

# Towards the Discovery of Down Syndrome Brain Biomarkers Using Generative Models

Jordi Malé<sup>1</sup>(✉) , Juan Fortea<sup>2</sup> , Mateus Rozalem Aranha<sup>2</sup> , Yann Heuzé<sup>3</sup> , Neus Martínez-Abadías<sup>4</sup> , and Xavier Sevillano<sup>1</sup> 

<sup>1</sup> HER - Human-Environment Research Group, La Salle - URL, Barcelona, Spain  
*jordi.male@salle.url.edu*

<sup>2</sup> Sant Pau Memory Unit, Hospital de Sant Pau i la Santa Creu, Barcelona, Spain

<sup>3</sup> Univ. Bordeaux, CNRS-PACEA, UMR 5199, Pessac, France

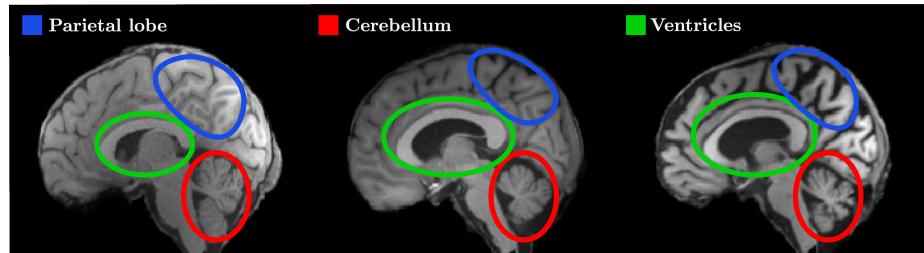
<sup>4</sup> Departament de Biologia Evolutiva, Ecología i Ciències Ambientals (BEECA),  
Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

**Abstract.** Brain imaging has allowed neuroscientists to analyze brain morphology in genetic and neurodevelopmental disorders, such as Down syndrome, pinpointing regions of interest to unravel the neuroanatomical underpinnings of cognitive impairment and memory deficits. However, the connections between brain anatomy, cognitive performance and comorbidities like Alzheimer's disease are still poorly understood in the Down syndrome population. The latest advances in artificial intelligence constitute an opportunity for developing automatic tools to analyze large volumes of brain magnetic resonance imaging scans, overcoming the bottleneck of manual analysis. In this study, we propose the use of generative models for detecting brain alterations in people with Down syndrome affected by various degrees of neurodegeneration caused by Alzheimer's disease. To that end, we evaluate state-of-the-art brain anomaly detection models based on Variational Autoencoders and Diffusion Models, leveraging a proprietary dataset of brain magnetic resonance imaging scans. Following a comprehensive evaluation process, our study includes several key analyses. First, we conducted a qualitative evaluation by expert neuroradiologists. Second, we performed both quantitative and qualitative reconstruction fidelity studies for the generative models. Third, we carried out an ablation study to examine how the incorporation of histogram post-processing can enhance model performance. Finally, we executed a quantitative volumetric analysis of subcortical structures. Our findings indicate that some models effectively detect the primary alterations characterizing Down syndrome's brain anatomy, including a smaller cerebellum, enlarged ventricles, and cerebral cortex reduction, as well as the parietal lobe alterations caused by Alzheimer's disease. These results provide preliminary evidence supporting the automatic, data-driven discovery of brain biomarkers for Down syndrome and its associated comorbidities.

**Keywords:** Generative Models · Magnetic Resonance Imaging · Brain Alteration Detection · Down syndrome · Autoencoder · Diffusion Models.

## 1 Introduction

Brain imaging techniques, especially magnetic resonance imaging (MRI), are essential for studying the complex neurocognitive phenotype of Down syndrome (DS) and its related comorbidities [7]. Neuroimaging studies have revealed that persons with DS show a reduced overall brain volume from birth, with disproportionately smaller hippocampus and cerebellum, malformations in the corpus callosum and ventriculomegaly [19], along with premature brain aging marked by accelerated brain volume loss and progressive atrophy [13]. These structural brain changes are linked to cognitive and functional impairments from birth and early-onset dementia in adulthood, as individuals with Down syndrome older than 40 years old are at a high risk of developing Alzheimer's disease (AD) due to accelerated neurodegeneration [6, 19]. For illustration purposes, Figure 1 shows an example of the Mid Sagittal Plane (MSP) of the brain MRI scan of: *i*) a person without DS (control euploid, or EU), *ii*) a person with DS with no signs of neurodegeneration, and *iii*) a person with DS with advanced signs of dementia. The image highlights the alterations caused by DS (enlarged ventricles and small cerebellum), and AD (deteriorated parietal lobe).



**Fig. 1:** Brain anatomy differences in the Mid Sagittal Plane (MSP) of three individuals: a control euploid (left), a person with DS with no signs of neurodegeneration (center), and a person with DS with advanced signs of dementia (right).

Despite advancements, the links between cognitive performance and brain anatomy, and between neuroinflammation and comorbidities like AD, remain unclear in DS. The rise of artificial intelligence (AI) offers the potential to develop automated techniques to assist clinicians in identifying disparities in brain anatomy between DS and EU subjects that lead to the discovery of diagnostic and prognostic brain biomarkers.

While supervised deep learning methods show great promise in brain analysis [15] and lesion detection [9, 10], they require large, annotated MRI datasets, which are difficult to obtain due to privacy and ethical issues, and to the labor-intensive nature of annotation. Furthermore, they tend to generalize poorly beyond the learned labels [20]. To address these limitations, unsupervised, weakly-supervised and self-supervised AI techniques are gaining interest for both detecting brain alterations [24] and generating synthetic brain images [18].

