International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6606
Article Received: 28 February 2023 Revised: 26 March 2023 Accepted: 16 April 2023

Generative Adversarial Networks as a Data Augmentation Tool for Handwritten Digits

Swarajya Madhuri Rayavarapu[1], Tammineni Shanmukha Prashanthi[2], Gottapu Santosh Kumar[3], Gottapu Sasibhushana Rao[4], Narendra Kumar Yegireddy[5]
[1]Department of Electronics and Communication
Andhra University, India
madhurirayavarapu.rs@andhrauniversity.edu.in
[2]Department of Electronics and Communication
Andhra University, India
prashanthitammineni.rs@andhrauniversity.edu.in
[3]Department of Civil Engineering
Gayatri Vidya Parishad College of Engineering, India
kumar.santou@gmail.com
[4]Department of Electronics and Communication, Andhra University, India
sasigps@gmail.com
[5]Department of Electrical and Electronics Engineering
Andhra University, India
narenyegireddy@gmail.com

Abstract— In the field of data processing, handwritten digit recognition (HDR) has proven to be of great use. However, due to the vast differences in how different people write, accurate recognition of such characters from images is a challenging job. The labelled samples necessary for supervised learning methods are not always easy to come by. For instance, a lot of labelled examples are needed to train a model in deep learning approaches, where all the feature extraction steps are learned within the artificial neural network. To get around this problem, data augmentation methods can be used to fill in the gaps using variations in an example's label that are already known. The Generative Adversarial Network (GAN) is able to generate random samples from the latent space that are statistically indistinguishable from the training set's actual examples. In this study, we leverage the powerful features of GAN to learn from the MNIST data set and produce digital images of handwriting.

Keywords- Artificial Intelligence, Handwritten digits, Generative Modelling, Machine learning.

I. INTRODUCTION

Automatic handwritten digit identification has become increasingly important in handwritten document analysis in recent years. It's integrated with systems that assist read old documents, mailing addresses, cheques, and it's used in a wide range of contexts. Handwritten digit identification is the ability of computers to recognize handwritten numbers. Since handwritten numbers are not exact and can be produced in various forms and sizes with overlaps, different ink thicknesses, etc., the machine has a hard time recognizing them. The handwritten digit identification system is one approach to this problem; it works by comparing an input image with a database of known digits.

Effective handwritten digit identification systems are developed using machine learning and deep learning methods. Past works have employed a wide range of deep learning methods, such as Convolutional Neural Networks (CNNs) [1], RNNs (Recurrent Neural Networks) [2] etc.

To identify images of handwritten digits, various methods have been developed, such as (Support vector machine) [3] and Random Forest classifier (RF)[4]. In addition to the aforementioned options, several algorithms detecting handwritten digits efficiently and accurately using handwriting recognition techniques.

Despite the growing popularity of deep learning approaches to hand digit identification, these kinds of tasks remain challenging and often necessitate the use of relatively large or big datasets in order to teach the algorithms the most important features that will improve their prediction accuracy. Some Digital handwritten databases include (ARDIS) [5] and Swedish Historical Birth Records (SHIBR) [6]. Due to poor performance on ARDIS and SHIBR datasets by existing state-of-the-art handwritten recognition models, more scientists are studying the problem. DIDA [7] is a brand-new collection of images of historical handwritten digits. It is a collection of scans of original Swedish historical papers.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6606
***Article Received**: 28 February 2023 **Revised**: 26 March 2023 **Accepted**: 16 April 2023*

Many tools and techniques have emerged in recent times to generate synthetic data. SDV (synthetic data vault), SynSys [8], and Synthia [9], [10] are just a few of the many new instruments being used in healthcare and medical research. Some approaches use sophisticated artificial neural networks, while others depend on tried-and-true statistical methods. The Variational Autoencoder (VAE) and the Generative Adversarial Network (GAN) are two of the most well-known and widely-used examples of deep generative models.

One popular form of generative modelling used to produce both time series data and pictures is the Generative Adversarial Network [11]. Finance [12], text to image translation [13], semantic segmentation [14], etc. are just a few of the many places you can find GANs in use.

In this study, we look into the application of generative modelling to the generation of digital photographs of handwriting. Section II discusses relevant prior work, Section III presents GAN and WGAN architecture, and Section IV explains datasets and simulation outcomes.

