International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Productivity Improvement with Generative AI Framework for Data Enrichment in Agriculture

Chhaya Narvekar[1], Dr. Madhuri Rao[2]
[1]Department Of Information Technology
Xavier Institute Of engineering
Mumbai, India
chhaya.n@xavier.ac.in
[2]Department Of AI & DS
Thadomal Shahani Engineering College
my_rao@yahoo.com
Mumbai, India

Abstract—The improvement in agricultural sector is essential for ensuring food security. Sector faces a multitude of challenges like climate change, resource limitations, and heightened food demand. To meet these challenges, there is an increasing demand for innovative solutions to enhance agricultural productivity, sustainability, and efficiency. This study presents an innovative framework that harnesses Generative Artificial Intelligence (GAI) to revolutionize agriculture. The objective is to conceptualize framework that integrates state-of-the-art AI techniques, encompassing deep learning and generative models, to provide farmers and stakeholders with data-driven insights and decision support tools. By leveraging GAI capabilities, study aims to address key agricultural issues, providing prototype implementation. Study concludes with various possible solution including crop yield prediction, disease identification, soil analysis, and resource optimization and future direction.

Keywords- Productivity, Generative Artificial Intelligence, Deep learning, simulation, visualization, synthetic data;

I. INTRODUCTION

Increasing agricultural productivity is essential for sustainable farming and global food security. It entails effectively using resources and cutting-edge technology to increase crop yields, reduce waste, and secure the productionof high-quality food [1]. Growing food demand is met by increased productivity, which also strengthens rural economies and lessens environmental harm. To do this, it includes utilizing cutting-edge technology, precision farmingtechniques, and data-driven insights to maximize resource allocation.

A. Parameters affecting production

Optimal crop choices tailored to the specific soil and climatic conditions of the region. Enhancing crop productivity through soil quality improvement and nutrient management. Pest and Disease Control: Vital measures to safeguard crops and prevent yield losses. Recognizing the impact of temperature, precipitation, and seasonal variations on agricultural outcomes. Embracing innovative machinery and technology to boost agricultural efficiency and productivity [2-3]. Ensuring fair pricing and market access to promote profitability and motivate producers. Effective allocation of financial resources and access to credit to support input investment. Education and Training Initiatives Farmer-focused programs for knowledge enhancement and the adoption of best agricultural practices. Government policies and initiatives.

Many researchers provided valuable insights into the factors influencing agricultural productivity. Ekbom (1998) has revealed that agricultural productivity is influenced by several key factors, including farm size, distance, labor availability, input costs, soil conservation quality, capital assets, and access to credit. [5] has identified variables such as the extent of cultivated land, fertilizer consumption, access to agricultural credit, and rainfall as factors positively impacting agricultural productivity has shed light on the significance of regulatory quality and efficiency in agriculture, demonstrating that reduced transaction costs and adherence to good regulatory practices are associated with higher agricultural productivity. Positive effects of policy changes, particularly the decoupling of payments, on farm productivity and sustainability outcomes. In summary, the determinants of agricultural productivity encompass a range of factors related to farm characteristics, input availability, regulatory conditions, and policy interventions [6-8].

B. Deep learning application and challenges

A number of agriculture issues are being solved by deep learning. It helps in early crop disease and pest identification, enabling prompt responses. Deep learning helps precision agriculture by maximizing resource allocation through in-the-moment data analysis, resulting in better irrigation, fertilization, and insect control. Deep learning algorithms alsoimprove food safety checks, track livestock health, predict crop yields, and

ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 September 2023

assess soil conditions. Additionally, it improves climate change adaption, optimizes the supply chain, and makes possible autonomous farming using robotics and automated vehicles. Deep learning is essentially promoting sustainability, effectiveness, and production in agriculture while providing answers to numerous business problems [9-11].