Autoencoders (AE) [21] are one of the most promising unsupervised AI approaches for detecting brain alterations on MRI. When trained on large, diverse samples of brain MRI scans of a reference group, AEs learn a reliable model of that group’s brain anatomy, which can then be used to detect alterations in new MRI scans [2]. Diffusion models (DM) have also been used for pixel-wise anomaly detection [26], as autoencoder models are often complicated to train and have difficulties to preserve the finest details of the images. These weakly supervised models rely only on image-level labels for training (for instance, healthy vs. pathological), and can generate high-quality anomaly maps.

In this work, we apply three state-of-the-art anomaly detection techniques based on generative models –Vector Quantized Variational Autoencoders (VQ-VAE) [16], Reverse Autoencoders (RAE) [2], and Denoising Diffusion Implicit Models (DDIM) [26], which have been previously used to detect severe alterations such as tumors– to detect subtle 2D brain alterations in MRI scans of individuals with DS. Our aim is to enhance the understanding of brain anatomical features in DS, paving the way for the discovery of biomarkers related to this syndrome and its associated comorbidities like AD. To achieve this, we leverage a proprietary dataset containing brain MRI scans of EU subjects and individuals with DS affected by varying degrees of neurodegeneration.

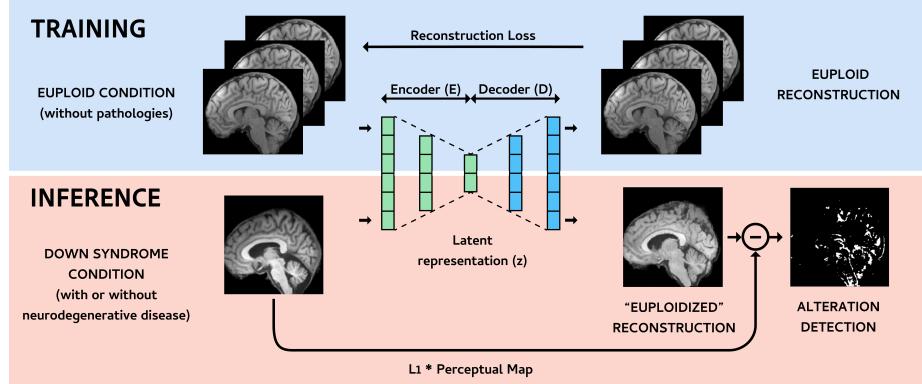
To validate the biological significance of the detected brain alterations and the effectiveness of the methods used, we conduct several evaluations: *i*) a qualitative evaluation by expert neuroradiologists, *ii*) a quantitative and qualitative evaluation of reconstruction fidelity on euploid control data, *iii*) an ablation study on the impact of histogram post-processing on model performance, and *iv*) a quantitative volumetric analysis using SynthSeg [4] to identify and quantify the regions most relevant to DS brain anatomy characterization.

The main contributions of this work are the following. First, to the best of our knowledge, this study pioneers the application of generative models for detecting brain alterations associated to different stages of neurodegeneration in DS. Second, we operate on mid-saggital plane views of the brain, in contrast to classic axial views used in anomaly detection, which entails a greater anatomical complexity. And third, we introduce a histogram matching post-processing step to enhance the quality of anomaly maps.

## 2 Generative Models for DS Brain Alteration Discovery

### 2.1 Variational Autoencoders

The central concept behind the usage of variational autoencoders for unsupervised brain alteration discovery in DS relies on two core principles: first, the model is trained exclusively on EU subjects, learning the anatomy distribution of euploid brains; second, a brain MRI of a person with DS is fed into the trained network, resulting in a reconstruction of an “euploidized” version of the input. These reconstructions are then histogram matched to the input image to reduce intensity heterogeneity. To compute the anomaly map, the reconstructions are



**Fig. 2:** Pipeline for DS brain alteration detection through Variational Autoencoders. L1 corresponds to Mean Absolute Loss.

compared to the inputs to detect alterations, combining pixel-wise and perceptual differences [27], along with histogram equalization, as proposed in [2]. Figure 2 illustrates this procedure.

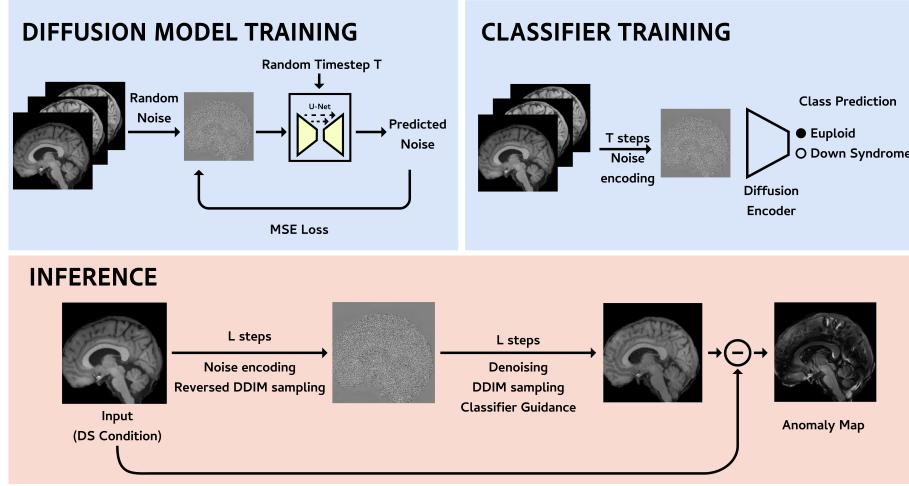
**Vector Quantized Variational Autoencoders (VQ-VAE)** [16] have demonstrated their ability to project high-resolution images into compressed latent representations [25]. They have been successfully used for unsupervised anomaly detection and other high-resolution reconstruction tasks, such as serving as an encoder for latent diffusion models.

