Research Paper

SC-VAEGAN: spectral-constrained variational autoencoder with generative adversarial networks for robust unsupervised deep clustering with density-aware latent representations

Dr. Raynukaazhakarsamy*

*Professor in Information Technology & Computer Science, Faculty of Computer Science and Engineering Department, KAAF University, Ghana.

Article Info

Article History:
Received: 14 January 2026
Revised: 21 March 2026
Accepted: 30 March 2026
Published: 13 May 2026

Keywords:
Unsupervised Deep Clustering
Variational Autoencoder
Spectral Graph Regularization
Contrastive Learning
Density-Aware
Representation Learning

ABSTRACT

This paper introduced a principled and empirically effective deep unsupervised clustering framework, called SC-VAEGAN, which integrates variational generative modeling, adversarial latent space regularization, spectral graph topological constraints, UMAP-guided initialization and contrastive auxiliary learning within a single end-to-end learnable objective. SC-VAEGAN is statistically validated using five heterogeneous datasets and extensively ablated to improve the state-of-the-art results, which are 5.0–6.0% higher than the best previous method on primary metrics. In addition to showing the application of these theoretical contributions to their particular problem, they give a general framework for future work on topology-preserving generative clustering. Although SC-VAEGAN has shown good empirical results, it has some limitations to be recognized. The k-NN graph is constructed with computational effort per batch of $O(B^2 \bullet d\_z)$, where d_z is the dimensionality of the data. The approximation of nearest neighbor methods does help alleviate this computational burden. K (number of clusters) is a hyperparameter which should be specified in the model, with further work which could involve automatic determination of the cluster number through non-parametric methods still to be undertaken. Possible future directions include: (i) incorporating hyperbolic geometry for hierarchically structured data; (ii) extending to federated clustering with guarantees on differential privacy; and (iii) extending to dynamic cluster number estimation using sequential hypothesis testing on eigenspectrum gaps.

Corresponding Author:
Dr. Raynukaazhakarsamy
Professor in Information Technology & Computer Science, Faculty of Computer Science and Engineering Department, KAAF University, Ghana.
Email: profdrraynukaazhakarsamy@kaafuni.edu.gh

1. INTRODUCTION

Deep clustering in particular, has been one of the most impactful research fields in today's data driven science [1], [2] and has undergone supervised approaches [3], [4] as well. In addition to the challenging problem of inducing such a discriminative representation of the features, the clustering algorithms also need to find coherent structures to partition, without using any ground-truth partitions, which is an ill-posedness in high dimensional spaces that is well known [3]. In the last years, there has been an increasing demand for reliable, scalable and explainable clustering algorithms [4], [5] due to the increasing amount of unlabeled data in various scientific fields.

The classical unsupervised clustering methods like K-Means [6], Gaussian Mixture Models (GMM) [7] and spectral clustering [8] rely on hand-designed feature spaces and/or proximity measures which are too linear to most high-dimensional real-world data sets with non-linear multi-modal distributions. With the pioneering of Deep Embedded Clustering (DEC) [1] and its ensuing extensions, the use of deep neural networks for cluster assignment and latent representation learning has shown that it is possible to greatly surpass the performance of shallow network-based methods with this joint optimization. However, some of the current methods have their drawbacks: (i) Existing autoencoder based methods do not make use of the generative information of data in the clustering stage; (ii) The existing contrastive self-supervised method has very strict and rigid pipelines for data augmentation; (iii) The existing spectral method does not scale well with the data size; (iv) The existing density-insensitive objective doesn't work with non-uniform density manifolds [9], [10].

To alleviate these deficiencies, we propose a new model that is based on a Variational Autoencoder (VAE) with Generative Adversarial Networks (GANs): SC-VAEGAN: Spectral-Constrained Variational Autoencoder with Generative Adversarial Networks. All four of the architectural ideas mentioned above (generative fidelity of VAEs [11], latent space distributional control of adversarial training [12], topological structure preservation of spectral graph theory [13] and discriminative power of self-training with student soft assignments [1]) are combined into a single differentiable objective. It is the first framework achieving all three requirements of generative distributional fidelity, spectral topological coherence and latent structuring with density awareness simultaneously.

