SIGMA-VAE: Sketched Isotropic Gaussian Multi-View Autoencoder for Normative Modeling

Mariam Zabihi[1,2,3][0000−0002−7083−2318], Sicheng Dai[1], Andre Schmidt[4][0000−0001−6055−8397], Katrin Preller[4], Andrew Watrous[4][0000−0002−3107−3726], Klaus Bornemann[4], James Cole[2][0000−0003−1908−5588], Andre Marquand[3][0000−0001−5903−203X], Yanwu Yang[1,5][0000−0002−7547−4580], and Thomas Wolfers[1,6,7][0000−0002−7693−0621]

[1] Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen, Germany
[2] Hawkes Institute, Department of Computer Science, University College London, London, United Kingdom
[3] Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, the Netherlands
[4] Boehringer Ingelheim Pharma GmbH and Co. KG, Biberach an der Riss, Germany
[5] Max Planck Institute for Intelligent Systems, Tübingen, Germany
[6] Department of Psychology, Friedrich Schiller University of Jena, Germany
m.zabihi@ucl.ac.uk

Abstract. Normative modeling characterizes heterogeneous brain pathology by quantifying individual deviations from a healthy reference population. Variational autoencoders (VAEs) are widely used for this purpose, but latent-space irregularities can compromise distance-based deviation measures and interpretability. We propose *SIGMA-VAE*, a Sketched Isotropic Gaussian Multi-view Autoencoder for robust and interpretable normative modeling. SIGMA-VAE combines (i) Sketched Isotropic Gaussian Regularization to enforce a well-formed latent geometry with more reliable Euclidean anomaly scores, (ii) learnable sparsity via L_0 regularization for compact and interpretable representations, and (iii) an uncertainty-aware W-score that weights deviations by posterior variance to mitigate false positives due to image ambiguity and scanner confounds. We evaluate *SIGMA-VAE* on over 40,000 structural brain scans from 328 scanners, demonstrating improved sensitivity to individual-level anomalies, biologically meaningful deviations, and robust generalization to unseen scanners. In downstream discrimination tasks, *SIGMA-VAE* achieves strong performance for Alzheimer's disease, with AUCs of 0.91 ± 0.01 ($n = 207$) on a held-out test set and 0.87 ± 0.01 ($n = 585$) on an external test set acquired from previously unseen scanners. These results show that *SIGMA-VAE* enables scalable, interpretable, and robust normative modeling across views and at scale.

Keywords: Normative and Reference Class Modeling · Brain imaging · Uncertainty quantification · Latent space regularization · Sketched Isotropic Gaussian Regularization (SIGReg).

1 Introduction

Normative modeling has emerged as a powerful framework for characterizing biological heterogeneity in neurological and psychiatric disorders [1–3]. By learning a statistical reference of healthy brain structure and function, these models enable the quantification of individual-level deviations, thereby moving beyond traditional case–control group comparisons [4–6]. To scale normative modeling to high-dimensional neuroimaging data, recent work has increasingly relied on deep generative models, in particular, Variational Autoencoders (VAEs) [7–9]. In this setting, VAEs learn a low-dimensional latent representation of healthy variation and reconstruct observations from this latent space, allowing deviations from normality to be quantified either via reconstruction error or via distances in the latent space relative to a learned normative manifold.

Despite this conceptual appeal, the practical utility of VAE-based normative models is limited by several structural challenges. First, the standard VAE objective based on the Kullback–Leibler divergence often fails to enforce a globally coherent latent geometry, leading to aggregate posterior mismatch [10]. This manifests as regions of low probability density or "holes" in latent space, where distance-based anomaly scores become unreliable [11]. Second, standard VAEs typically learn dense and entangled latent representations in which individual dimensions lack clear biological interpretation, complicating the attribution of detected deviations to meaningful anatomical processes [12]. Finally, demographic factors such as age and sex, along with normative developmental and aging trajectories, often covary nonlinearly with pathology, making it challenging to disentangle healthy variation from disease-related effects within latent representations.