VQ-VAEs enhance the representation learning capabilities of standard VAEs by introducing a discrete latent space. This discrete space provides more robust encoding and mitigates issues like posterior collapse, which are commonly observed in traditional VAEs.

In the VQ-VAE framework, the encoder  $E$  projects the input image  $\mathbf{x} \in \mathbb{R}^{H \times W \times D}$  into a latent representation space  $\mathbf{z} \in \mathbb{R}^{h \times w \times d \times n_z}$ , where  $n_z$  is the dimensionality of the latent embedding vector. Each spatial code  $\mathbf{z}_{ijk} \in \mathbb{R}^{n_z}$  is then quantized to its nearest vector  $e_k \in \mathbb{R}^{n_z}$  from a codebook containing  $K$  vectors, resulting in the quantized latent representation  $\mathbf{z}_q$ . The elements of the codebook are learned online, along with the other model parameters. Based on the quantized latent space, a decoder  $G$  reconstructs the input image  $\hat{\mathbf{x}} \in \mathbb{R}^{H \times W \times D}$ .

The VQ-VAE framework employs two main types of losses: an  $L1$  loss for reconstruction fidelity and a commitment loss to ensure the encoder outputs remain close to the codebook entries. This commitment loss prevents the codebook from becoming too sparsely populated and maintains robust quantization.

In our approach, we trained a VQ-VAE exclusively on euploid control data, focusing on minimizing the reconstruction error and maintaining high-quality latent representations through the combination of these losses.



**Fig. 3:** Pipeline for DS brain alteration detection using Diffusion Models.

**Reverse Autoencoders (RAE)** [2] have demonstrated effectiveness in mitigating reconstruction errors by incorporating a reversed multi-scale embedding loss into the encoder and computing anomaly scores based on residual and perceptual differences. In the evaluated approach, the architecture was trained using the Evidence Lower Bound (ELBO) for both the encoder and decoder, along with a reversed embedding similarity coefficient to ensure that the input representations align with the embeddings of the generated reconstructions. Inspired by knowledge distillation methods [23], this reverse comparison is conducted at multiple levels.

Thus, the autoencoder was trained to minimize the learning objectives of the encoder (Equation 1) and the decoder (Equation 2).

$$L_{E_\theta}(x, z) = ELBO(x) - \frac{1}{\alpha} (\exp(\alpha ELBO(D_\theta(z))) + \lambda L_{Reversed}(x)), \quad (1)$$

$$L_{D_\theta}(x, z) = ELBO(x) + \gamma ELBO(D_\theta(z)) \quad (2)$$

where  $L_{reversed}$  is the reconstruction embedding error,  $\lambda$  was empirically set to  $5 \times 10^{-3}$ , and  $\gamma$  and  $\alpha$  are hyperparameters set to 0.5 [2].

## 2.2 Diffusion Models

The fundamental principle of Diffusion Models (DM) for automatic DS brain alteration discovery is image-to-image translation [11], which involves transforming a brain image of a person with DS into a brain image of a euploid person. For effective detection, it is essential that only the altered regions are modified, while

the rest of the image remains unchanged. The anomaly map is then defined by the differences between the original and the translated images.

Diffusion Models (DM) have been recently applied to various tasks in medical imaging [11], demonstrating excellent performance in anomaly detection, among other applications.

In this work, we employed denoising diffusion implicit models (DDIM) and followed the methodology presented in [26]. We trained a DDIM on a dataset containing images of EU and DS subjects. For evaluation, we defined a noise level  $L \in \{1, \dots, T\}$  and a gradient scale  $s$ . Given an input image  $x$ , we encoded it to a noisy image  $x_L$  for  $t \in \{0, \dots, L-1\}$ . This iterative noising process allowed us to induce anatomical information of the input image. Subsequently, the image was denoised for  $t \in \{L, \dots, 1\}$  steps. We applied classifier guidance as introduced in [5] to steer the image generation toward the desired euploid class  $h$ . For this purpose, we trained a classifier network  $C$  on the noisy images  $x_t$  for  $t \in \{1, \dots, T\}$  to predict the class label of  $x$ . During the denoising process, the scaled gradient  $s\nabla_{x_t} \log C(h|x_t, t)$  of the classifier was used to update  $\varepsilon_\theta(x_t, t)$ . This iterative noising and denoising scheme is described in [26]. We generated an image  $x_0$  of the desired class  $h$  that retained the basic structure of  $x$ . The anomaly map was then defined by the difference between  $x$  and  $x_0$ . The choice of the noise level  $L$  and the gradient scale  $s$  was crucial for balancing detail-preserving image reconstruction and the freedom for translation to a euploid subject. Figure 3 presents a visual schematic of the training and inference processes of the diffusion model for discovering DS alterations.

### 3 Experimental Setup

Different datasets were used to train the proposed models. MSP images of 1,113 structural T1 MRI scans of euploid brains that are publicly available from the Human Connectome Project (HCP) were used to train the VQ-VAE and the RAE<sup>5</sup>. This dataset was divided into 90% for training (1000 subjects) and 10% for evaluation (113 subjects). In contrast, the DDIM and the classifier were trained using the HCP dataset and 93 T1 MRI scans of subjects with DS obtained from a Philips 3 Tesla X Series Achieva scanner and provided by Hospital Sant Pau Memory Unit (Barcelona, Spain). The study was approved by the Sant Pau Hospital Research Ethics Committee, following the standards for medical research in humans recommended by the Declaration of Helsinki. All participants or their legally authorised representative gave written informed consent before enrolment.