Deep learning for agricultural data has a number of limitations: The size and quality of agricultural datasets are constrained, which make it difficult to train deep learning models. Less data can result in overfitting. Data Variability: Because of factors including varying weather patterns, different types of soil, and different crop varieties, agriculturedata is highly variable. Because of this heterogeneity, deep learning models may have trouble efficiently generalizing under various settings [12-14]. Data Labelling: Labelling agricultural data for supervised learning can be a time- consuming and expensive operation. As an example, professional expertise and considerable effort are required when annotating photographs for activities like crop disease diagnosis or pest identification [15]. Imbalances in data classes,such as those between healthy and sickly crops, can bring biasinto models [16]. Deep learning algorithms might perform less well for minority classes since they favor the majority class.

C. Generative Models

These models use unsupervised learning to create novel samples from complex, high-dimensional probability distributions that are estimated. They are applicable to a variety of data formats, including audio, text, image and video [17-18].Two noteworthy projects include StyleGAN by NVIDIA, which can generate realistic human faces, and the GPT-2 language model developed by OpenAI, capable of generating original text based on a given input.

Generative models are frequently classified as: Models that explicitly train the likelihood distribution P(X|Y) through loss functions are said to be "likelihood-based," and their goalis to closely resemble the distribution of the input data. One illustration is variational autoencoders (VAE).

Implicit models: These models, like Generative Adversarial Networks (GANs), provide results that resemblethe input data without explicitly learning the likelihood distribution [19-20]. For example, GANs can produce images that are convincing enough to trick a discriminator.

D. Generative AI Prospective Solution for Agriculture

Some of the difficulties discussed in the context of agriculture could be resolved by generative AI: Data augmentation: Generative algorithms are capableof producing artificial data samples that can be used to supplement the scarce agricultural datasets [21]. Inadequate data can be addressed in this way, lowering the possibility of overfitting when deep learning models are trained. Data generation for variation: Generative models, like Generative Adversarial Networks (GANs), provide a wide range of accurate data samples [22]. This help deep learning models learn to generalize across the diversity of agriculture data, such as various crop varieties, soil types, and weather patterns. Generative AI is capable of producing labelled data in semi-supervised learning environments [23]. This strategy will make manual data labelling less time-consuming and more efficient. To solve the problem of data imbalance, generative models can be employed to create artificial minority class samples [24]. This makes it possible to increase model performance for underrepresented classes and ensuresthat deep learning methods do not favor the majority class. Data Enhancement for Transfer Learning: Using generative AI, relevant training data can be produced for particular agricultural regions or situations, facilitating transfer learning. Models that have been previously trained on a single dataset can be improved upon using the generated data. Anomaly Detection: By learning the typical distribution of agricultural data, generative AI can help in anomaly detection. It can assist in locating unexpected data patterns orvariances, which may point to problems like agricultural diseases or erratic weather.

Resource Allocation Optimization: By simulating multiple scenarios and selecting the most effective strategies based on generated data, generative models can be used to optimize resource allocation, such as the use of water and fertilizer. By simulating environmental circumstances, generative models may test and validate agricultural tactics under a variety of situations without the requirement for actual field experiments [25].

II. PROPOSED FRAMEWORK

A. Novel Generative AI Framework

Generative AI is a subset of artificial intelligence that utilizes data-driven algorithms to produce novel outputs. Introducing an innovative framework empowered by Generative AI (GAI) to revolutionize the enhancement of agriculture productivity as shown in Figure 1. This cutting-edge framework integrates advanced AI techniques, including deep learning and generative models, to provide farmers and stakeholders with data-driven insights and decision support tools. By harnessing the potential of AI, our proposal aims to address crucial agricultural challenges such as crop yield prediction, disease detection, soil analysis, and resource optimization [26-27]

1. Data Generation Layer: Functions as a multifunctional hub for data collecting. It compiles various agricultural data from a variety of sources, including photographs, satellite photos, and textual information from recognized agricultural colleges, governmental and research organizations. For the later application of deep learning algorithms, this layer serves as a vital data harmonization point. It enables the construction of complete datasets through

seamless data gathering and fusion, enabling our models to provide data- driven insights, predictive analytics, and informed decision-making within the agriculture industry. The foundation of our data-driven strategy to promote innovation and raise agricultural output is this adaptable layer.