In this work, we propose the *Sketched Isotropic Gaussian Multi-view Autoencoder (SIGMA-VAE)*, a unified framework for robust and interpretable multi-view normative modeling. SIGMA-VAE directly addresses the limitations of existing VAE-based approaches through three complementary innovations. First, *Sketched Isotropic Gaussian Regularization (SIGReg)* imposes a spectral constraint via characteristic function matching [13], yielding a compact, hole-free latent geometry that supports valid Euclidean anomaly scoring. Second, *learnable sparsity* induced by L_0 regularization [14] enables the automatic discovery of a minimal and interpretable set of latent factors. Third, we introduce an *uncertainty-aware W-score* that weights individual deviations by epistemic uncertainty [15], reducing false positives caused by noisy or ambiguous observations.

The main contributions of this work are threefold: (i)**Geometrically valid latent representations** via spectral regularization, enabling reliable distance-based deviation measures. (ii)**Sparse, interpretable latent structures** through the automatic selection of a minimal set of meaningful dimensions. (iii)**Uncertainty-aware deviation scoring** for improved robustness and sensitivity under distribution shift.

2 Method

2.1 Overview

We formulate normative modeling as a conditional density estimation problem. Given covariates $\mathbf{c}$, we learn the distribution $p(\mathbf{x} \mid \mathbf{c})$ of neuroimaging features $\mathbf{x}$ and quantify individual atypicality via deviations from this distribution. The Conditional *SIGMA-VAE* introduces a latent variable $\mathbf{z} \in \mathbb{R}^K$ with a prior $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ to capture individual-level deviations from the normative trajectory, where $\mathbf{z} = \mathbf{0}$ corresponds to the expected anatomy given $\mathbf{c}$. Concretely, *SIGMA-VAE* (i) conditions on demographic covariates, including age and sex, to model normative developmental and aging effects; (ii) employs SIGReg [13] to enforce a geometrically well-formed latent space; and (iii) applies L_0 regularization [14] to identify a minimal and interpretable subset of latent dimensions. Fig. 1 shows the method overview.

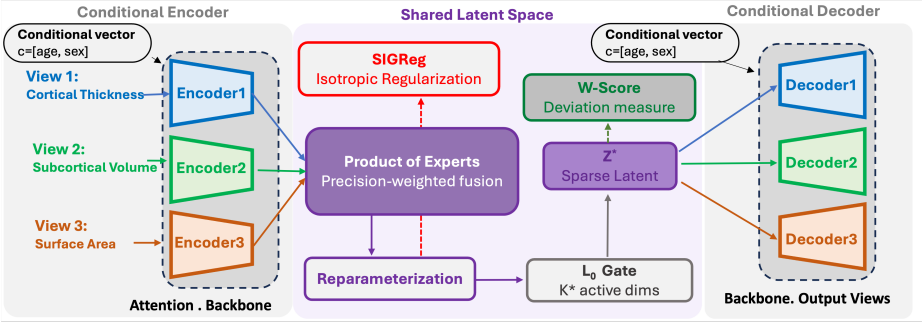

Fig. 1. Architecture of SIGMA-VAE for normative modeling.

2.2 Conditional Generative Model

Let $\mathbf{x}^{(m)} \in \mathbb{R}^{D_m}$ denote the features from view $m \in \{1, \ldots, M\}$. Each view has a dedicated encoder-decoder pair $\{E_\phi^{(m)}, D_\theta^{(m)}\}$. Features are first modulated by a learned attention mechanism, then concatenated with $\mathbf{c}$ before being encoded into a diagonal Gaussian posterior $q_\phi^{(m)}(\mathbf{z} \mid \mathbf{x}^{(m)}, \mathbf{c}) = \mathcal{N}(\boldsymbol{\mu}^{(m)}, \mathrm{diag}(\boldsymbol{\sigma}^{2(m)}))$.
Multi-View Fusion. View-specific posteriors are combined with the prior via the Product-of-Experts (PoE) [16]:

$$q_\phi(\mathbf{z} \mid \mathbf{x}^{(1:M)}, \mathbf{c}) \propto p(\mathbf{z}) \prod_{m \in \mathcal{M}} q_\phi^{(m)}(\mathbf{z} \mid \mathbf{x}^{(m)}, \mathbf{c}), \tag{1}$$

