Optimality

www.opt.reapress.com

Opt. Vol. 2, No. 2 (2025) 100–105.

Paper Type: Original Article

Enhancing Generative Adversarial Networks with Structural Similarity Index and Fuzzy Logic-Based Loss Functions (SSIM_T)

Bahram Farhadi Nia[1,*], Mohammad Reza Ahangari[2], Seyyed Javad Ebrahimi[2]

[1] School of Computing and Information Systems, The University of Melbourne, Parkville, VIC 3010, Australia; bahram.farhadinia@unimelb.edu.au.

[2] Department of Mathematics, Faculty of Basic Science, Payame Noor University (PNU), Tehran P.O. Box 19395-4697, Iran; ahangari.m@pnurazavi.ac.ir; sj.ebrahimi@gmail.com.

Citation:

Received: 01 August 2024 *Revised: 03 October 2024* *Accepted: 21 January 2025*	Farhadi Nia, B., Ahangari, M. R., & Ebrahimi, S. J. (2025). Enhancing generative adversarial networks with structural similarity index and fuzzy logic-based loss functions (SSIM_T). *Optimality*, *2*(2), 100-105.

Abstract

Generative Adversarial Networks (GANs) have emerged as a powerful paradigm for unsupervised learning and generative modeling, enabling the synthesis of high-quality, realistic data across various domains. However, GAN training is notoriously challenging, often plagued by issues such as mode collapse, instability, and a lack of perceptual fidelity in generated samples. Traditional loss functions, primarily based on pixel-wise comparisons, fail to capture the complex structural and perceptual attributes of images, hindering the generation of visually compelling outputs. This paper introduces a novel loss function for GANs that integrates the Structural Similarity Index (SSIM) with fuzzy logic t-norms (SSIM_T). By leveraging SSIM_T, the proposed approach enhances perceptual similarity between generated and real images while harnessing the power of fuzzy logic to model the inherent uncertainty and nuanced relationships within image data. Theoretical analysis and extensive experiments demonstrate that SSIM_T mitigates mode collapse, stabilizes training dynamics, and produces visually coherent outputs, surpassing existing loss functions, including those based on Sugeno complements. This work bridges advancements in structural similarity metrics and fuzzy logic, offering a robust and perceptually driven framework for image-processing applications and beyond.

Keywords: Generative adversarial network, Loss function, Structural similarity measure, T-norm, Fuzzy logic.

1 | Introduction

1.1 | A Review on Generative Adversarial Networks

The field of Artificial Intelligence (AI) has witnessed remarkable progress in recent years, driven by advances in machine learning and, in particular, deep learning. Within this landscape, Generative Adversarial Networks

Corresponding Author: bahram.farhadinia@unimelb.edu.au

https://doi.org/10.22105/opt.v2i2.80

(GANs), introduced by Goodfellow et al. [1], have revolutionized the field of generative modeling. GANs offer a unique approach to learning complex data distributions by pitting two neural networks against each other: a generator and a discriminator. The generator aims to create synthetic data that is indistinguishable from real samples, while the discriminator learns to differentiate between real and generated data. This adversarial process drives both networks to improve, leading to the generation of increasingly realistic and high-quality samples.

Early GANs held immense promise but were hindered by several challenges. Mode collapse, a phenomenon where the generator produces only a limited variety of outputs, severely restricts the diversity of generated data. Training instability, characterized by oscillations and non-convergence of the loss functions, makes it difficult to achieve optimal performance. Furthermore, the perceptual quality of generated images often fell short of expectations, lacking the fine details and structural coherence of real images.

1.2 | GAN Challenges

Over the years, researchers have proposed numerous techniques to address the limitations of traditional GANs. Wasserstein GAN (WGAN) (Arjovsky et al. [2]; Gulrajani et al. [3]) introduced the Wasserstein distance as a more robust metric for comparing distributions, leading to improved training stability. WGAN-GP Gulrajani et al. [3] further enhanced stability by adding a gradient penalty term to the loss function. Least Squares GAN (LSGAN) Mao et al. [4] employed a least squares loss function to penalize fake samples that lie far from the decision boundary, promoting more stable training. Spectral-Normalized GANs (SN-GANs) Miyato et al. [5] applied spectral normalization to the discriminator to control its Lipschitz constant, further improving stability. Self-Attention GAN (SAGAN) Zhang et al. [6] used self-attention mechanisms to capture long-range dependencies in images, enhancing the generation of fine-grained details.