Models were evaluated using multiple datasets. For the EU control reconstruction fidelity evaluation, we tested the methods using three different datasets

---

<sup>5</sup> Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

of control euploid brain structural T1 MRI scans: *i*) 160 scans provided by Hospital Sant Pau Memory Unit (Barcelona, Spain), *ii*) 338 scans from the OASIS-3 dataset [12], *iii*) 506 scans from the IXI dataset [14] and *iv*) 113 scans from the HCP dataset (the 10% of the dataset). For the DS alteration detection, models were evaluated using scans of individuals with DS, which were categorized into three subgroups based on their degree of neurodegeneration: *i*) 61 DS subjects without signs of AD (no AD, mean age: 38 years), *ii*) 11 DS subjects with incipient signs of dementia (prodromal AD, mean age: 52 years), and *iii*) 12 DS subjects with advanced signs of dementia (AD, mean age: 52 years). Nine subjects were excluded from the evaluation due to uncertain neurodegeneration diagnosis.

All the MRI scans went through a preprocessing pipeline comprising bias correction, skull-stripping using Synthstrip [8], and affine registration to the MNI-152 template. From each 3D volume, we extracted the MSP (the MRI slice that separates the brain into two almost-identical hemispheres) automatically through a multi-scale search algorithm that finds the plane that maximizes brain symmetry measured in terms of cross-correlation [22]. This algorithm [15] employs a two-scale process: initially, it evaluates the cross-correlation of planes at a  $\frac{1}{4}$  scale of the volume, then refines the search using the full volume. The multi-scale search ultimately identifies the MSP as the plane that maximizes brain symmetry on both sides.

To improve the model’s generalization to new data, data augmentation was employed during training through affine rotations of  $\pm 10^\circ$ . Additionally, to address the significant class imbalance in the Down syndrome (DS) subject evaluations, all available DS mid-sagittal plane (MSP) images were used, along with the two adjacent sagittal slices on either side. This approach resulted in five slices per volume, yielding a total of 465 testing images (93 DS subjects with 5 slices per subject).

The VQ-VAE was implemented using the MONAI Generative framework<sup>6</sup> [17]. The RAE was implemented based on the public official implementation<sup>7</sup>). The DDIM was implemented using the MONAI framework and the public official implementation<sup>8</sup> [17]. Our code was developed using Python 3.9 and PyTorch, with experiments conducted on NVIDIA Tesla V100 GPUs.

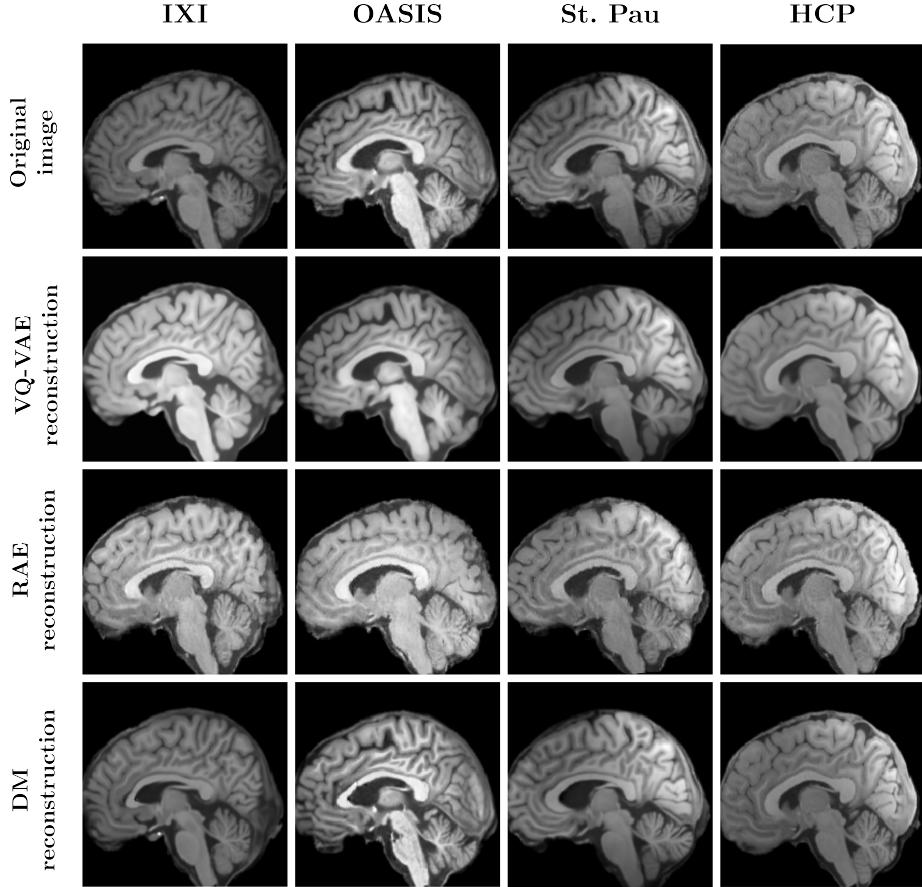
The VQ-VAE was trained for 1000 epochs, and the RAE was trained for 2000 epochs, both using the Adam optimizer. The VQ-VAE was trained with a fixed learning rate of  $5E - 4$ , while the RAE was trained with an initial learning rate of  $5E - 4$ , but was gradually decreased to  $1E - 6$  throughout training. The DM and the classifier were trained for 4000 epochs, also using the Adam optimizer, with a fixed learning rate of  $2E - 5$ .

---

<sup>6</sup> <https://github.com/Project-MONAI/GenerativeModels>

<sup>7</sup> <https://github.com/ci-ber/RA>

<sup>8</sup> <https://gitlab.com/cian.unibas.ch/diffusion-anomaly>



**Fig. 4:** Qualitative example of the models’ performance on euploid subjects.

## 4 Results

### 4.1 Evaluation of euploid brain anatomy reconstruction

The following experiments evaluate the reconstruction of EU brain anatomy in both qualitative and quantitative terms.