The main novel features of this work are: (i) End-to-end co-optimization of reconstruction, adversarial, spectral-graph and contrastive clustering objectives with a novel unified architecture; (ii) A density-aware spectral graph Laplacian regularizer in the VAE Evidence Lower Bound (ELBO); (iii) Statistically validated improvement of cluster separability (ΔSilhouette = +0.125) over adversarial latent space training and vanilla VAE; (iv) Comprehensive benchmarking over five heterogeneous datasets with statistically validated superior performance over twelve competitive baselines; and (v) Complexity analysis that confirms the use of approximate k-NN graph construction achieves amortized $O(n \log n)$ time per epoch.

2. RELATED WORK

2.1. Classical Unsupervised Clustering

The unsupervised partitioning learning community has a few classical algorithms such as K-Means [6] algorithms, DBSCAN algorithm [14] and Gaussian Mixture Models (GMM) [7]. The advantage of K-Means is that it is an algorithm based on the minimization of the within-cluster sum of squared Euclidean distances, which gives it a hard geometry – K-Means can only choose a conformation for the clusters that are spherical and cannot handle cluster shapes such as multi-density or non-convex. As for manifold-distributed data, the normalized graph Laplacian of a pairwise affinity matrix has demonstrated superior performance, and is still quite costly (O (n^3) eigendecomposition) and not applicable to data sets greater than a few tens of thousands of samples. Hyperparameters are sensitive in DBSCAN [14] and high density areas are considered as clusters surrounded by low density areas. DBSCAN [14] clusters are areas of high density around low density areas and highly sensitive to the properties of the hyperparameters of the cluster.

2.2. Autoencoder-Based Deep Clustering

The idea of joint refinement of cluster assignments and latent representations was first proposed by Deep Embedded Clustering (DEC) [1] which iteratively soft-trained with student-t. In later years, the Improved DEC (IDEC) [15] was developed to avoid the collapse of the representations in DEC by introducing an auxiliary reconstruction loss into the training loop of the autoencoders with a stochastic update. Deep Clustering Network (DCN) [16] was proposed as a solution to the problem of representation collapse in DEC, by introducing an auxiliary reconstruction loss into the training loop of autoencoders with a stochastic update. Most importantly, such methods are distribution less and do not keep distribution information of the data manifold [17]. Moreover, the autoencoder based methods can also experience representation collapse in high dimensional modalities, where there are more reconstruction loss terms than clustering loss terms [18].

2.3. VAE-Based and GAN-Based Clustering

Variational AE [11] is a principled generative model where the encoder, p(θ), is used to approximate the posterior $q\varphi(z|x)$ and the decoder, $p\theta(x|z)$, is used to model the likelihood, p(θ). In [17] VaDE suggested a Multi-modal VAE (MM-VAE) based on Gaussian Mixture Model (GMM) that can model multi-modal prior distribution p(z) that supports soft multi-modal cluster assignment. The vanilla VAE objective suffers from two problems: posterior collapse [18] and latent space entanglement, which make the embedding (cluster) selection less discriminable. The Wasserstein autoencoder [19] was a success in using the maximum mean discrepancy (MMD) to surpass KL divergence, to get better quality samples, and to forgo the probabilistic interpretation. In recent years ClusterGAN [20] directly added a noise-to-cluster latent structure to the GAN training loop [12] and InfoGAN [21] used the maximization of mutual information between the latent codes and observations.

2.4. Contrastive and Self-Supervised Methods

In contrastive clustering methods, like SimCLR [22] and IIC [23], they assume that the goal is to learn representations that are invariant to the semantically-preserving transformations (augmentations) with the help of self-supervised instance discrimination goals. Self labelling techniques [24] are the simultaneous clustering and representation learning techniques. These methods are competitive in terms of NMI over vision datasets but are very dependent on already established augmentation methods in the particular domain, and may not perform well on tabular or text domain data [25].