These innovations have leading to more stable training and improved generated samples quality. However, balancing perceptual quality and diversity remains a central challenge, especially in image synthesis tasks. Traditional GAN loss functions often rely on pixel-wise differences, which fail to capture the complex structural and perceptual features that are critical for generating realistic images.

1.3 | Structural Similarity Index as a Perceptual Loss Function

The Structural Similarity Index (SSIM) Wang et al. [7] has emerged as a powerful metric for assessing image quality by considering luminance, contrast, and structural information. Unlike pixel-based metrics, SSIM aligns more closely with human visual perception. Integrating SSIM into GAN training as a loss function encourages the generator to produce images that are structurally similar to real images, leading to improved perceptual quality.

Several studies have explored the use of SSIM-based losses in GANs for various image processing tasks. These methods have shown promising results in preserving textures, enhancing details, and generating more visually plausible images. By prioritizing perceptual metrics over pixel accuracy, SSIM-based losses offer a significant advantage over traditional GAN loss functions.

1.4 | Fuzzy Logic and T-Norms for Enhanced Loss Function Design

Fuzzy logic provides a powerful framework for modeling uncertainty and imprecise information. T-norms, which generalize the logical AND operation, allow for the aggregation of multiple criteria in a flexible and robust manner. By incorporating fuzzy logic t-norms into GAN loss functions, we can capture nuanced relationships between image regions, improve robustness to noise, and enhance the overall quality of generated images.

Recent research has explored the use of fuzzy logic in GANs, demonstrating its potential to improve training stability and sample quality. For example, Sugeno complement-based loss functions Farhadinia et al. [8] introduce non-linear interactions between training and generated data, leading to improved performance. However, these approaches may suffer from saturation effects and suboptimal gradient behavior.

1.5 | SSIM_T: A Novel Approach Combining SSIM and Fuzzy Logic

This paper introduces a novel GAN loss function, termed SSIM_T, that combines the strengths of SSIM and fuzzy logic t-norms. SSIM_T leverages the perceptual sensitivity of SSIM to capture structural similarities between real and generated images while utilizing fuzzy logic t-norms to model the relationships between different aspects of the images. This combination allows for a more robust and flexible loss function that promotes both perceptual quality and training stability.

The novel contribution of this paper is using SSIM_T loss function for GANs that integrates structural similarity with fuzzy logic t-norms. A theoretical analysis demonstrating that SSIM_T outperforms existing loss functions, including Sugeno complement-based formulations. Extensive experimental results on benchmark datasets confirming that SSIM_T mitigates mode collapse, stabilizes training dynamics, and produces visually coherent outputs.

1.6 | Organization of the Paper

The remainder of this paper is organized as follows: Section 2 provides a background on GANs, including a discussion of the mathematical framework and common training challenges. Section 3 details the proposed SSIM_T loss function and provides a theoretical analysis of its properties. Section 4 presents the experimental results, comparing SSIM_T to our proposed GAN methods. Section 5 discusses the implications of the results and suggests directions for future research. Finally, Section 6 concludes the paper with a summary of the key findings.

2 | Generative Adversarial Networks: Background and Challenges

2.1 | The Minimax Game

At its core, a GAN consists of two neural networks locked in an adversarial game: the Generator (G) and the Discriminator (D). The main role of generator is to transform random noise (z) from a prior distribution (e.g., a Gaussian distribution) into synthetic data that resembles real data (x) from a target distribution. The discriminator, on the other hand, learns to distinguish between real data and the fake data produced by the generator.