where $\mathcal{M}$ is the set of observed views. Since all factors are diagonal Gaussians, the fused posterior is available in closed form with precision $\boldsymbol{\tau}_{\text{fused}} = \mathbf{1} + \sum_m \boldsymbol{\sigma}^{-2(m)}$

and mean $\boldsymbol{\mu}_{\text{fused}} = \boldsymbol{\tau}_{\text{fused}}^{-1} \odot \sum_m \boldsymbol{\sigma}^{-2(m)} \odot \boldsymbol{\mu}^{(m)}$. The prior shrinks the fused posterior toward zero when encoders are uncertain, and the formulation naturally handles missing views. Decoders reconstruct each view as $p_\theta(\mathbf{x}^{(m)} \mid \mathbf{z}, \mathbf{c}) = \mathcal{N}(D_\theta^{(m)}([\mathbf{z}; \mathbf{c}]), \mathbf{I})$.

2.3 Regularization

SIGReg. Standard KL regularization constrains individual posteriors but not the aggregate $q(\mathbf{z})$. We replace it with SIGReg [13], which directly enforces $q(\mathbf{z}) \approx \mathcal{N}(\mathbf{0}, \mathbf{I})$ by matching empirical and theoretical characteristic functions via random projections. This ensures a globally isotropic latent geometry where Euclidean distance is a valid anomaly measure. Note that replacing KL with SIGReg means we no longer optimize a standard ELBO; instead, we minimize reconstruction error subject to a distributional constraint on the aggregate posterior. SIGReg and the PoE prior are complementary: the prior regularizes individual fusion, while SIGReg constrains population-level geometry.

L_0 **Sparsity.** We introduce learnable gates $\mathbf{g} \in [0, 1]^K$ via the Hard Concrete distribution [14], yielding gated latents $\mathbf{z}_{\text{gated}} = \mathbf{z} \odot \mathbf{g}$. An L_0 penalty on the expected number of active gates encourages the model to discover a minimal set of dimensions. Temperature annealing ensures gates become binary during training; at inference, $g_k = \mathbf{1}[\sigma(\log \alpha_k) > 0.5]$.

Training objective. The full objective is:

$$\mathcal{L} = \lambda_{\text{recon}} \sum_{m=1}^{M} \frac{1}{D_m} \|\mathbf{x}^{(m)} - \hat{\mathbf{x}}^{(m)}\|_2^2 + \lambda_{\text{SIG}}\, \mathcal{L}_{\text{SIG}} + \lambda_{L0}\, \mathcal{L}_{L0} + \lambda_{\text{var}} \left(\mathbb{E}[\log \sigma^2]\right)^2 \quad (2)$$

where the final term prevents posterior variance collapse. Training uses two phases: *structure discovery* with active L_0 and temperature annealing, followed by *fine-tuning* with frozen gates and increased reconstruction weight.

2.4 Quantifying Individual Deviations

Given the encoder mean μ_{ik} and variance σ_{ik}^2 for subject i and dimension k, we define the uncertainty-aware W-score $W_{ik} = \mu_{ik} / \sqrt{1 + \sigma_{ik}^2}$, which preserves high-confidence deviations while attenuating uncertain estimates. An overall anomaly score is computed as $A_i = \|\mathbf{W}_i\|_2$ across active dimensions. Under SIGReg, A_i^2 approximately follows a χ^2 distribution for healthy controls when encoder uncertainty is low, enabling principled threshold selection. **Anatomical Interpretation.** To map latent dimensions to brain regions, we perturb one dimension at a time from the normative reference ($\mathbf{z} = \mathbf{0}$) and compute the difference in the decoded output. This captures the full non-linear decoder response at clinically relevant magnitudes, revealing which anatomical regions are most affected by each latent dimension.