II. Literature Review

In the realm of machine learning, deep learning is one of the fastest growing fields, especially as it pertains to improving performance in pattern recognition and character detection. Due to the larger number of hidden units and connections, deep neural networks (DNNs) are the most time-consuming DL framework [2].

Researchers have extensively investigated several CNN algorithms due to the fact that handwritten numerals can be found in a variety of styles and orientations.

The authors in [2] compared CNN over RNN and in their paper, they found CNN better compared to RNN. The authors in [3] compared various classifiers such as K-Nearest Neighbors, RF, SVM, BP Neural Network algorithms, CNN. In these, they compared recognition rate and advantages and disadvantages of each algorithm is studied. In these, CNN got a good recognition rate.

In their study, the authors in [14] presented a convolutional neural network (CNN) model utilising the architecture of Alex Net to accurately detect characters with distorted and disordered visual features with a good accuracy.

III. Architecture

A. *Generative Adversarial Networks*

The Generative Adversarial Networks (GAN) is an unsupervised learning technique that utilises the widely-accepted zero-sum game theory for two players. It is introduced by Goodfellow in 2014. The primary aim of GANs is to estimate the distribution of authentic samples by means of a discriminator, while the Generator's function is to produce novel samples based on the genuine data samples.

Generator - In a GAN, the generator component uses input from the discriminator or takes noise as input to creates a fake data.

Discriminator - The discriminator component within a Generative Adversarial Network functions as a classifier. The objective is to differentiate authentic data from the data that is artificially generated by the generator. The selection of a network architecture for data classification should be based on its suitability to the specific data type.

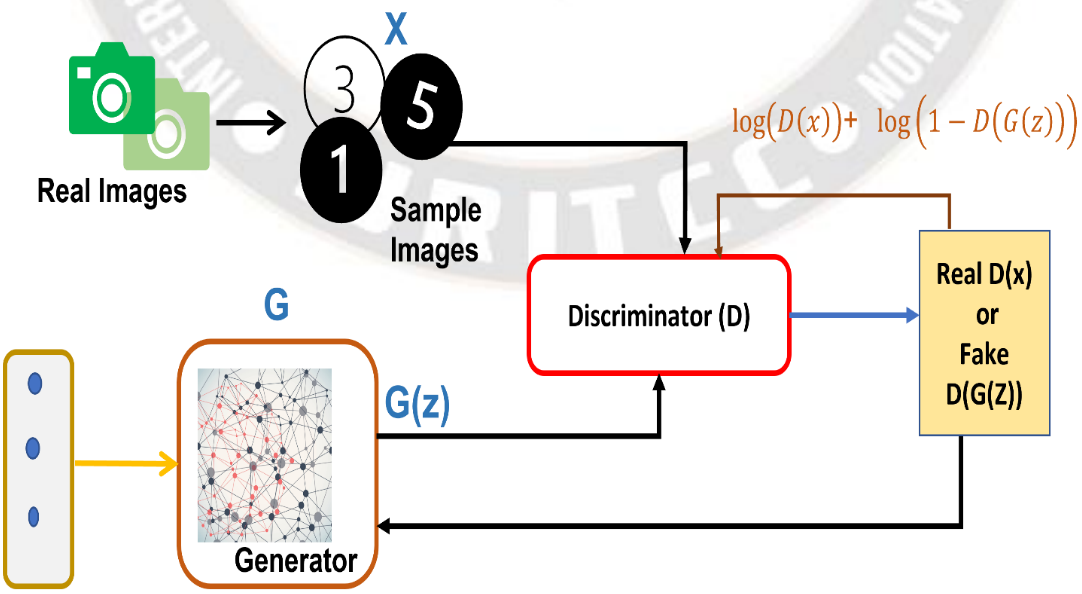

Figure 2. High level GAN Architecture

$$min_G max_D \mathrm{V}(\mathrm{D},\mathrm{G}) = E_x \log\big(D(x)\big) + E_z \log\Big(1 - D\big(G(z)\big)\Big) \quad (1)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6606
***Article Received**: 28 February 2023 **Revised**: 26 March 2023 **Accepted**: 16 April 2023*

The loss function of the GANs is given in the equation 1. Let us consider latent space or noise as Z, It is given as the input to the generator, The generator is represented by G whereas the discriminator is represented by D. The generated images are given to the discriminator in addition to the images from real data. D(x) represents real data where as D(G(z)) represents for the generated data.

