance without

either component dominating. Using our trained model, we generated 2000 MRI volumes to match the approximate size of our augmented training dataset. The generation process demonstrated efficient performance, requiring an average of 0.5281 seconds per complete volume and 0.0252 seconds per individual slice. The synthetic dataset was evaluated by a medical professional. 20 MRI samples (10 real and 10 generated) were provided to the professional. He was able to correctly categorize 30% of the MRIs, which supports the excellent generation quality of the synthetic volumes. Figure 6 presents representative examples of the generated MRI volumes.

Finally, the 3D U-Net architecture served as an evaluation framework for our proposed pipeline. To ensure reproducibility, we systematically fixed all random seed components throughout our implementation. The experimental design involved training the segmentation model with varying proportions of the training dataset while maintaining a constant ratio (1.0) of synthetic data. Given the dimensional discrepancy between generated volumes (64×64×12) and real training data (64×64×64), we applied zero-padding to the synthetic volumes to maintain dimensional consistency without introducing confounding information. As illustrated in Figure 7, the presence of synthetic data significantly enhanced segmentation performance as measured by the Dice similarity score. Notably, the impact was most pronounced in low-resource scenarios (with only 2 or 5 MRI volumes), where the segmentation accuracy improved substantially. Although the marginal benefit diminished as the proportion of real training data increased, the synthetic dataset consistently improved segmentation performance across all experimental configurations and effectively mitigated overfitting. These findings suggest that integrating our generation pipeline with state-of-the-art pituitary adenoma segmentation networks could yield further performance improvements.

6 Conclusions

This paper presents a novel generative augmentation pipeline for pituitary adenoma MRI segmentation that effectively addresses the challenge of limited training data. Our approach combines a key-slice extraction network with a Conditional Diffusion Projected GAN to generate high-quality synthetic MRI volumes complete with corresponding segmentation masks. The experimental results demonstrate that incorporating our synthetic data consistently improves segmentation performance across various training scenarios. The most significant benefits were observed in extremely data-limited settings, where the inclusion of generated volumes substantially increased segmentation accuracy. This confirms the practical utility of our approach for real-world medical imaging applications where annotated data is scarce. While our current implementation focuses specifically on pituitary adenoma, the underlying methodology applies to other anatomical regions and pathologies. The ability to generate anatomically accurate synthetic medical images with corresponding segmentation masks represents an important step toward overcoming data limitations in specialized healthcare domains. By reducing the dependency on large, manually annotated datasets, our approach can accelerate the development and deployment of deep learning-based tools for clinical practice, potentially improving diagnostic capabilities in resource-constrained environments.

References

[1] Halima Hamid N Alrashedy, Atheer Fahad Almansour, Dina M Ibrahim, and Mohammad Ali A Hammoudeh. Braingan: brain mri image generation and classification framework using gan architectures and cnn models. *Sensors*, 22(11):4297, 2022.

[2] Martin Černỳ, Jan Kybic, Martin Májovskỳ, Vojtěch Sedlák, Karin Pirgl, Eva Misiorzová, Radim Lipina, and David Netuka. Fully automated imaging protocol independent system for pituitary adenoma segmentation: a convolutional neural network—based model on sparsely annotated mri. *Neurosurgical Review*, 46(1):116, 2023.

[3] Raffaele Da Mutten, Olivier Zanier, Olga Ciobanu-Caraus, Stefanos Voglis, Michael Hugelshofer, Athina Pangalu, Luca Regli, Carlo Serra, and Victor E Staartjes. Automated volumetric assessment of pituitary adenoma. *Endocrine*, 83(1):171–177, 2024.

[4] Adrian F Daly and Albert Beckers. The epidemiology of pituitary adenomas. *Endocrinol Metab Clin North Am*, 49(3):347–355, 2020.

[5] Zolnamar Dorjsembe, Hsing-Kuo Pao, Sodtavilan Odonchimed, and Furen Xiao. Conditional diffusion models for semantic 3d brain mri synthesis. *IEEE Journal of Biomedical and Health Informatics*, 2024.

[6] Ahmed Elazab, Changmiao Wang, Syed Jamal Safdar Gardezi, Hongmin Bai, Qingmao Hu, Tianfu Wang, Chunqi Chang, and Baiying Lei. Gp-gan: Brain tumor growth prediction using stacked 3d generative adversarial networks from longitudinal mr images. *Neural Networks*, 132:321–332, 2020.