Figure 4 presents qualitative results of the trained models on the euploid datasets mentioned in Section 3. These results demonstrate that all models preserve the anatomy of EU subjects without altering any regions, as the models are trained on subjects from the same distribution. The diffusion model is evaluated by performing image-to-image translation on a EU subject, and it maintains the brain’s anatomy throughout the noising-denoising process.

To quantitatively assess the fidelity of the reconstructions, Table 1 presents the Structural Similarity Index Measure (SSIM) and Mean Squared Error (MSE)

**Table 1:** Quantitative evaluation of the generative models on euploid subjects datasets, measuring the performance in terms of the SSIM and MSE metrics.

Method	OASIS		IXI		St. Pau		HCP	
	SSIM ↑	MSE ↓						
VQ-VAE	0.748	0.006	0.724	0.003	0.725	0.002	0.755	0.0010
VQ-VAE + HM	<b>0.921</b>	<b>0.001</b>	<b>0.905</b>	<b>0.002</b>	<b>0.919</b>	<b>0.001</b>	0.938	0.0005
RAE	0.512	0.020	0.492	0.020	0.604	0.008	0.605	0.0080
RAE + HM	0.512	0.015	0.491	0.018	0.613	0.0060	0.614	0.0060
DM	0.769	0.009	0.785	0.006	0.854	0.005	0.930	0.0006
DM + HM	0.899	0.004	0.902	0.005	0.901	0.004	<b>0.973</b>	<b>0.0003</b>

of all models across different datasets. SSIM was chosen as it is sensitive to changes in image structure, making it suitable for evaluating the preservation of anatomical details in reconstructed brain images. MSE, on the other hand, offered a straightforward measure of pixel-wise differences, allowing for the detection of overall reconstruction accuracy. These complementary metrics together provide a robust assessment of model performance in maintaining both structural and quantitative integrity of the brain images. For this analysis, a subset of 10 subjects from each dataset (40 subjects in total) was used for inference due to time constraints. Additionally, we studied the effects of histogram matching to enhance the reconstruction’s output and maintain the primary intensities of the input image. Histogram matching was performed using the model’s output as the source histogram and the model’s input as the target histogram. This technique improved both SSIM and MSE across all models, with notable improvements in the VQ-VAE and RAE models.

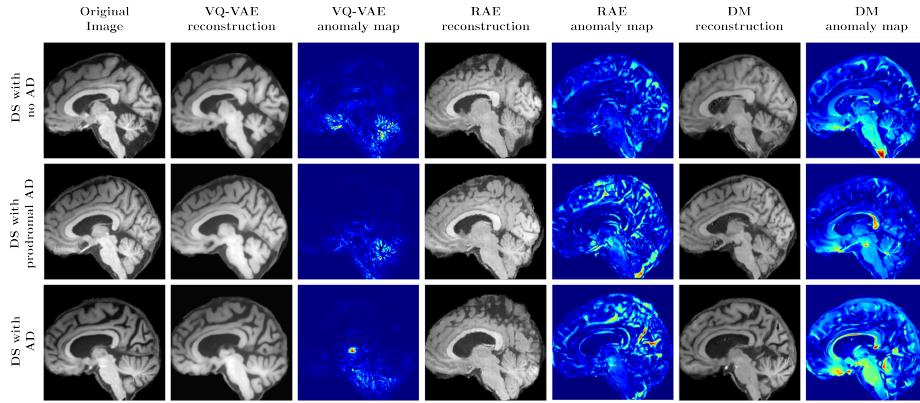
To thoroughly investigate how histogram matching affects model performance, we conducted ablation studies to evaluate its impact on reconstruction fidelity in the autoencoder-based models. This study is presented in Table 2. We implemented four variants per model: *i*) the original model without any modification (RAE, VQ-VAE), *ii*) a new model trained with data matched to a template histogram (HRAE, HVQ-VAE), *iii*) the original model with histogram matching applied from the model’s output to the model’s input (RAE + HIST M, VQ-VAE + HIST M), and *iv*) a new model trained with data matched to a template histogram and with histogram matching from the model’s output to the model’s input (HRAE + HIST M, HVQ-VAE + HIST M). Results show that applying various histogram-based techniques can significantly improve the model’s reconstruction fidelity, thereby enhancing performance in both SSIM and MSE.

#### 4.2 Evaluation of DS brain alteration detection

The following experiments evaluate the capability of the three generative models to detect brain alterations in individuals with DS affected by different degrees of neurodegeneration.

**Table 2:** Ablation study evaluating the impact of histogram modifications on the performance of various UAD architectures across euploid datasets. Performance is measured using SSIM and MSE metrics.