2. Data Enrichment Layer: This is crucial in our system for augmenting the reliability and excellence of agricultural data. It addresses important issues such data imbalance, correcting differences across various data classes, and assuring fair representation. Additionally, it reduces data scarcity by using cutting-edge methods to synthesize missing data and fill in dataset gaps. Additionally, this layer combines transformer-based models to allow for text-based question-answer production, enhancing data understanding and interpretation. The Data Enrichment Layer resolves these problems and improves data completeness, which increases the overall resilience and dependability of our agricultural deep learning models and enables them to make more precise predictions and defensible conclusions in actual agricultural scenarios.

3. Model Training and Application Layer: We integrate deep learning to handle a variety of agricultural difficulties at the Model Training and Application Layer. Our models are excellent at jobs like crop classification through image analysis, correctly identifying various crop varieties. They are also adept at identifying diseases, quickly diagnosing and localizing crop problems to allow for prompt interventions. Additionally, this layer enables us to accurately anticipate agricultural yields, assisting farmers in resource allocation optimization. Additionally, NLP enable smooth interactions via a Chabot interface, providing stakeholders with an approachable manner to get information and suggestions. Our data-driven strategy has culminated in this layer, which improves agricultural productivity and decision-making while fostering accessibility and usability.

4. Simulation and Interface Layer: This layer of framework, introduced an innovative approach to agriculture by simulating diverse scenarios with varying parameters. Leveraging synthetic data, we create a dynamic environment that mirrors real-world agricultural conditions, allowing us to explore different situations and their outcomes. What sets this layer apart is its interactive feedback loop, where stakeholders actively participate in shaping and evaluating these simulated scenarios. By involving stakeholders in decision-making and learning from their input, we foster collaboration and adaptability within the agricultural ecosystem. This layer empowers stakeholders to make informed decisions.