The network architecture of generator and discriminator are shown in Figure 3(a) and Figure 3(b).

This training procedure is done for the allotted time period, or epochs. Nash equilibrium is the optimal stopping point for the training process. However, that is challenging to accomplish because the loss values for D and G fluctuate quite a bit. When the generator loss is nearly zero and the discriminator loss is nearly half, we cease training the GAN.

B. Wasserstein -GAN

In order to stabilise learning in generative models, Wasserstein Generative Adversarial Networks (WGANs) [17.569] were suggested. To improve upon the initial GAN's discriminator, WGANs implement a critic network. In the first generation of GANs, a discriminator would estimate probabilities to tell generated data from the actual thing. To counter this, the critic first makes an estimate of the distribution of both the actual and generated data, and then uses the Wasserstein distance between the two as a metric to find the minimum value. The produced data is of higher quality and more stable after this optimisation is applied.

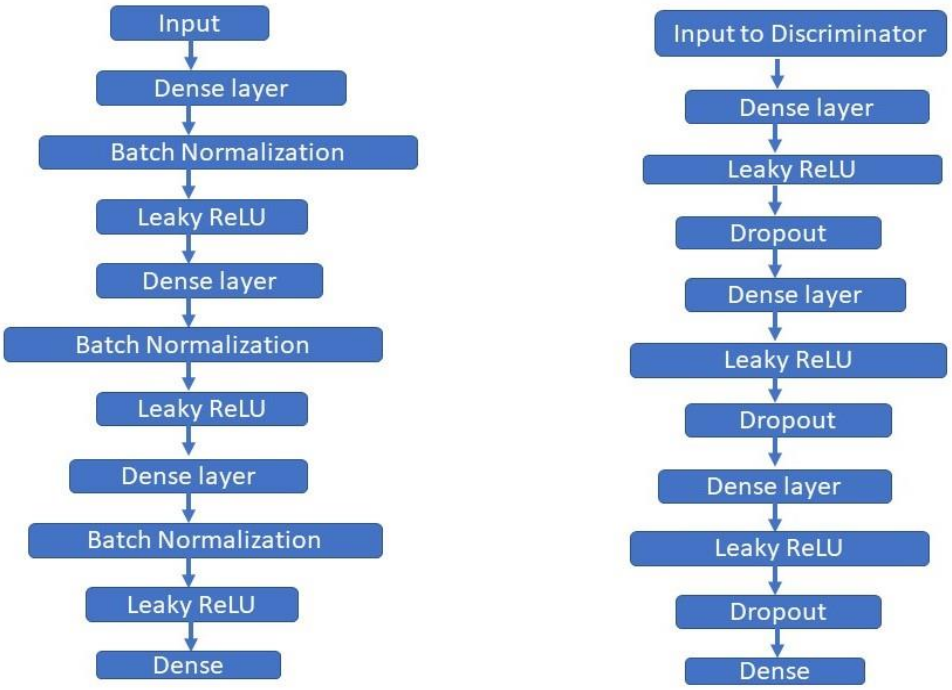

Figure 3. Generator and discriminator network architecture

IV. Experimental Results

The proposed system was built on top of the Tensor Flow library for Python 3.7. This experiment was done in Intel I7 processor laptop with an inbuilt NVIDA GEFORCE GTX GPU.

A. MNIST Dataset

The MNIST dataset, or Mixed National Institute of Standards and Technology, is a massive collection of scribbled digits compiled by NIST. There are 60,000 examples in the training set, and 10,000 in the test collection. In order to determine the form of a 28x28x1 image, the handwritten numbers dataset uses a training sample set with a buffer size of 60k and a batch size of 256. The data is stored in one place, apart from the folders of scanned handwriting documents.

B. Simulation Results

Figure 4, Figure 5 shows generated or artificial hand written images for 10^{th} 60^{th} epochs using GAN and WGAN respectively.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6606
***Article Received**: 28 February 2023 **Revised**: 26 March 2023 **Accepted**: 16 April 2023*

V. CONCLUSION

This study employs a GAN network to both generate and recognize handwritten digits. The handwritten digital data produced by GAN has potential applications in combat training, as it has the ability to enhance the precision of digital recognition on a continuous basis. Through an examination of empirical data, it has been determined that the Generative Adversarial Network (GAN) model exhibits a high level of proficiency in effectively addressing both the recognition and generation of handwritten digits, while also demonstrating exceptional resilience.