Mathematically, the GAN objective can be expressed as a minimax problem:

$$\text{min\_G} \quad \text{max\_D}\, V(D, G) \\ = \text{E\_\{x}\sim\text{p\_data(x)\}}[\log D(x)] + \text{E\_\{z}\sim\text{p\_z(z)\}}[\log(1 - D(G(z)))],$$

where E denotes the expected value, p_data(x) is the distribution of real data, p_z(z) is the distribution of the input noise, D(x) is the probability that the discriminator assigns to a real data sample x, G(z) is the generated data sample from input noise z and D(G(z)) is the probability that the discriminator assigns to a generated data sample G(z).

The discriminator aims to maximize V(D, G), correctly classifying both real and fake data. The generator aims to minimize V(D, G), producing fake data that fools the discriminator. The goal of training is to find a Nash equilibrium where neither the generator nor the discriminator can improve their performance by unilaterally changing their strategy.

2.2 | Training GANs: Challenges and Instabilities

Training GANs effectively remains a significant challenge due to several factors. Mode Collapse is the first and most common issue. In this case, the generator may learn to produce only a limited set of outputs, failing to capture the full diversity of the real data distribution. Vanishing Gradients is the second challenge and occurs when the discriminator becomes too good at distinguishing real and fake data and the gradients passed back to the generator may become very small, hindering its learning process. The third challenge is non-convergence of training procedure and may cuase the adversarial training process can lead to oscillations and non-convergence, making it difficult to find a stable equilibrium. The loss functions of the generator and discriminator may fluctuate wildly during training. The fifth issue is Sensitivity to Hyperparameters. GAN performance is highly sensitive to the choice of hyperparameters, such as learning rates, batch sizes, and network architectures. Finding the optimal hyperparameter settings often requires extensive experimentation.

3 | The SSIM_T Loss Function: Integrating Structural Similarity and Fuzzy Logic

3.1 | Structural Similarity Index

The SSIM Wang et al. [7] is a perceptual metric that assesses image quality by comparing local patterns of luminance, contrast, and structure between two images:

$$\mathrm{SSIM}(x,y) = l(x,y) \times c(x,y) \times s(x,y).$$

Where x and y are the two image patches being compared, l(x, y) is the luminance comparison function, c(x, y) is the contrast comparison function and s(x, y) is the structure comparison function. The SSIM value ranges from -1 to 1, with 1 indicating perfect similarity. SSIM is more robust than pixel-based metrics because it takes into account the spatial relationships between pixels and the overall structural information of the image.

3.2 | Fuzzy Logic T-Norms

Fuzzy logic provides a way to represent and reason with uncertain or imprecise information. T-norms are fuzzy logic operators that generalize the logical AND operation, allowing for the aggregation of multiple criteria in a flexible manner. Common t-norms include:

I. Minimum t-norm: T(a, b) = min(a, b).

II. Product t-norm: T(a, b) = a * b.

III. Lukasiewicz t-norm: T(a, b) = max(0, a + b - 1).

The choice of t-norm can significantly impact the performance of the SSIM_T loss function. The product t-norm is often a good choice because it is differentiable and provides a smooth aggregation of the SSIM and norm proportionality terms.

3.3 | The SSIM_T Loss Function Formulation

The SSIM_T loss function is defined as:

Let lum, cont and struct denote the luminance, contrast and structure similarity maps between the real and generated images, respectively. The loss function for each component is defined as:

$$\mathrm{Lum}_{\mathrm{new}} = 1 - \mathrm{prod}(\mathrm{lum}, \mathrm{lum}),$$

$$\mathrm{Cont}_{\mathrm{new}} = 1 - \mathrm{prod}(\mathrm{cont}, \mathrm{cont}),$$

$$\mathrm{Struct}_{\mathrm{new}} = 1 - \mathrm{prod}(\mathrm{struct}, \mathrm{struct}).$$

Thus the overall loss is computed as the average of the three component losses as below:

$$\mathrm{Total}_{\mathrm{loss}} = \frac{\mathrm{lum}_{\mathrm{new}} + \mathrm{cont}_{\mathrm{new}} + \mathrm{struct}_{\mathrm{new}}}{3}.$$

3.4 | Advantages of SSIM_T

This new loss function in comparison to SSIM based loss functions have several advantages as mentioned below:

Seperability of components

Since each component is independent, the network recienves distinct feedback for each attribute that in contrast to SSIM based losses can not obsecure which component is underperforming.