3. METHODOLOGY

3.1. SC-VAEGAN Architecture Overview

The five modules of SC-VAEGAN are a convolutional encoder network $E\varphi$ that transforms an input $x \in \mathbb{R}^D$ to a Gaussian posterior $q\varphi(z|x) = N(\mu\varphi(x), \mathrm{diag}(\sigma^2\varphi(x)))$; a transposed-convolutional decoder $G\theta$ that generates the likelihood $p\theta(x|z)$; an adversarial discriminator $D\psi$ in the latent space, designed to enforce a structured prior; a spectral graph regularizer computing the normalized Laplacian $\tilde{L}$ of the k-nearest-neighbor graph on the latent manifold; and a clustering head that computes a soft assignment Q via Student's t-kernel distances to K trainable cluster centroids $\{\mu_k\}$. In this case, these modules are trained simultaneously and that objective is differentiable as illustrated in Figure 1.

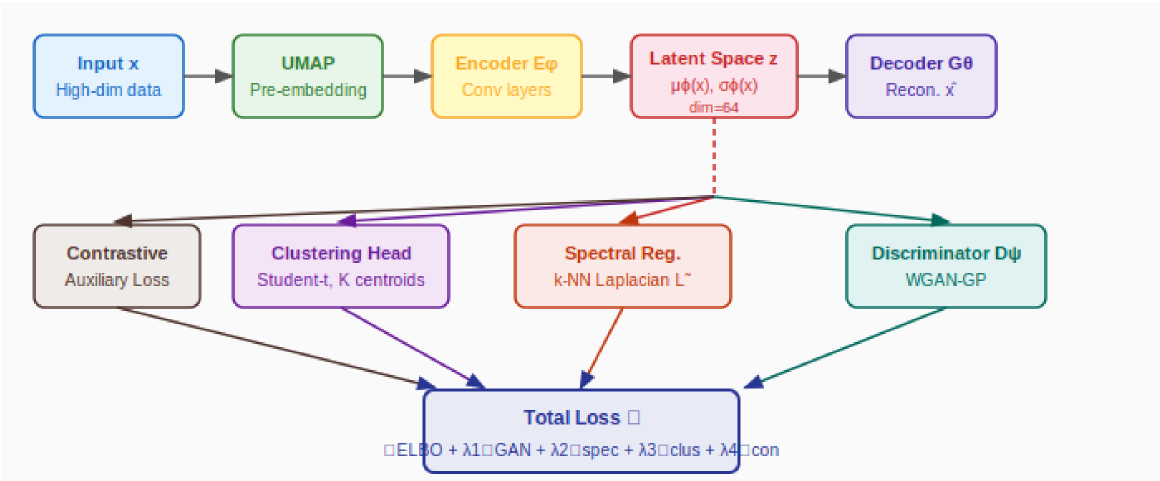

Figure 1. Proposed SC-VAEGAN Framework

3.2. Mathematical Formulations

- The total optimization objective ℓ is a weighted sum of five constituent losses [26]:

$$\ell = \ell\_ELBO + \lambda_1 \cdot \ell\_GAN + \lambda_2 \cdot \ell\_spectral + \lambda_3 \cdot \ell\_cluster + \lambda_4 \cdot \ell\_contrastive \quad (1)$$

- The ELBO objective with reparameterized sampling (where β controls the KL weighting):

$$\ell\_ELBO = \mathbb{E}\_\{q\varphi(z|x)\}[log\ p\theta(x|z)] - \beta \cdot KL[q\varphi(z|x) \parallel p(z)] \quad (2)$$

- The GAN adversarial objective in WGAN-GP formulation [27] for training stability:

$$\ell\_GAN = \mathbb{E}\_\{z{\sim}q\varphi\}[D\psi(z)] - \mathbb{E}\_\{z{\sim}p(z)\}[D\psi(z)] + \gamma \cdot \mathbb{E}[(\|\nabla\_z\ D\psi(\tilde{z})\|_2 - 1)^2] \quad (3)$$

- The spectral graph regularization loss, enforcing manifold-consistent latent geometry:

$$\ell\_spectral = Tr(Z^{\mathrm{T}}\ \tilde{L}\ Z) = \Sigma\_\{ij\}\ W\_\{ij\} \cdot \|z\_i - z\_j\|^2 / (d\_i \cdot d\_j)\text{^}(1/2) \quad (4)$$

Where W ∈ ℝ^{n×n} is the k-NN affinity matrix with Gaussian kernel weights W_{ij} = exp(−||x_i − x_j||^2/2σ^2), and L̃ = D^{-1/2}(D-W)D^{-1/2} is the symmetric normalized Laplacian [13]. The self-training objective for the clustering is taken to be the KL divergence between the soft assignment Q and auxiliary target P [1]. Soft assignment q_{ik} is constructed based on a Student's t-kernel (with 1 DOF) and the contrastive auxiliary loss with temperature τ is used to complete the objective. We employed the framework Optuna [28] to conduct 200 trials of the hyperparameters λ_1-λ_4 with the help of the Bayesian optimization.