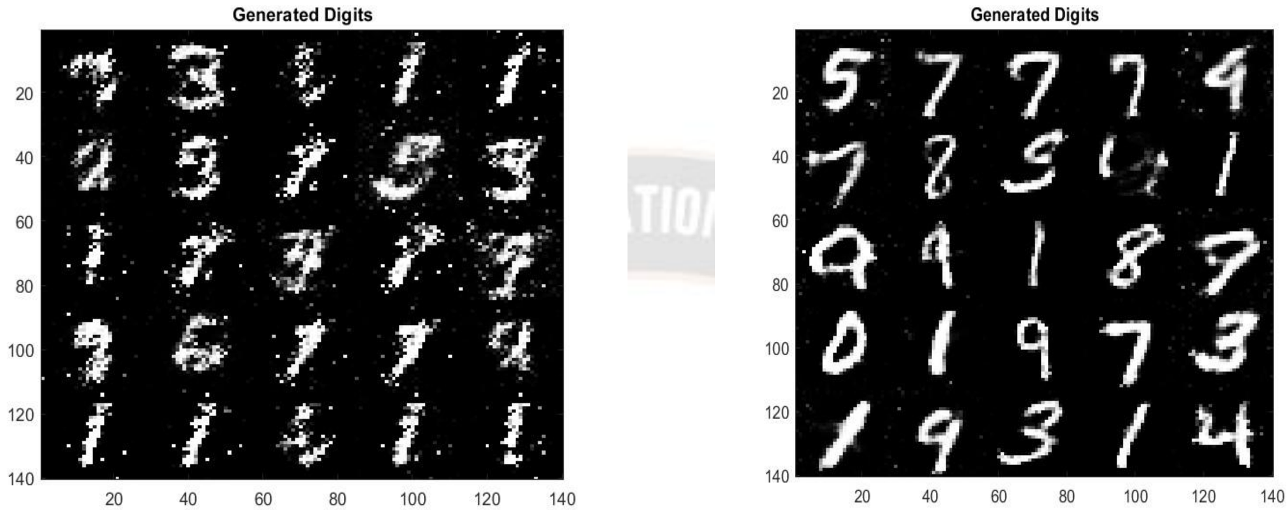

Figure 4(a) & 4(b). Generated Handwritten digit images for 10th and 60th epochs by using GAN

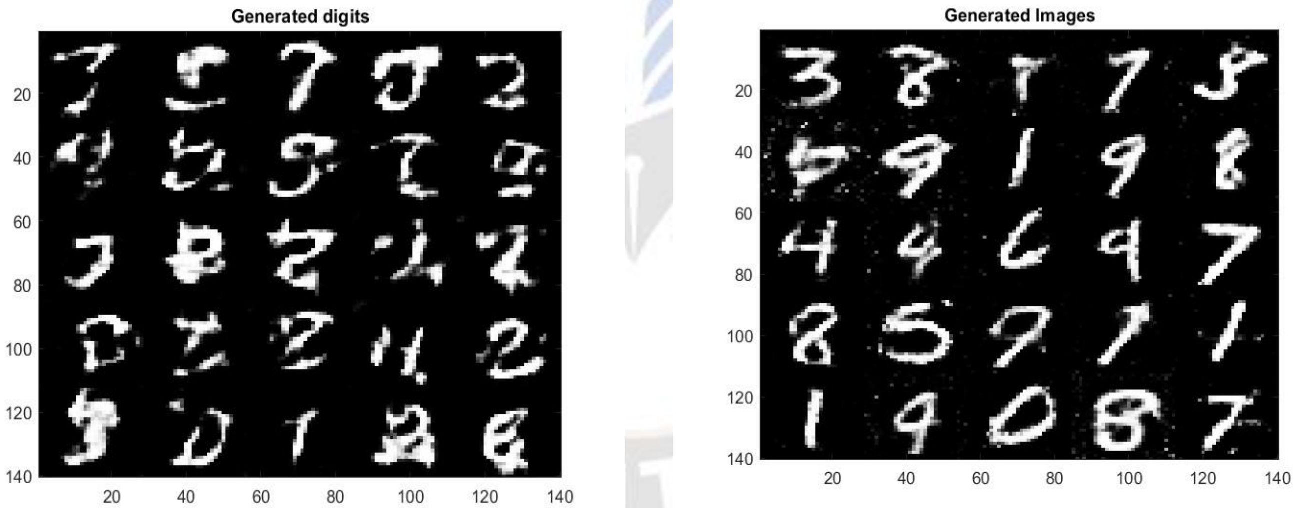

Figure 3. Generated Handwritten digit images for 10th and 60th epochs by using WGAN

