Tiako P.F. (ed) *Smart and Sustainable Applications*. Chronicle of Computing. OkIP

https://doi.org/10.55432/978-1-6692-0005-5_10

Optimal Transport-based Loss Functions for Machine Learning

Bernard Kamsu Foguem[a]*, Pierre F. Tiako[b], Cheick A. K. A. Kounta[c]

[a]University of Toulouse, France
[b]Tiako University in Oklahoma City, USA
[c]University of Science and Technology of Bamako, Maly

ABSTRACT

This short paper briefly reports the essential facets of the article (Kamsu-Foguem & al., 2022) presented and discussed as a Journal First paper. The article overviews generative neural networks whose loss functions are based on optimal transport with the Wasserstein distance. This tool of mathematical origin allows interesting automatic learning to be obtained in a reasoning time under Lipschitz constraints. As the proposed studies are based on Wasserstein Generative Adversarial Networks (WGAN), we conclude this report with a short discussion on how WGAN currently supports critical, intelligent applications in our society and nearly all industry sectors.

Keywords: Geometry, Neural Network, Deep Learning, WGAN, Intelligent Applications

I. INTRODUCTION

This Optimal transport defines a set of geometric tools with interesting properties (comparison and morphism of probability measures), making it particularly suitable for solving high-dimensional artificial learning problems.

In mathematics and, more specifically, in probability theory and statistics, the optimal transport distance (Wasserstein distance or Kantorovitch distance) is defined between probability measures in a Riemannian space.

Below, we briefly present an overview of the optimal transport theory, which is applied in machine learning, particularly for optimizing the loss function with this Wasserstein distance during learning.

II. PROPOSED INNOVATION

This section provides a brief historical overview of the evolution of Wasserstein Generative Adversarial Networks. On the one hand, there was the mathematical work of Monge (1781) on the search for transport distance at minimum cost and Kantorovitch (1942) on the duality of minimization and maximization of this cost. On the other hand, he worked in scientific computing on the competition of two artificial neural networks with the structure of Generative Adversarial Networks (GAN) (Goodfellow & al., 2014), including a generator network and another discriminator network.

III. METHODS

The combination of these mathematical and computational approaches by Arjovsky & al. (2017) gave birth to WGAN. This combination effectively approximates the minimum cost distance at the theoretical level and promotes the stability of machine learning at the practical level. Therefore, there has been substantial development in variations of WGAN, with various applications in machine learning, some for classification and others for regression (Peyre & Cuturi, 2019). The conclusion section discusses the variety of such innovative applications and their impact on our society and primary industries.

IV. LIMITATIONS

To improve the understanding of the development of WGANs and the relationships between different variants, a detailed analysis of the adversarial loss functions of the variants was carried out (Kamsu-Foguem & al., 2022). To clarify these developments, we have proposed a framework showing this temporal evolution with the intersections.

V. FUTURE WORK

Understanding the development of Wasserstein Generative Adversarial Networks and the relationship between the different variants is done through the prism of the loss function, which fully reflects it in the proliferation of structures and learning applications. The most promising perspectives concern using optimal transport to improve transparency, particularly by guaranteeing fairness, which is seen as a global sensitivity analysis (Bénesse & al. 2022).

VI. CONCLUSION

WGAN tools, as presented above in a broad view, are currently supporting critical, intelligent applications in our society and nearly all industry sectors. The impact of WGAN-related tools and techniques in machine learning improvement is enormous, based on their numerous adoptions and adaptations to support these selected few industries. In the medical field, this technique can help medical intelligent applications obtain better image quality for radiologists to make live-saving decisions (Yang & al., 2018; Lei & al., 2021; Skandarani & al., 2023; Koohi-Moghadam & al., 2023). In Cybersecurity, the technique strengthens intrusion and detection of intelligent applications, preventing malicious traffic from penetrating and disrupting critical network and system infrastructures. (Wang & Wang, 2019; Chauhan & al. 2021; Vo & al., 2024). Mechanical equipment diagnosis for fault detection and prevention can effectively diminish catastrophic failures and significant economic losses. Several works of WGAN-related tools (Liu & al., 2020; Li & al., 2022; Zhan & al., 2022) have significantly contributed to the subject. The oil, gas, and energy industries support millions of jobs all over the globe, provide lower energy costs for consumers, and ensure our energy security. The economic benefits of this industry worldwide are huge. The technics discussed in this paper have heavily contributed to the oil, gas, and energy industry, particularly during the exploration phase using seismic inversion smart applications (Wang & al.,

* Corresponding author E-mail: Bernard.Kamsu-Foguem@enit.fr

Presented at the 2023 OkIP International Conference on Automated and Intelligent Systems (CAIS) in Oklahoma City, OK, USA, and Online, on October 2-5, 2023

2022) or to help improve renewable energy usage (Ma & al., 2023).