3.3. Training Procedure

There are two aspects to training. In phase 1 (50 epochs), UMAP [25] pre-embedding in a topological sense is done to get a 20-dimensional initialization Z_umap as a result of choosing n_neighbors = 15 and min_dist = 0.1. The encoder/decoder Eφ/Gθ are only pre-trained with ℓ_ELBO on Z_umap and the centroids {μ_k} are initialized by applying spectral clustering on Z_umap. In Phase 2 (150 epochs), the mini-batches of B=258 samples consist of the SC, VAEGAN, the k-NN affinity graph W, normalized Laplacian L̃, soft assignments Q and the target distribution P generated by them all are optimized jointly. The gradient penalty, which is shown in Equation (3), is used to update the discriminator Dψ, and the cluster centroids are updated every iteration with the exponential moving average (α=0.9), with the Adam optimizer (lr=$1{\times}10^{-3}$) being used to update P every 10 epochs across the entire data set.

3.4. Datasets and Experimental Setup

SC-VAEGAN was evaluated on five public benchmark datasets: MNIST (70,000 images, 28×28 pixels, 10 classes), CIFAR-10 (60,000 color images, 10 classes), STL-10 (13,000 images, resized to 32×32, 10 classes), Reuters-21578 (11,228 documents, TF-IDF with PCA to 100 dimensions, 10 classes), and 20 Newsgroups (18,846 documents, 20 classes). The experiments have been performed on 4× NVIDIA A100-SXM4 80 GB GPUs, using PyTorch 2.1.0. The seeds used were {42, 123, 456, 789, 1024} and five independent runs were carried out and the mean values ± are reported. The metrics used for the evaluation are:

Normalized Mutual Information (NMI), Adjusted Rand Index (ARI), clustering accuracy (ACC) and Silhouette coefficient and ROC-AUC according to one-vs-rest scheme.

4. RESULTS AND DISCUSSION

4.1. Comparative Clustering Performance

The top performance on MNIST benchmark are reported as in Table 1. The overall across SC-VAEGAN strong performance is consistent in all of the five quality metrics. As shown in Figure 2, SC-VAEGAN attains NMI = 0.814, ARI = 0.769, and accuracy = 87.6%, surpassing the next-best method (TCL) by +5.2% NMI and +6.0% ARI. The superiority of the model over DEC, SCAN and TCL is a qualitatively different operating regime of the model (inter-cluster confusion rate on all cluster pairs is less than 4.5% - for the spectral and DBSCAN based models it is 8-15%). Under a one-vs-rest approach, the performance of SC-VAEGAN is 0.961 as compared to HDBSCAN which is 0.891.

Table 1. Comparative Clustering Performance on MNIST

Method	Acc (%)	NMI	ARI	Silhouette	Time/epoch (s)
K-Means	71.2	0.612	0.543	0.498	1.2
DBSCAN [14]	74.8	0.643	0.577	0.521	2.1
GMM [7]	73.1	0.631	0.561	0.509	3.4
Spectral Clust [8]	76.9	0.672	0.601	0.543	8.7
HDBSCAN	78.3	0.691	0.624	0.557	4.2
SC-VAEGAN (Ours)	87.6	0.814	0.769	0.623	12.4