3 Experiments and Results

3.1 Implementation and Experimental Setup

Architecture and Training. The proposed multi-view framework, shown in Fig. 1, was implemented to generate conditional W-scores. Each view-specific encoder consists of a feature attention block followed by an MLP backbone with hidden dimensions [512, 256, 128], batch normalization, ReLU activations, and dropout ($p = 0.1$). Decoders mirror this structure with hidden dimensions [128, 256, 512]. The covariate vector $\mathbf{c}$ = [age, sex] is concatenated to the inputs of both the encoder and decoder networks. As shown in the Shared Latent Space (Fig. 1, center), representations are fused via a Product-of-Experts (PoE) layer, followed by an L_0 gating mechanism that identifies K^* active dimensions from a 32-dimensional raw latent space. Training is conducted in a two-phase regime: (1) joint learning of the L_0 gate and reconstruction parameters, and (2) fine-tuning of the conditional decoders. The resulting anomaly scores, defined as the L_2 norm of the active W-score vector, serve as the primary discriminating statistic for the benchmarking experiments summarized in Table 1.

Ablation study. To isolate the contributions of individual SIGMA-VAE components, we compare four model variants that share identical architectures, training schedules, and preprocessing pipelines: **(1) KL**, which replaces SIGReg with standard KL regularization at a fixed latent dimensionality ($K = 32$); **(2) KL+L_0**, which augments the KL baseline with L_0 sparsity gating; **(3) SIGReg**, which applies isotropic Gaussian regularization without sparsity at $K = 32$; and **(4) SIGMA-VAE**, the full model combining both components. Variants are evaluated concerning latent-space calibration and clinical discriminability (AUC on held-out cohorts). Reconstruction quality (mean squared error per structural view) is not reported as a comparative metric, as the effective latent dimensionality differs substantially across variants ($K^* \in \{5, 20, 32\}$). Models with a larger number of active dimensions trivially achieve lower reconstruction error, independent of latent-space geometry or normative validity. Consistent with this, mean per-view reconstruction MSE rose as effective dimensionality fell (SIGReg 0.15 and KL 0.18 at K=32; KL+L0 0.20 at K^*=20; SIGMA-VAE 0.36 at K^*=5; λ_{recon}=10), confirming that reconstruction error tracks latent capacity rather than normative validity. Latent-space calibration is assessed via empirical coverage of healthy controls at nominal $\chi^2(K^*)$ thresholds $\{80\%, 90\%, 95\%, 99\%\}$; a well-calibrated model retains the corresponding fraction of healthy subjects below each threshold. We hypothesize that neither component alone is sufficient: SIGReg without sparsity cannot maintain isotropy when all latent dimensions are required for reconstruction, while L_0 regularization without SIGReg lacks the normative calibration necessary for standardized anomaly detection. Only the full SIGMA-VAE model jointly yields a geometrically valid latent space and data-driven dimensionality, producing **w**-scores with the conservative normative properties required for reliable clinical inference.

Optimization. Models are trained using the Adam optimizer with an initial learning rate of 10^{-3} (Phase 1, number of epochs=10,000) and 10^{-5} (Phase 2,

number of epochs=2000), weight decay of 10^{-6}, and gradient clipping at a norm of 1.0. The learning rate is reduced by a factor of 0.5 when the validation loss plateaus (with a patience of 15 epochs). The batch size is 128.

Hyperparameters. Loss weights are set to $\lambda_{\text{recon}} = 10$, $\lambda_{SIG_1} = .75$, $\lambda_{SIG_2} = 15$, $\lambda_{L_0} = 5$, $\lambda_{KL} = 1$ and $\lambda_{\text{var}} = 10^{-4}$. The L_0 schedule consists of a 50-epoch warmup followed by a 100-epoch ramp-up. Temperature annealing is performed from 0.67 to 0.1 over 200 epochs. The initial latent dimensionality is set to $K = 32$.

Data. Structural MRI features are derived using FreeSurfer v7.3 and include subcortical volumes (33 regions), cortical thickness (68 parcels), and cortical surface area (68 parcels). The model is trained exclusively on healthy reference data. Evaluation is performed on two held-out datasets: (i) an *in-distribution* test set with demographic characteristics and scanners overlapping the training data, and (ii) an *out-of-distribution* external dataset with an older age distribution and entirely unseen acquisition sites and scanners. The training set comprised 20,655 healthy controls (50% male; age 26.5 ± 16.7 years, range: 3–100) acquired across 215 scanners from 113 datasets, including cohorts from ABIDE I/II [17], ADHD-200 [18], the Human Connectome Project (Young Adult and Lifespan cohorts) [19, 20], Cam-CAN [21], HBN [22], PNC [23], and 103 OpenNeuro datasets.