References

Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais; Mémoires de l'Académie des Sciences de Paris, pages 666–704, 1781.

Kantorovich, L. (1942). On the translocation of masses. C.R. (Doklady) Acad. Sci. URSS (N.S.), 37:199–201, 1942.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2014). Generative Adversarial Networks. ArXiv14062661 Cs Stat.

Arjovsky M, Chintala S, Bottou L (2017) Wasserstein GAN. ArXiv170107875 Cs Stat

Peyré, G., Cuturi, M. (2019). Computational Optimal Transport: With Applications to Data Science. Foundations and Trends in Machine Learning: Vol. 11: No. 5-6, pp 355-607.

Kamsu-Foguem, B., Msouobu Gueuwou, S.L. & Kounta, C.A.K.A. (2022). Generative Adversarial Networks based on optimal transport: a survey. Artificial Intelligence Review. https://doi.org/10.1007/s10462-022-10342-x,

Bénesse, C., Gamboa, F., Loubes, JM. Boissin, T. Fairness seen as global sensitivity analysis. Machine Learning (2022). https://doi.org/10.1007/s10994-022-06202-y

Lei,K., Mardani,M., Pauly, J.M., & Vasanawala, S.S. (2021). Wasserstein GANs for MR Imaging: From Paired to Unpaired Training. IEEE Transactions on Medical Imaging, vol. 40, no. 1, pp. 105-115, doi: 10.1109/TMI.2020.3022968.

Skandarani, Y., Jodoin, P.M., & Lalande, A. (2023). GANs for Medical Image Synthesis: An Empirical Study. J Imaging. Mar 16;9(3):69. doi: 10.3390/jimaging9030069.

Koohi-Moghadam, M., & Bae, K.T. (2023) Generative AI in Medical Imaging: Applications, Challenges, and Ethics. J Med Syst. Aug 31;47(1):94. doi: 10.1007/s10916-023-01987-4.

Yang, Q., Yan, P., Zhang, Y., Yu. H., Shi, Y., Mou, X., Kalra, M.K., Zhang, Y., Sun, L., & Wang, G. (2018). Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. IEEE Trans Med Imaging.Jun;37(6):1348-1357. doi: 10.1109/TMI.2018.2827462.

Chauhan, R., Sabeel, U., Izaddoost, A., & Shah-Heydari, S.(2021) Polymorphic Adversarial Cyberattacks Using WGAN. Journal of Cybersecurity and Privacy. 1(4):767-792. https://doi.org/10.3390/jcp1040037

Wang J.-T., & Wang, C.-H. (2019). High Performance WGAN-GP based Multiple-category Network Anomaly Classification System. In The 2019 International Conference on Cyber Security for Emerging Technologies (CSET), Doha, Qatar, 2019, pp. 1-7, doi: 10.1109/CSET.2019.8904890.

Vo, H.V., Du, H.P. & Nguyen, H.N. (2024). APELID: Enhancing real-time intrusion detection with augmented WGAN and parallel ensemble learning, Computers & Security,
Volume 136, 103567, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2023.103567.

Liu, W., Duan, L., Tang, Y., & Yang, J. (2020). Data Augmentation Method for Fault Diagnosis of Mechanical Equipment Based on Improved Wasserstein GAN. 2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan), Jinan, China, pp. 103-111, doi: 10.1109/PHM-Jinan48558.2020.00025.

Li, M., Zou, D., Luo, S., Zhou, Q., Cao, L., & Huaping Liu, H. (2022). A new generative adversarial network based imbalanced fault diagnosis method. Measurement, Volume 194, 111045, ISSN 0263-2241,
https://doi.org/10.1016/j.measurement.2022.111045.

Zhan, L., Xu, X., Qiao, X., Qian, F., & Luo, Q. (2022). Fault Feature Extraction Method of a Permanent Magnet Synchronous Motor Based on VAE-WGAN. Processes. 10(2):200. https://doi.org/10.3390/pr10020200

Wang, Z., Wang, S., Zhou, C., & Cheng, W. (2022). Dual Wasserstein generative adversarial network condition: A generative adversarial network-based acoustic impedance inversion method. GEOPHYSICS 87: R401-R411.
https://doi.org/10.1190/geo2021-0600.1

Ma, X., Liu, Y., Yan, J., & Wang, H. (2023). A WGAN-GP-Based Scenarios Generation Method for Wind and Solar Power Complementary Study. Energies 16(7):3114. https://doi.org/10.3390/en16073114