REFERENCES

[1] F. Siddique, S. Sakib, and M. A. B. Siddique, ‘Recognition of handwritten digit using convolutional neural network in python with tensorflow and comparison of performance for various hidden layers’, in 2019 5th international conference on advances in electrical engineering (ICAEE), 2019, pp. 541–546.

[2] A. Hasib Uddin, J. Khatun, M. Afroz Meghna and P. Mahmud, "Bangla Handwritten Digit Recognition using RNN-CNN Hybrid Approach," 2022 25th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2022, pp. 288-293, doi: 10.1109/ICCIT57492.2022.10055089.

[3] W. Liu, J. Wei, and Q. Meng, ‘Comparisions on KNN, SVM, BP and the CNN for Handwritten Digit Recognition’, in 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA), 2020, pp. 587–590.

[4] S. Bernard, S. Adam and L. Heutte, "Using Random Forests for Handwritten Digit Recognition," Ninth International Conference on Document Analysis and Recognition (ICDAR

2007), Curitiba, Brazil, 2007, pp. 1043-1047, doi: 10.1109/ICDAR.2007.4377074.

[5] Kusetogullari, H., Yavariabdi, A., Cheddad, A. et al. ARDIS: a Swedish historical handwritten digit dataset. Neural Comput & Applic 32, 16505–16518 (2020). https://doi.org/10.1007/s00521-019-04163-3.

[6] Abbas Cheddad, Hüseyin Kusetogullari, Agrin Hilmkil, Lena Sundin, Amir Yavariabdi, Mustapha Aouache, Johan Hall; "SHIBR-The Swedish Historical Birth Records: A Semi-Annotated Dataset," Neural Computing & Applications, 33:15863–15875, Springer, 2021.

[7] https://didadataset.github.io/DIDA/

[8] J. Dahmen and D. Cook, 'SynSys: A Synthetic Data Generation System for Healthcare Applications', Sensors, vol. 19, no. 5, 2019.

[9] D. Meyer and T. Nagler, 'Synthia: multidimensional synthetic data generation in Python', Journal of Open Source Software, vol. 6, no. 65, p. 2863, 2021.

[10] D. Meyer, T. Nagler, and R. J. Hogan, 'Copula-based synthetic data augmentation for machine-learning emulators', Geoscientific Model Development, vol. 14, no. 8, pp. 5205–5215, 2021.

[11] J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[12] F. Eckerli and J. Osterrieder, 'Generative Adversarial Networks in finance: an overview', arXiv [q-fin.CP]. 2021.

[13] J. Peng et al., "Knowledge-Driven Generative Adversarial Network for Text-to-Image Synthesis," in IEEE Transactions on Multimedia, vol. 24, pp. 4356-4366, 2022, doi: 10.1109/TMM.2021.3116416.

[14] K. Liu, Z. Ye, H. Guo, D. Cao, L. Chen, and F.-Y. Wang, 'FISS GAN: A generative adversarial network for foggy image semantic segmentation', IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 8, pp. 1428–1439, 2021.

[15] S. Aly and S. Almotairi, "Deep Convolutional Self-Organizing Map Network for Robust Handwritten Digit Recognition," in IEEE Access, vol. 8, pp. 107035-107045, 2020, doi: 10.1109/ACCESS.2020.3000829.

[16] H. Mansourifar, A. Moskovitz, B. Klingensmith, D. Mintas and S. J. Simske, "GAN-Based Satellite Imaging: A Survey on Techniques and Applications," in IEEE Access, vol. 10, pp. 118123-118140, 2022, doi: 10.1109/ACCESS.2022.3221123.

[17] M. Arjovsky, S. Chintala, and L. Bottou, 'Wasserstein GAN', arXiv [stat.ML]. 2017.