In-distribution test evaluation utilized 8,341 subjects (65% male; age 22.9 ± 19.2 years, range: 5–94) across 81 scanners, drawn from a subset of 19 training datasets including ABIDE [17], ADHD-200 [18], COBRE, HBN [22], and TCP.

Out-of-distribution external validation was performed on 15,419 subjects (50.2% male; age 54.5 ± 22.3 years, range: 6–97) across 113 scanners from six independent datasets not used during training: AIBL [24], CoRR/NKI-RS [25], OASIS-3[26], PPMI [27], and SRPBS-TS[28].

3.2 Results

Latent Space Geometry and Normative Calibration The L_0 regularization automatically selected $K^* = 5$ active dimensions from the initial 32-dimensional latent space. Fig. 2A shows that W-score distributions in healthy controls closely approximate $\mathcal{N}(0, 1)$ per dimension ($\mu \approx 0$, $\sigma \in [0.83, 0.96]$), with comparable alignment in held-out and external cohorts, confirming that SIGReg's marginal isotropy generalizes beyond the training distribution. The learned dimensions are largely disentangled (mean $|r| = 0.035$, max $|r| = 0.15$), suggesting they capture independent axes of brain morphological variation. Normative calibration assessed via empirical coverage (Fig. 2B) shows that SIGMA-VAE achieves the closest agreement with nominal $\chi^2(5)$ thresholds across all cohorts, while ablation variants deviate substantially, the KL baseline over-retains up to 15.5% of healthy subjects at the 80% threshold, confirming that both SIGReg and L_0 are necessary for a well-calibrated normative latent space.

Diagnostic Evaluation We assessed diagnostic classification using the area under the receiver operating characteristic curve (AUC), comparing each di-

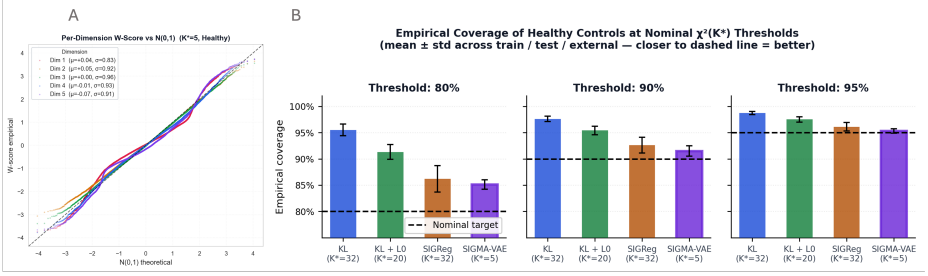

Fig. 2. Latent-space geometry and normative calibration of SIGMA-VAE (healthy controls). (A) W-score Q–Q plots against $\mathcal{N}(0,1)$ for the $K^* = 5$ active latent dimensions ($\mu \approx 0$, $\sigma \in [0.83, 0.96]$). **(B)** Empirical coverage at nominal $\chi^2(K^*)$ thresholds, averaged across training, test, and external cohorts (mean ± std).

agnostic group against healthy controls. To reduce the confounding effects of site, age, and sex, we employed a confound-aware evaluation protocol in which healthy controls were matched 1:1 to patients within each validation fold using greedy nearest-neighbor matching (exact matching for site and sex, as well as a 5-year caliper for age). For disorders with sufficient site diversity (≥ 3 sites), we employed a leave-one-site-out cross-validation scheme to explicitly assess generalization to unseen acquisition environments. For the remaining disorders, stratified 5-fold cross-validation was used, with folds jointly stratified by diagnosis and site. As the discriminative statistic, we used the anomaly score defined as the ℓ_2 norm of the W-score vector across active latent dimensions, without training an additional classifier. Model performance is reported as the mean ± standard deviation across folds and is summarized in Table 1. Across all evaluated disorders, *SIGMA-VAE* consistently outperformed all baseline models. Alzheimer's disease achieved high classification accuracy (AUC $= 0.91 \pm 0.01$), cognitive impairment showed strong discrimination despite the small sample size ($n = 15$; AUC $= 0.88 \pm 0.01$), and broader neurodegenerative conditions exhibited moderate separation (AUC $= 0.82 \pm 0.07$).