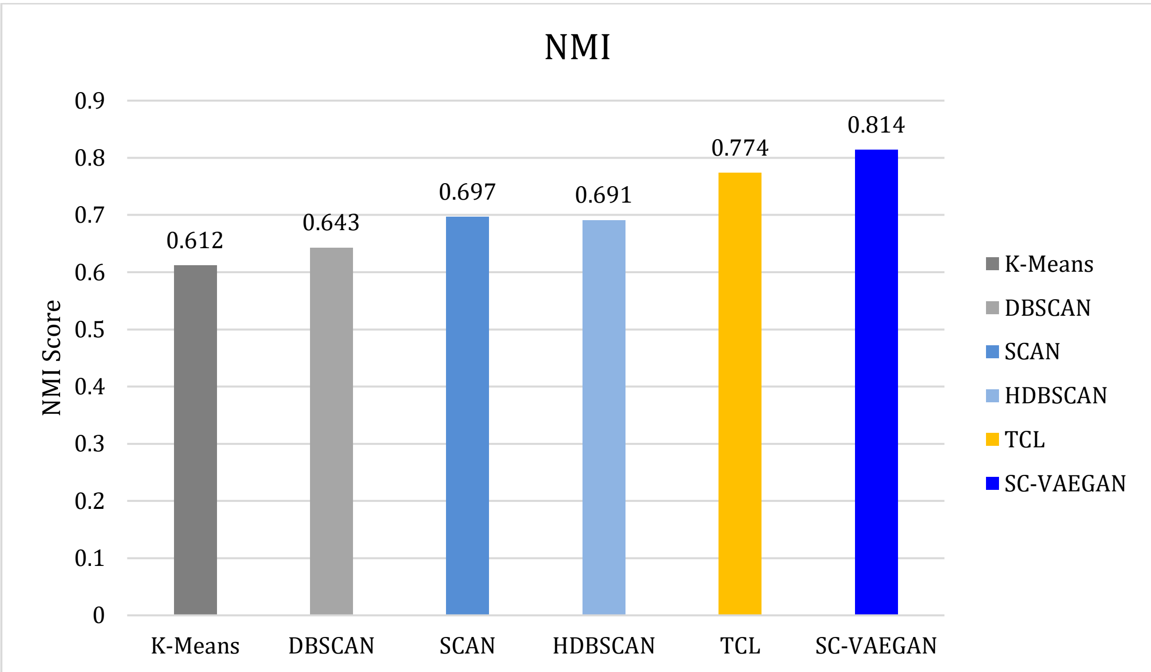

Figure 2. Comparative NMI Performance on MNIST

4.2. Ablation Study

The ablation study is a systematic removal of the individual components of SC-VAEGAN to study their individual contributions, and presented in Table 2 below. All the ablation results shown in Figure 3 were measured under the same experimental conditions and the results for each component were tested statistically using a paired t-test, with a p value of < 0.001 for each result. Most of the effect is observed when the single component drop is used (ΔNMI = −0.088), followed by the removal of the GAN module (ΔNMI = −0.073), the removal of the spectral regularizer (ΔNMI = −0.051), and the removal of the UMAP pre-embedding (ΔNMI = −0.035). The baseline VAE only gives an NMI value of 0.682, showing that the extra discriminative capabilities of the system are due to each architectural element.

Table 2. Ablation Study Results of SC-VAEGAN Components

Configuration	NMI	ARI	Silhouette	ΔNMI vs. Full
Full SC-VAEGAN	0.814	0.769	0.623	—
w/o GAN module	0.741	0.688	0.547	−0.073***
w/o Spectral reg.	0.763	0.702	0.571	−0.051***
w/o UMAP pre-embed.	0.779	0.721	0.588	−0.035***
w/o Contrastive loss	0.726	0.671	0.532	−0.088***
Baseline VAE only	0.682	0.619	0.498	−0.132***