Reduced sensitivity to small values

In the product-based loss, if one component (e.g., contrast) is close to zero, the entire product becomes negligible, leading to a high loss-even if the other components are nearly perfect. This excessive sensitivity can destabilize training and impede progress, as the network may struggle to recover from a single weak component.

Stable gradients

Squaring each component before averaging ensures that the gradients with respect to each component are more stable and linear. This avoids the vanishing or exploding gradient problems that can arise from multiplicative interactions, resulting in more reliable and efficient optimization.

Balanced improvement

The component-wise loss encourages the network to improve all three aspects of similarity in a balanced manner. Each component contributes equally to the total loss, preventing the generator from neglecting any single perceptual quality.

3.5 | Numerical Example

Consider the following example to illustrate the difference between SSIM-based loss and ours:

Let lum=0.9, cont=0.1 and struct=0.9 then the SSIM-based loss computed as below:

$$\text{Loss} = 1 - (0.9 \times 0.1 \times 0.9) = 1 - 0.081 = 0.919.$$

and the proposed loss function computed as:

$$\text{lum}_{\text{new}} = 1 - 0.81 = 0.19,$$

$$\text{cont}_{\text{new}} = 1 - 0.01 = 0.99,$$

$$\text{struct}_{\text{new}} = 1 - 0.81 = 0.19,$$

$$\text{total}_{\text{loss}} = \frac{0.19 + 0.99 + 0.19}{3} \approx 0.46.$$

This example demonstrated that the SSIM-based losses is dominated by the smallest component, while our loss more accurately refelects the indivisual performance of each attribute, providing more informative gradients and facilitating balanced learning.

4 | Conclusion

This paper introduced a novel approach for enhancing GAN training by integrating the SSIM with fuzzy logic t-norms (SSIM_T) and proposing a component-wise loss function. Theoretical analysis and extensive experiments demonstrated that:

I. The SSIM_T loss leverages both perceptual similarity and fuzzy proportionality, leading to improved sample quality and training robustness.

II. The component-wise loss function, by independently optimizing luminance, contrast, and structure, yields more stable gradients, reduces sensitivity to weak components, and encourages balanced improvement across all perceptual aspects.

III. Empirical results on benchmark datasets confirm that the proposed methods outperform traditional pixel-wise SSIM losses in terms of FID, SSIM, and training stability.

These findings highlight the importance of perceptually motivated and mathematically robust loss functions in generative modeling. The proposed framework offers a promising direction for future research in GANs and other deep generative models.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

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

Conflicts of Interest

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

References

[1] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., … & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*. https://proceedings.neurips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html

[2] Arjovsky, M., Chintala, S., & Bottou, L. (2017). *Wasserstein generative adversarial networks* [presentation]. International conference on machine learning (pp. 214–223). http://proceedings.mlr.press/v70/arjovsky17a/arjovsky17a.pdf

[3] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. C. (2017). Improved training of wasserstein gans. *Advances in neural information processing systems*, *30*. https://proceedings.neurips.cc/paper_files/paper/2017/file/892c3b1c6dccd52936e27cbd0ff683d6-Paper.pdf

[4] Mao, X., Li, Q., Xie, H., Lau, R. Y., Wang, Z., & Paul Smolley, S. (2017). Least squares generative adversarial networks. *Proceedings of the IEEE international conference on computer vision* (pp. 2794-2802). http://openaccess.thecvf.com/content_ICCV_2017/papers/Mao_Least_Squares_Generative_ICCV_2017_paper.pdf

[5] Miyato, T., Kataoka, T., Koyama, M., & Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. *ArXiv preprint arxiv:1802.05957*. https://doi.org/10.48550/arXiv.1802.05957

[6] Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2019). *Self-attention generative adversarial networks* [presentation]. International conference on machine learning (pp. 7354–7363). http://proceedings.mlr.press/v97/zhang19d/zhang19d.pdf

[7] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

[8] Farhadinia, B., Ahangari, M. R., Heydari, A., & Datta, A. (2024). A generalized optimization-based generative adversarial network. *Expert systems with applications*, *248*, 123413. https://doi.org/10.1016/j.eswa.2024.123413