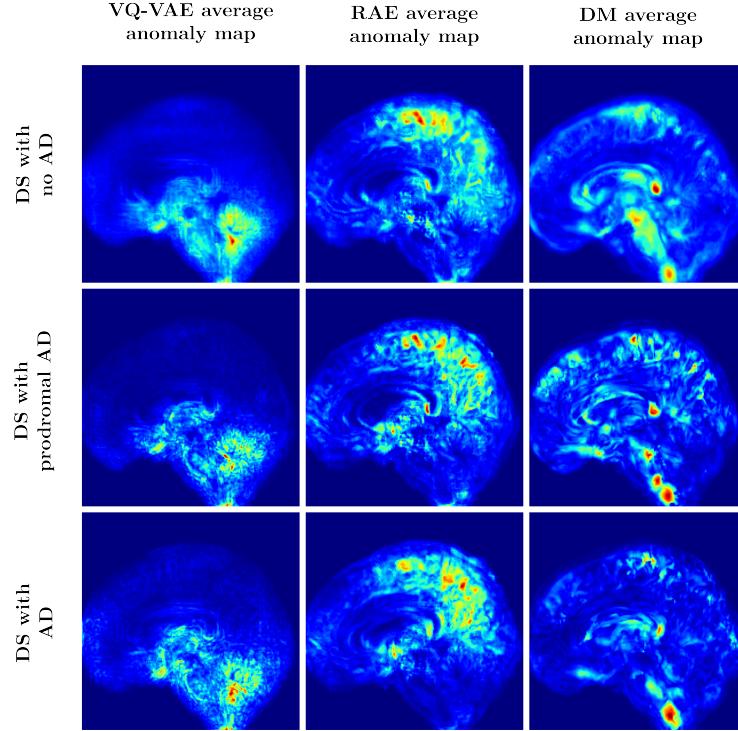
Method	OASIS		IXI		St. Pau		HCP	
	SSIM ↑	MSE ↓						
RAE	0.513	0.0198	0.496	0.0220	0.520	0.0170	0.596	0.0080
HRAE	0.525	0.0230	0.514	0.0250	0.542	0.0188	0.624	0.0066
RAE + HIST M	0.512	0.0150	0.496	0.0166	0.520	0.0134	0.604	0.0066
HRAE + HIST M	0.550	0.0113	0.523	0.0135	0.540	0.0108	0.622	0.0059
VQ-VAE	0.756	0.0050	0.742	0.0040	0.743	0.0030	0.770	0.0010
HVQ-VAE	0.759	0.0049	0.747	0.0038	0.773	0.0030	0.826	0.0010
VQ-VAE + HIST M	0.907	<b>0.0014</b>	<b>0.904</b>	0.0015	0.907	<b>0.0010</b>	0.934	<b>0.0005</b>
HVQ-VAE + HIST M	<b>0.908</b>	0.0015	0.894	<b>0.0010</b>	<b>0.908</b>	<b>0.0010</b>	<b>0.938</b>	<b>0.0005</b>



**Fig. 5:** Qualitative example of the models’ performance across subjects with DS affected by different degrees of neurodegeneration: no AD (top row), prodromal AD (center row) and established AD (bottom row).

First, Figure 5 presents qualitative results of “euploidized” reconstructions and anomaly maps between EU and DS subjects. Each row of the figure displays the results for each DS subgroup (no AD, prodromal AD and AD). From left to right, we portray the original MRI scan of a particular subject, the reconstruction and anomaly map obtained through VQ-VAE, the reconstruction and anomaly map obtained through RAE, and the reconstruction and anomaly map obtained through DM. Then, Figure 6 presents the averaged anomaly maps obtained for each DS subgroup by RAE and DM, using all subjects in each subgroup.

To qualitatively validate these results, two neurologists and radiologists with high expertise in DS from Sant Pau Hospital (Barcelona) analyzed the reconstructions and the anomaly maps. Their analysis confirmed that both the RAE and DM successfully generate “euploidized” reconstructions of the DS subjects, with increased cerebellum and decreased ventricles on all three DS groups, as



**Fig. 6:** Average anomaly maps across subjects with DS affected by different degrees of neurodegeneration: no AD (top row), prodromal AD (center row) and established AD (bottom row).

these are two of the main features of DS brain anatomy. On the other hand, they stated VQ-VAE reconstructions did not produce “euploidized”, as the model reconstructs almost the same input image. Moreover, the experts highlighted that DM reconstructions preserved better fine anatomical details of the brain, whereas RAE struggled with detecting fine-grain anomalies. This suggests the superior performance of DM in maintaining the integrity of brain structures, as confirmed in the evaluation in Section 4.1.

However, the DM average anomaly maps highlight differences in the brain stem between EU and DS subjects. These differences are not true anatomical alterations; instead, they arise because the model struggles to reconstruct the region depending on the intensity of the input image, leading to high pixel-wise discrepancies that do not correspond to actual anatomical differences. Investigating how the methods in Table 2 can enhance the diffusion model’s performance could significantly improve the quality of the average anomaly maps. This is particularly important because the DM reconstructions already outperform those of the other models, as shown in Figure 5.

**Table 3:** Top-12 average volumes (in mm<sup>3</sup>) of subcortical regions showing the greatest differences between Down Syndrome (DS) and Euploid (EU) subjects. The DS subjects are further divided into subgroups based on the presence of Alzheimer’s Disease (AD) neurodegeneration.

Region	Euploid	Down syndrome		
		No AD	Prodromal AD	AD
L inferior lateral ventricle	895.7	1178.4 (32%)	2155.8 (141%)	2530.7 (183%)
L lateral ventricle	16094.5	21115.9 (31%)	32380.5 (101%)	34557.2 (115%)
R inferior lateral ventricle	863.5	1100.2 (27%)	2190.0 (154%)	2473.9 (187%)
R lateral ventricle	14694.6	18468.8 (26%)	26112.0 (78%)	30709.1 (109%)
L cerebellum white matter	25148.2	19576.1 (-22%)	20434.9 (-19%)	18534.4 (-26%)
R cerebellum white matter	25288.2	19762.5 (-22%)	20626.1 (-18%)	18758.7 (-26%)
R putamen	7274.4	8641.8 (19%)	7956.2 (9%)	6594.8 (-9%)
L putamen	7499.4	8801.7 (17%)	8095.3 (8%)	6852.4 (-9%)
L cerebellum cortex	67715.1	57297.9 (-15%)	59742.7 (-12%)	50625.0 (-25%)
R cerebellum cortex	68025.5	58030.0 (-15%)	58642.9 (-14%)	51427.8 (-24%)
L hippocampus	5868.5	5329.8 (-9%)	4621.2 (-21%)	3781.7 (-36%)
R hippocampus	6036.3	5531.5 (-8%)	4913.5 (-19%)	3835.1 (-37%)

More interestingly, the experts remarked that the evaluated models effectively detected brain alterations between DS subgroups caused by different degrees of neurodegeneration. Indeed, anomaly maps show the presence of alterations in the parietal lobes of prodromal AD and AD subgroups, which is one of the main brain areas affected by Alzheimer’s disease [1].