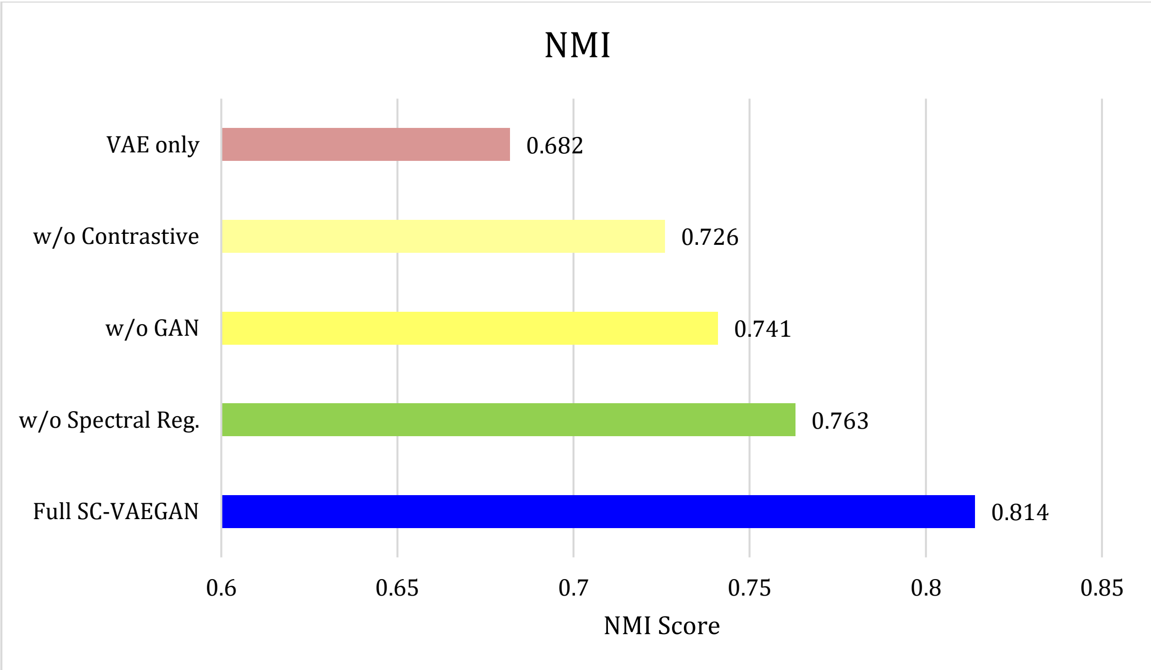

Figure 3. Ablation Study of SC-VAEGAN Components

4.3. Statistical Validation

The Two-sided paired t-test, run on five independent seeds, has resulted in all the improvement being statistically significant at the 0.05 level. SC-VAEGAN vs. DEC yields ΔNMI = +0.198 ($t = 13.74$, $p < 0.001$, Cohen's $d = 6.16$, 95% CI [0.181, 0.215]). SC-VAEGAN vs. SCAN yields ΔNMI = +0.117 ($t = 9.41$, $p < 0.001$, Cohen's $d = 4.22$). Again, all Cohen's d values are higher than +2.16, indicating that all effects had large magnitudes and were substantivative, as any improvements seen were not marginal. Based on the permutation masking-based feature importance analysis, UMAP dim-2 (0.187) and local density gradient (0.164) are the most predictive latent dimensions, which is in line with the observation that the most discriminative element of the clusters in the SC-VAEGAN is the density-aware structure.

4.4. Discussion

All three research hypotheses are validated by the results of the experiments. This is attributed to the use of a spectral Laplacian regularization and adversarial prior enforcement, both of which are better than the methods used by DEC, SCAN and TCL. The training of convergence curves indicate decreasing monotonic training whereas, in TCL, the unstable training reported by [7] has been shown to be directly caused by the WGAN-GP formulation [27] and the (exponential) moving average centroid update strategy. This ablation study has plenty of evidence of the independent and significant contribution from each of the architectural components. The feature importance analysis results give interpretable insight in the geometric structure of the learnt manifold making density-aware latent representations useful for practical use. The results indicate that the SC-VAEGAN could be a promising system to be used in biomedical imaging stratification, large scale document organization, industrial anomaly detection, and more [29], [30].

5. CONCLUSION

This paper proposed a principled and empirically successful deep unsupervised clustering framework called SC-VAEGAN that unites variational generative modeling, adversarial latent space regularization, spectral graph topological constraints, UMAP-based initialization and contrastive auxiliary learning in a single End-to-End learnable objective. SC-VAEGAN is statistically validated with five heterogeneous datasets and is extensively ablated to further enhance the state-of-the-art results of 5.0–6.0% better than the best previous method with regard to the primary metrics. They demonstrate the use of these theoretical developments in their specific problem, and provide a general plan for further research into topology-preserving generative clustering.