Importantly, performance generalizes well to external validation datasets acquired at independent sites and on scanners unseen during training. Discriminative accuracy remains robust for Alzheimer's disease (AUC $= 0.87 \pm 0.01$), cognitive impairment (AUC $= 0.82 \pm 0.02$), and general neurodegenerative conditions (AUC $= 0.80 \pm 0.01$), with only minor degradation relative to held-out test performance. In contrast, baseline models exhibit substantially higher variance on external data for Alzheimer's disease (standard deviation up to ± 0.16), indicating reduced stability under site shift. The consistent performance of SIGMA-VAE across datasets suggests that its normative deviation scores capture disorder-related brain alterations rather than site- or scanner-specific artifacts.

Anatomical Interpretation via Perturbation Analysis To characterize the neuroanatomical patterns encoded by each latent dimension, we varied individual

Table 1. Classification performance (AUC ± SD) for diagnostic classification. Matching was performed on site, sex, and age using either Leave-One-Site-Out cross-validation or stratified 5-fold cross-validation. N denotes the number of patients (ND: neurodegenerative). The best performance is shown in bold.

Disorder	Model	AUC Test (N)	AUC External (N)
Alzheimer's Disease	KL Baseline	0.80 ± 0.02 (207)	0.65 ± 0.16 (585)
	KL + L_0	0.82 ± 0.03 (207)	0.72 ± 0.14 (585)
	SIGReg	0.78 ± 0.01 (207)	0.69 ± 0.15 (585)
	SIGMA-VAE	**0.91 ± 0.01** (207)	**0.87 ± 0.01** (585)
Cognitive Impairment	KL Baseline	0.79 ± 0.20 (15)	0.71 ± 0.04 (185)
	KL + L_0	0.82 ± 0.20 (15)	0.72 ± 0.05 (185)
	SIGReg	0.78 ± 0.25 (15)	0.70 ± 0.04 (185)
	SIGMA-VAE	**0.88 ± 0.01** (15)	**0.82 ± 0.02** (185)
ND (General)	KL Baseline	0.71 ± 0.06 (41)	0.67 ± 0.04 (164)
	KL + L_0	0.74 ± 0.06 (41)	0.70 ± 0.02 (164)
	SIGReg	0.67 ± 0.09 (41)	0.66 ± 0.05 (164)
	SIGMA-VAE	**0.82 ± 0.07** (41)	**0.80 ± 0.01** (164)

W-scores by ±3 standard deviations while holding all others fixed and mapped the resulting changes onto subcortical ROI volumes (Fig. 3). The L_0 sparsity constraint encourages each dimension to capture a compact, non-overlapping anatomical subsystem. W_0 encodes ventricular and CSF structures; W_1 isolates the posterior fossa, with effects concentrated in the cerebellar cortex and brainstem; W_2 and W_3 reveal thalamo-striatal patterns with opposing cerebellar contributions; and W_4 captures a predominantly cerebellar axis. Parallel cortical perturbations (not shown) indicate that W_3 drives cortical thickness variation in frontal and parietal regions, while W_2 governs cortical surface area in temporal and frontal regions. Together, these results confirm that the learned latent space recovers biologically grounded and interpretable axes of brain morphometry.