For further confirmation of these results, we validated the biological significance of these findings by conducting an automatic volumetric analysis of subcortical structures using Synthseg [4] on 84 T1 MRI scans of DS subjects (divided into three groups based on neurodegeneration diagnosis) and 77 T1 MRI scans of EU subjects, all obtained from the same scanner to ensure data consistency. Table 3 presents the average volumes (in mm<sup>3</sup>) of the 12 subcortical structures that have higher relative volume differences between DS and EU subjects. For each DS subgroup, we present the relative volume difference (in %) with respect to the EU subjects.

We observe that the largest differences are congruent with the anomaly maps, focusing on the ventricles, cerebellum and hippocampus. This underscores the reliability and accuracy of the generative models in detecting and characterizing these alterations.

## 5 Conclusions

The primary objective of this study was to take the first steps towards automatic, unsupervised, and data-driven biomarker definition by comparing euploid subjects with individuals with Down syndrome. To that end, we have evaluated three state-of-the-art generative models for detecting anomalies in brain MRI

scans, targeting individuals with Down syndrome at three different neurodegeneration stages. Moreover, a study on the impact of histogram processing on these models has been conducted, resulting in improved quality of the model's reconstructions.

Three generative model-based anomaly detection approaches were compared: a vector-quantized variational autoencoder, a reverse autoencoder, and a diffusion model. These models were evaluated using euploid data from different datasets to assess the fidelity of EU reconstructions both qualitatively and quantitatively. Additionally, they were tested on a proprietary dataset of brain MRI scans, with qualitative evaluations by expert neuroradiologists and quantitative assessments through a subcortical volumetric analysis to determine the biological significance of our findings.

Our experimental results demonstrate that models effectively detect regions with diagnostic potential by comparing reconstructions with the original images. Additionally, diffusion models surpass autoencoders in preserving fine anatomical details and identifying subtle anomalies. This highlights the clinical value of diffusion models, particularly in applications where maintaining the integrity of brain structures is crucial. Notably, different models detected different regions, suggesting that further research is needed to explore how these models can be used in tandem to detect subtle differences, as proposed in [3]. This could involve combining the strengths of each model, such as developing an ensemble approach that integrates the outputs of multiple models to improve detection accuracy and robustness.

Furthermore, our results indicate that leveraging traditional image processing techniques can significantly enhance the performance of generative models for image reconstruction. Employing various histogram-based methods has demonstrated improvements in both the reconstructions and anomaly maps.

Future research will focus on several areas to enhance the models and broaden their applicability to various patient groups. First, we will explore new datasets with a greater number of Down syndrome subjects to improve the models' performance and generalization ability. Additionally, we will apply the models to identify biomarkers for other conditions, such as bipolar disorder and schizophrenia. As our results demonstrate, studying how traditional image processing techniques can complement generative models is also crucial. Furthermore, we will develop three-dimensional extensions of the models to define biomarkers based on overall volumes, which will be significantly more useful than the current approach that only utilizes the mid-sagittal plane. Finally, the study of the models using EU subjects has revealed some model-induced errors, which should be considered when interpreting differences in Down Syndrome patients.

## Acknowledgements

This work was partly supported by Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) of the Generalitat de Catalunya (2021 SGR01396, 2021 SGR00706), Agencia Española de Investigación (PID2020-113609RB-C21/AEI/

10.13039/501100011033), the Fondation Jerome Lejeune under grant 2020b cycle-Project No.2001, and the Joan Oró grant (FI2024) from the DRU of the Generalitat de Catalunya and the European Social Fund (2024 FI-200014).

## References

1. Ahulló-Fuster, M.A., Ortiz, T., Varela-Donoso, E., Nacher, J., Sánchez-Sánchez, M.L.: The parietal lobe in alzheimer's disease and blindness. *Journal of Alzheimer's Disease* **89**(4), 1193–1202 (2022). <https://doi.org/10.3233/JAD-220498>, <https://content.iospress.com/articles/journal-of-alzheimers-disease/jad220498>
2. Bercea, C.I., Wiestler, B., Rueckert, D., Schnabel, J.A.: Generalizing unsupervised anomaly detection: Towards unbiased pathology screening. In: *Medical Imaging with Deep Learning* (2023)
3. Bercea, C.I., Wiestler, B., Rueckert, D., Schnabel, J.A.: Reversing the abnormal: Pseudo-healthy generative networks for anomaly detection. *arXiv preprint arXiv:2303.08452* (2023)
4. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., Iglesias, J.E.: Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. *Medical Image Analysis* **86**, 102789 (2023). <https://doi.org/https://doi.org/10.1016/j.media.2023.102789>, <https://www.sciencedirect.com/science/article/pii/S1361841523000506>
5. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis (2021)
6. Fortea, J., Vilaplana Martínez, E., Carmona-Iragui, M., Benejam, B., Videla, L., Barroeta, I., Fernández, S., Altuna Azkargorta, M., Pegueroles, J., Montal, V., Valldeneu, S., Gimenez, S., González-Ortiz, S., Muñoz Llahuna, L., Estellés, T., Illán-Gala, I., Belbin, O., Camacho, V., Wilson, L., Lleó, A.: Clinical and biomarker changes of alzheimer's disease in adults with down syndrome: a cross-sectional study. *The Lancet* **395**, 1988–1997 (06 2020). [https://doi.org/10.1016/S0140-6736\(20\)30689-9](https://doi.org/10.1016/S0140-6736(20)30689-9)
7. Hamner, T., Udhmani, M.D., Osipowicz, K., Lee, N.R.: Pediatric brain development in down syndrome: A field in its infancy. *Journal of the International Neuropsychological Society* **24**, 966 – 976 (2018), <https://api.semanticscholar.org/CorpusID:44074892>
8. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. *NeuroImage* **260**, 119474 (2022). <https://doi.org/https://doi.org/10.1016/j.neuroimage.2022.119474>, <https://www.sciencedirect.com/science/article/pii/S1053811922005900>
9. Kamnitsas, K., Ferrante, E., Parisot, S., Ledig, C., Nori, A.V., Criminisi, A., Rueckert, D., Glocker, B.: Deepmedic for brain tumor segmentation. In: *Brain-Les@MICCAI* (2016), <https://api.semanticscholar.org/CorpusID:23448264>
10. Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B.: Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis* **36**, 61–78 (2017). <https://doi.org/https://doi.org/10.1016/j.media.2016.10.004>, <https://www.sciencedirect.com/science/article/pii/S1361841516301839>
11. Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Fayyaz, M., Hacihamoglu, I., Merhof, D.: Diffusion models for medical image analysis: A comprehensive survey (2023)