[7] Xiaoliang Jiang, Junjian Xiao, Qile Zhang, Lihui Wang, Jinyun Jiang, and Kun Lan. Improved u-net based on cross-layer connection for pituitary adenoma mri image segmentation. *Mathematical Biosciences and Engineering: MBE*, 20(1):34–51, 2022.

[8] Euijin Jung, Miguel Luna, and Sang Hyun Park. Conditional gan with 3d discriminator for mri generation of alzheimer's disease progression. *Pattern Recognition*, 133:109061, 2023.

[9] Ali Mahboubisarighieh, Hossein Shahverdi, Shabnam Jafarpoor Nesheli, Mohammad Alipoor Kermani, Milad Niknam, Mohanna Torkashvand, and Seyed Masoud Rezaeijo. Assessing the efficacy of 3d dual-cyclegan model for multi-contrast mri synthesis. *Egyptian Journal of Radiology and Nuclear Medicine*, 55(1):118, 2024.

[10] Debadyuti Mukherkjee, Pritam Saha, Dmitry Kaplun, Aleksandr Sinitca, and Ram Sarkar. Brain tumor image generation using an aggregation of gan models with style transfer. *Scientific reports*, 12(1):9141, 2022.

[11] Shaoyan Pan, Chih-Wei Chang, Junbo Peng, Jiahan Zhang, Richard LJ Qiu, Tonghe Wang, Justin Roper, Tian Liu, Hui Mao, and Xiaofeng Yang. Cycle-guided denoising diffusion probability model for 3d cross-modality mri synthesis. *arXiv preprint arXiv:2305.00042*, 2023.

[12] Wei Peng, Ehsan Adeli, Tomas Bosschieter, Sang Hyun Park, Qingyu Zhao, and Kilian M Pohl. Generating realistic brain mris via a conditional diffusion probabilistic model. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 14–24. Springer, 2023.

[13] Francesco Piccialli, Vittorio Di Somma, Fabio Giampaolo, Salvatore Cuomo, and Giancarlo Fortino. A survey on deep learning in medicine: Why, how and when? *Information Fusion*, 66:111–137, 2021.

[14] Walter HL Pinaya, Petru-Daniel Tudosiu, Jessica Dafflon, Pedro F Da Costa, Virginia Fernandez, Parashkev Nachev, Sebastien Ourselin, and M Jorge Cardoso. Brain imaging generation with latent diffusion models. In *MICCAI Workshop on Deep Generative Models*, pages 117–126. Springer, 2022.

[15] Yu Qian, Yue Qiu, Cheng-Cheng Li, Zhong-Yuan Wang, Bo-Wen Cao, Hong-Xin Huang, Yi-Hong Ni, Lu-Lu Chen, and Jin-Yu Sun. A novel diagnostic method for pituitary adenoma based on magnetic resonance imaging using a convolutional neural network. *Pituitary*, 23:246–252, 2020.

[16] W Rui, W Gao, N Qiao, X Chen, M Han, Y Wu, T Xin, J Yang, Y Zhao, and Z Yao. Automatic pituitary adenoma segmentation and identification of cavernous sinus invasion via multitask learning. *Clinical Radiology*, 80:106756, 2025.

[17] Axel Sauer, Kashyap Chitta, Jens Müller, and Andreas Geiger. Projected gans converge faster. *Advances in Neural Information Processing Systems*, 34:17480–17492, 2021.

[18] Xujun Shu, Yijie Zhou, Fangye Li, Tao Zhou, Xianghui Meng, Fuyu Wang, Zhizhong Zhang, Jian Pu, and Bainan Xu. Three-dimensional semantic segmentation of pituitary adenomas based on the deep learning framework-nnu-net: A clinical perspective. *Micromachines*, 12(12):1473, 2021.

[19] Nicholas A Tritos and Karen K Miller. Diagnosis and management of pituitary adenomas: a review. *Jama*, 329(16):1386–1398, 2023.

[20] Zhendong Wang, Huangjie Zheng, Pengcheng He, Weizhu Chen, and Mingyuan Zhou. Diffusion-gan: Training gans with diffusion. *arXiv preprint arXiv:2206.02262*, 2022.

[21] Qile Zhang, Xiaoliang Jiang, Xiuqing Huang, and Chun Zhou. Msr-net: Multi-scale residual network based on attention mechanism for pituitary adenoma mri image segmentation. *IEEE Access*, 2024.

[22] Lingting Zhu, Zeyue Xue, Zhenchao Jin, Xian Liu, Jingzhen He, Ziwei Liu, and Lequan Yu. Make-a-volume: Leveraging latent diffusion models for cross-modality 3d brain mri synthesis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 592–601. Springer, 2023.