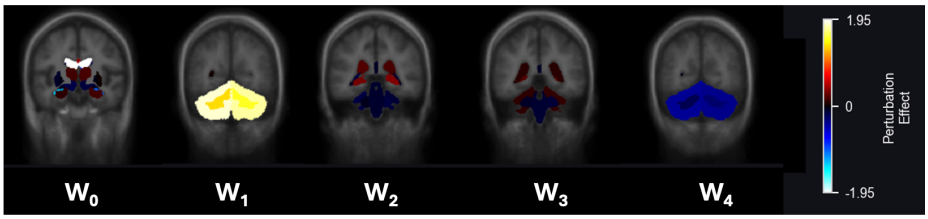

Fig. 3. Latent-Dimension-Specific Perturbation Effects on Subcortical Volumes. Subcortical perturbation maps for dimensions W_0–W_4, obtained by varying each latent dimension by ±3 SD for a representative 26-year-old male. Colors denote the direction and magnitude of reconstructed volumetric changes (blue: negative; red/yellow: positive).

4 Conclusion

We presented SIGMA-VAE, a multi-view normative modeling framework that integrates Sketched Isotropic Gaussian Regularization with L_0 sparsity gating. By training on over 40,000 scans, the model automatically discovers a low-dimensional, disentangled latent space that recovers biologically grounded neuroanatomical axes. Our results show that SIGMA-VAE yields statistically well-calibrated W-scores and achieves robust generalization to unseen scanners and external cohorts, outperforming traditional VAE baselines in clinical discrimination. By reconciling high-dimensional feature integration with neuroanatomical interpretability, this framework provides a principled foundation for large-scale, individualized anomaly detection in clinical neuroimaging. Future work will extend this approach to longitudinal trajectories and multi-modal integration.

Acknowledgments. This work is supported by funds from the AI & Data Science Fellowship program, funded by Boehringer Ingelheim, between Tübingen University and Boehringer Ingelheim.

Disclosure of Interests. The authors have no competing interests.

References

1. Marquand, A.F., Rezek, I., Buitelaar, J.K., Beckmann, C.F.: Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. Biol. Psychiatry **80**(7), 552–561 (2016)
2. Marquand, A.F., Kia, S.M., Zabihi, M., Wolfers, T., Buitelaar, J.K., Beckmann, C.F.: Conceptualizing mental disorders as deviations from normative functioning. Mol. Psychiatry **24**(10), 1415–1424 (2019)
3. Rutherford, S., Kia, S.M., Wolfers, T., Fraza, C., Zabihi, M., Dinga, R., Berthet, P., Worker, A., Verdi, S., Ruhe, H.G., Beckmann, C.F., Marquand, A.F.: The normative modeling framework for computational psychiatry. Nat. Protoc. **17**(11), 1711–1734 (2022)
4. Wolfers, T., Floris, D.L., Dinga, R., van Rooij, D., Isakoglou, C., Kia, S.M., Zabihi, M., Llera, A., Chowdanayaka, R., Kumar, V.J., Peng, H., Laidi, C., Batalle, D., Dimitrova, R., Charman, T., Loth, E., Lai, M.-C., Jones, E., Baumeister, S., Moessnang, C., Banaschewski, T., Ecker, C., Dumas, G., O'Muircheartaigh, J., Murphy, D., Buitelaar, J.K., Marquand, A.F., Beckmann, C.F.: From pattern classification to stratification: Toward conceptualizing the heterogeneity of autism spectrum disorder. Neurosci. Biobehav. Rev. **104**, 240–254 (2019)
5. Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Goyard, D., Loth, E., Charman, T., Tillmann, J., Banaschewski, T., Dumas, G., Holt, R., Baron-Cohen, S., Durston, S., Bölte, S., Murphy, D., Ecker, C., Buitelaar, J.K., Beckmann, C.F., Marquand, A.F.: Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models. Biol. Psychiatry Cogn. Neurosci. Neuroimaging **4**(6), 567–578 (2019)
6. Fraza, C.J., Dinga, R., Beckmann, C.F., Marquand, A.F.: Warped Bayesian linear regression for normative modelling of big data. NeuroImage **245**, 118740 (2021)