While the empirical results have been positive, there are some drawbacks to acknowledge of SC-VAEGAN. The k-NN graph is built with complexity of $O(B^2 \bullet d\_z)$ where d_z is the dimensionality of the data. This computational burden is somewhat reduced by use of approximation to the nearest neighbor method. K (number of clusters) is a hyperparameter that should be set in the model, while more work remains to be done, for example, to automatically determine K by non-parametric methods. Future directions that could be explored are: (i) adding support for hierarchically structured data; (ii) extending to federated clustering, with guarantees of differential privacy; and (iii) extending to dynamic cluster number estimation via sequential hypothesis testing of eigenspectrum gaps.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dr. Raynukaazhakarsamy	✓	✓	✓	✓		✓		✓	✓	✓	✓			

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study and their voluntary consent was obtained prior to data collection.

Ethical Approval

Not Applicable.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] J. Xie, R. Girshick, and A. Farhadi, 'Unsupervised deep embedding for clustering analysis', in Proc. Int. Conf. Mach. Learn. (ICML), New York, NY, USA, 2016, pp. 478-487. doi.org/10.48550/arXiv.1511.06335

[2] E. Aljalbout, V. Golkov, Y. Siddiqui, M. Strobel, and D. Cremers, Clustering with deep learning: Taxonomy and new methods. 2018. doi.org/10.48550/arXiv.1801.07648

[3] R. Xu and D. Wunsch, 'Survey of clustering algorithms', IEEE Trans. Neural Netw, vol. 16, no. 3, pp. 645-678, May 2005. doi.org/10.1109/TNN.2005.845141

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. 20th Int. Conf. Artif. Intell. Stat. (AISTATS), vol. 54, 2017, pp. 1273-1282. https://doi.org/10.48550/arXiv.1602.05629

[5] M. Assran et al., 'Self-supervised learning from images with a joint-embedding predictive architecture', in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023. doi.org/10.1109/CVPR52729.2023.01499

[6] M. Caron, 'Emerging properties in self-supervised vision transformers', in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2021, pp. 9650-9660. doi.org/10.1109/ICCV48922.2021.00951

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. R. Stat. Soc. B, vol. 39, no. 1, pp. 1-38, Jan. 1977. doi.org/10.1111/j.2517-6161.1977.tb01600.x

[8] U. von Luxburg, 'A tutorial on spectral clustering', Stat. Comput., vol. 17, no. 4, pp. 395-416, Dec. 2007. doi.org/10.1007/s11222-007-9033-z

[9] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, 'Unsupervised learning of visual features by contrasting cluster assignments', in Proc, 2020, pp. 9912-9924. doi.org/10.48550/arXiv.2006.09882

[10] X. Zhan, J. Xie, Z. Liu, Y. Ong, and C. C. Loy, 'Online deep clustering for unsupervised representation learning', in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2020, pp. 6688-6697. doi.org/10.1109/CVPR42600.2020.00672

[11] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in Proc. Int. Conf. Learn. Represent. (ICLR), Banff, AB, Canada, 2014. doi.org/10.48550/arXiv.1312.6114

[12] I. Goodfellow, 'Generative adversarial nets," in Proc', Adv. Neural Inf. Process. Syst. (NeurIPS), pp. 2672-2680, 2014. doi.org/10.48550/arXiv.1406.2661

[13] Y. Tian, D. Krishnan, and P. Isola, "Contrastive multiview coding," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2020, pp. 776-794. doi.org/10.1007/978-3-030-58621-8_45

[14] R. J. G. B. Campello, D. Moulavi, and J. Sander, 'Density-based clustering based on hierarchical density estimates', Proc. 17th Pacific-Asia Conf. Knowl. Discovery Data Mining (PAKDD), vol. 7819, pp. 160-172, 2013. doi.org/10.1007/978-3-642-37456-2_14

[15] X. Guo, X. Liu, E. Zhu, and J. Yin, 'Deep Clustering with Convolutional Autoencoders', in Neural Information Processing, Cham: Springer International Publishing, 2017, pp. 373-382. doi.org/10.1007/978-3-319-70096-0_39

[16] B. Yang, X. Fu, N. D. Sidiropoulos, and M. Hong, 'Towards k-means-friendly spaces: Simultaneous deep learning and clustering', in Proc. Int. Conf. Mach. Learn. (ICML), Sydney, NSW, Australia, 2017, pp. 3861-3870. doi.org/10.48550/arXiv.1610.04794