12. LaMontagne, P.J., Benzinger, T.L., Morris, J.C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A.G., Raichle, M.E., Cruchaga, C., Marcus, D.: Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. medRxiv (2019). <https://doi.org/10.1101/2019.12.13.19014902>
13. Lao, P., Handen, B., Betthauser, T., Cody, K., Cohen, A., Tudorascu, D., Stone, C., Phd, J., Johnson, S., Klunk, W., Christian, B.: Imaging neurodegeneration in down syndrome: brain templates for amyloid burden and tissue segmentation. *Brain Imaging and Behavior* **13**, 1–9 (04 2019). <https://doi.org/10.1007/s11682-018-9888-y>
14. London, I.C., London, U.C., of Edinburgh, U.: Information extraction from images (ixi) dataset. <https://brain-development.org/ixi-dataset/>, accessed: 2024-05-21
15. Malé, J., Heuzé, Y., Fortea, J., Abadias, N.M., Sevillano, X.: Single-landmark vs multi-landmark deep learning approaches to brain mri landmarking: a case study with healthy controls and down syndrome individuals. In: 34th British Machine Vision Conference 2023, BMVC 2023, Aberdeen, UK, November 20-24, 2023. BMVA (2023), <https://papers.bmvc2023.org/0754.pdf>
16. van den Oord, A., Vinyals, O., Kavukcuoglu, K.: Neural discrete representation learning (2018)
17. Pinaya, W.H.L., Graham, M.S., Kerfoot, E., Tudosiu, P.D., Dafflon, J., Fernandez, V., Sanchez, P., Wolleb, J., da Costa, P.F., Patel, A., Chung, H., Zhao, C., Peng, W., Liu, Z., Mei, X., Lucena, O., Ye, J.C., Tsafaris, S.A., Dogra, P., Feng, A., Modat, M., Nachev, P., Ourselin, S., Cardoso, M.J.: Generative ai for medical imaging: extending the monai framework (2023)
18. Pinaya, W.H.L., Tudosiu, P.D., Dafflon, J., da Costa, P.F., Fernandez, V., Nachev, P., Ourselin, S., Cardoso, M.J.: Brain imaging generation with latent diffusion models (2022)
19. Rodrigues, M., Nunes, J., Figueiredo, S., Campos, A., Geraldo, A.: Neuroimaging assessment in down syndrome: a pictorial review. *Insights into Imaging* **10** (12 2019). <https://doi.org/10.1186/s13244-019-0729-3>
20. Ruff, L., Kauffmann, J.R., Vandermeulen, R.A., Montavon, G., Samek, W., Kloft, M., Dietterich, T.G., Muller, K.R.: A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE* **109**(5), 756–795 (May 2021). [https://doi.org/10.1109/JPROC.2021.3052449](https://doi.org/10.1109/jproc.2021.3052449), <http://dx.doi.org/10.1109/JPROC.2021.3052449>
21. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* **323**(6088), 533–536 (Oct 1986). <https://doi.org/10.1038/323533a0>, <https://doi.org/10.1038/323533a0>
22. Ruppert, G., Teverovskiy, L., Yu, C.P., Falcão, A., Liu, Y.: A new symmetry-based method for mid-sagittal plane extraction in neuroimages. pp. 285–288 (03 2011). <https://doi.org/10.1109/ISBI.2011.5872407>
23. Salehi, M., Sadjadi, N., Baselizadeh, S., Rohban, M.H., Rabiee, H.R.: Multiresolution knowledge distillation for anomaly detection (2020)
24. Tschuchnig, M.E., Gadermayr, M.: Anomaly Detection in Medical Imaging - A Mini Review, p. 33–38. Springer Fachmedien Wiesbaden (2022). [https://doi.org/10.1007/978-3-658-36295-9\\_5](https://doi.org/10.1007/978-3-658-36295-9_5), [http://dx.doi.org/10.1007/978-3-658-36295-9\\_5](http://dx.doi.org/10.1007/978-3-658-36295-9_5)
25. Tudosiu, P.D., Pinaya, W.H.L., Graham, M.S., Borges, P., Fernandez, V., Yang, D., Appleyard, J., Novati, G., Mehra, D., Vella, M., Nachev, P., Ourselin, S., Car-

- doso, J.: Morphology-preserving autoregressive 3d generative modelling of the brain (2022), <https://arxiv.org/abs/2209.03177>
- 26. Wolleb, J., Bieder, F., Sandkühler, R., Cattin, P.C.: Diffusion models for medical anomaly detection (2022)
  - 27. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric (2018)