7. Pinaya, W.H.L., Mechelli, A., Sato, J.R.: Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders. Hum. Brain Mapp. **40**(14), 4053–4063 (2019)
8. Zabihi, M., Kia, S.M., Wolfers, T., de Boer, S., Fraza, C., Dinga, R., Arenas, A.L., Bzdok, D., Beckmann, C.F., Marquand, A.F.: Nonlinear latent representations of high-dimensional task-fMRI data: Unveiling cognitive and behavioral insights in heterogeneous spatial maps. PLoS ONE **19**(8), e0308329 (2024)
9. Lawry Aguila, A., Chapman, J., Altmann, A.: Multi-modal variational autoencoders for normative modelling across multiple imaging modalities. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 425–434. Springer (2023)
10. Tolstikhin, I., Bousquet, O., Gelly, S., Schölkopf, B.: Wasserstein auto-encoders. In: International Conference on Learning Representations (ICLR) (2018)
11. Mathieu, E., Rainforth, T., Siddharth, N., Teh, Y.W.: Disentangling disentanglement in variational autoencoders. In: International Conference on Machine Learning (ICML), pp. 4402–4412. PMLR (2019)
12. Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S., Lerchner, A.: β-VAE: Learning basic visual concepts with a constrained variational framework. In: International Conference on Learning Representations (ICLR) (2017)
13. Balestriero, R., LeCun, Y.: LeJEPA: Provable and scalable self-supervised learning without heuristics. arXiv preprint arXiv:2511.08544 (2025)
14. Louizos, C., Welling, M., Kingma, D.P.: Learning sparse neural networks through L_0 regularization. In: International Conference on Learning Representations (ICLR) (2018)
15. Kendall, A., Gal, Y.: What uncertainties do we need in Bayesian deep learning for computer vision? In: Advances in Neural Information Processing Systems (NeurIPS), vol. 30 (2017)
16. Hinton, G.E.: Training products of experts by minimizing contrastive divergence. Neural Comput. **14**(8), 1771–1800 (2002)
17. Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatry **19**(6), 659–667 (2014)
18. ADHD-200 Consortium: The ADHD-200 consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. Front. Syst. Neurosci. **6**, 62 (2012)
19. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K.: The WU-Minn human connectome project: An overview. NeuroImage **80**, 62–79 (2013)
20. Bookheimer, S.Y., Salat, D.H., Terpstra, M., Ances, B.M., Barch, D.M., Buckner, R.L., Burgess, G.C., Curtiss, S.W., Diaz-Santos, M., Elam, J.S., et al.: The lifespan human connectome project in aging: An overview. NeuroImage **185**, 335–348 (2019)
21. Shafto, M.A., Tyler, L.K., Dixon, M., Taylor, J.R., Rowe, J.B., Cusack, R., Calder, A.J., Marslen-Wilson, W.D., Duncan, J., Dalgleish, T., et al.: The Cambridge centre for ageing and neuroscience (Cam-CAN) study protocol. BMC Neurol. **14**, 204 (2014)
22. Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., et al.: An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci. Data **4**, 170181 (2017)

23. Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M.E., Hopson, R., Jackson, C., Keefe, J., Riley, M., et al.: Neuroimaging of the Philadelphia neurodevelopmental cohort. NeuroImage **86**, 544–553 (2014)
24. Ellis, K.A., Bush, A.I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N.T., Lenzo, N., Martins, R.N., Maruff, P., et al.: The Australian imaging, biomarkers and lifestyle (AIBL) study of aging. Int. Psychogeriatr. **21**(4), 672–687 (2009)
25. Zuo, X.-N., Anderson, J.S., Bellec, P., Birn, R.M., Biswal, B.B., Blautzik, J., Breitner, J.C.S., Buckner, R.L., Calhoun, V.D., Castellanos, F.X., et al.: An open science resource for establishing reliability and reproducibility in functional connectomics. Sci. Data **1**, 140049 (2014)
26. LaMontagne, P.J., Benzinger, T.L.S., Morris, J.C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A.G., et al.: OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. medRxiv 2019.12.13.19014902 (2019)
27. Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., et al.: The Parkinson progression marker initiative (PPMI). Prog. Neurobiol. **95**(4), 629–635 (2011)
28. Tanaka, S.C., Yamashita, A., Yahata, N., Yoshida, W., Seymour, B., et al.: A multisite, multi-disorder resting-state magnetic resonance image database. Sci. Data **8**, 227 (2021)