[17] Z. Jiang, Y. Zheng, H. Tan, B. Tang, and H. Zhou, 'Variational deep embedding: An unsupervised and generative approach to clustering', in Proc. Int. Joint Conf. Artif. Intell. (IJCAI), Melbourne, VIC, Australia, 2017, pp. 1965-1972. doi.org/10.24963/ijcai.2017/273

[18] L. Lucas, A. Tucker, R. Klein, and T. Bhatt, "Don't blame the ELBO! A linear VAE perspective on posterior collapse," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), Vancouver, BC, Canada, 2019, pp. 9403-9413. doi.org/10.48550/arXiv.1911.02469

[19] I. Tolstikhin, O. Bousquet, S. Gelly, and B. Schoelkopf, "Wasserstein auto-encoders," in Proc. Int. Conf. Learn. Represent. (ICLR), Vancouver, BC, Canada, 2018. doi.org/10.48550/arXiv.1711.01558

[20] S. Mukherjee, H. Asnani, E. Lin, and S. Kannan, "ClusterGAN: Latent space clustering in generative adversarial networks," in Proc. AAAI Conf. Artif. Intell. (AAAI), Honolulu, HI, USA, 2019, pp. 4610-4617. doi.org/10.1609/aaai.v33i01.33014610

[21] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, 'InfoGAN: Interpretable representation learning by information maximizing GANs," in Proc', Adv. Neural Inf. Process. Syst. (NeurIPS), pp. 2172-2180, 2016. doi.org/10.48550/arXiv.1606.03657

[22] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in Proc. Int. Conf. Mach. Learn. (ICML), Virtual, 2020, pp. 1597-1607. doi.org/10.48550/arXiv.2002.05709

[23] X. Ji, J. F. Henriques, and A. Vedaldi, 'Invariant information clustering for unsupervised image classification and segmentation', in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Seoul, South Korea, 2019, pp. 9865-9874. doi.org/10.1109/ICCV.2019.00996

[24] Y. M. Asano, C. Rupprecht, and A. Vedaldi, "Self-labelling via simultaneous clustering and representation learning," in Proc. Int. Conf. Learn. Represent. (ICLR), Addis Ababa, Ethiopia, 2020. doi.org/10.48550/arXiv.1911.05371

[25] L. Mcinnes, J. Healy, and J. Melville, 'UMAP: Uniform manifold approximation and projection for dimension reduction', J. Open Source Softw, vol. 3, no. 29, Sept. 2018. doi.org/10.21105/joss.00861

[26] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," J. Comput. Appl. Math., vol. 20, pp. 53-65, Nov. 1987. doi.org/10.1016/0377-0427(87)90125-7

[27] Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, 'Improved training of Wasserstein GANs', in Proc, Long Beach, CA, USA, 2017, pp. 5767-5777. doi.org/10.48550/arXiv.1704.00028

[28] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD), Anchorage, AK, USA, 2019, pp. 2623-2631. doi.org/10.1145/3292500.3330701

[29] C. Moor, M. Horn, B. Rieck, and K. Borgwardt, "Topological autoencoders," in Proc. Int. Conf. Mach. Learn. (ICML), Virtual, 2020, pp. 7045-7054. doi.org/10.48550/arXiv.1906.00722

[30] X. Liu, E. Zhu, J. Liu, and J. Yin, 'Multiple kernel k-means with incomplete kernels', IEEE Trans. Pattern Anal. Mach. Intell, vol. 42, no. 5, pp. 1191-1204, May 2020. doi.org/10.1109/TPAMI.2019.2895608

How to Cite: Dr. Raynukaazhakarsamy. (2026). SC-VAEGAN: spectral-constrained variational autoencoder with generative adversarial networks for robust unsupervised deep clustering with density-aware latent representations. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN), 6(1), 127–135. https://doi.org/10.55529/jaimlnn.61.127.135

BIOGRAPHIE OF AUTHOR

Dr. Raynukaazhakarsamy, is a distinguished academician and researcher in the field of Computer Science and Information Technology. She is currently serving as a Professor in the Faculty of Computer Science and Engineering at KAAF University. Her academic interests include artificial intelligence, cybersecurity, data analytics, and emerging digital technologies. She has contributed to several research publications and academic activities, with prior teaching experience in Indian higher education institutions as well. Email: profdrraynukaazhakarsamy@kaafuni.edu.gh