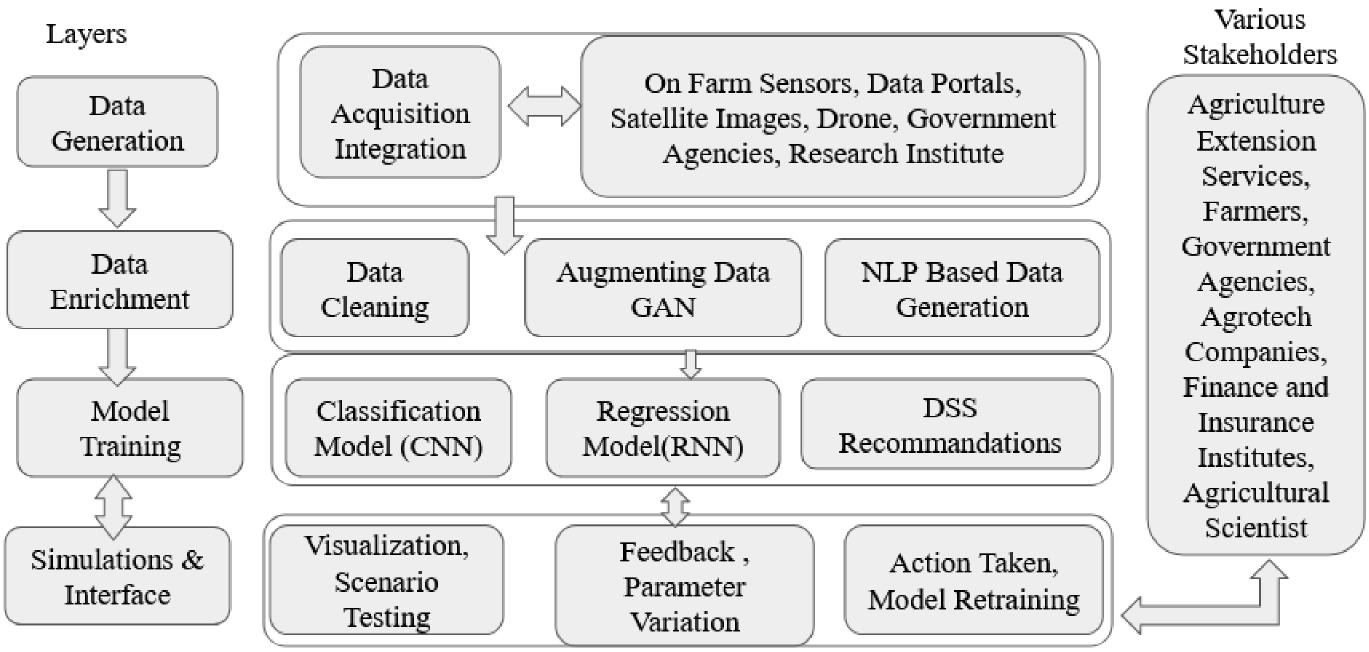

Figure 1: Proposed Generative AI based Framework for Data Enrichment

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Through accurate forecasts of crop yields, disease outbreaks, and weather patterns, farmers will benefit from enhanced decision-making, maximizing resource allocation and increasing productivity. A flexible tool will be available to agricultural researchers so they can investigate different farming methods and find cutting-edge strategies to improve crop resilience and sustainability as shown one scenario in Figure 2. This paradigm can be usedby policymakers to develop evidence-based agricultural policies that address local and global food security issues. Agribusinesses and other supply chain participants will learn crucial information about market trends and changes in demand, enabling them to streamline distribution and reduce waste.

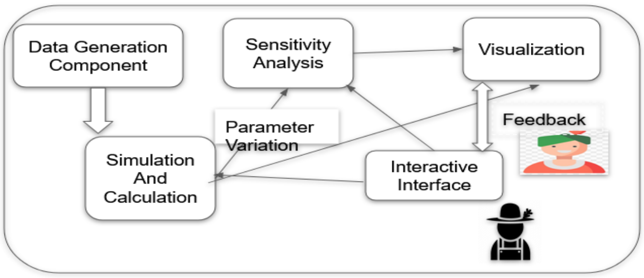

Figure 2: Sample Use case

III. PROTOTYPE USECASE WITH GAI

A. *Marathi Farmer Assistance Chatbot*

One of the challenges faced by farmers is the absence of comprehensive guidance throughout the various stages of crop management. The integration of technology into agriculture has led to innovative solutions that address challenges faced by farmers. Among these, the utilization of chatbots powered by artificial intelligence (AI) and natural language processing (NLP) has emerged as a remarkable solution. These chatbots offer real-time assistance and information to farmers, covering diverse aspects such as crop management, weather predictions, pest control, and market prices. One significant contribution is the introduction of the Marathi generative chatbot "MARGEN,"[28] designed to meet the conversational needs in agriculture.

Prototype is implemented intending to providing advisory to farmers, and enhance livelihoods by utilizing the power of BERT. Farmers may readily get essential agricultural information and professional guidance in their own language with the Marathi Farmer Assistance शेतकरीमित्र Chatbot. Despite their lack of technological expertise, farmers may easily connect with the chatbot. Objectives are Create a chatbot शेतकरीमित्र, to offer farmers individualized support. Provide farmers with a reliable source of information and assistance regarding crops, fertilizers, and other queries. Information about government schemes .

Methodology: The Krishi Diary of Maharashtra, provided the chatbot with a wealth of training data. The chatbot has been trained todeliver accurate and pertinent information on a variety of agricultural topics, such as crop cultivation, pest control, irrigation methods, soil management, weather forecasts, and more using this customized data. Employed deep learning techniques, specifically transformer-based models, to develop the chatbot, शेतकरीमित्र. Trained the chatbot model using a dataset of question-answer pairs specifically curated for farming and agriculture-related queries. Incorporated a pre-trained language model, such as

BERT, to enhance the chatbot's language understanding and response generation capabilities.

Precision, F1-score and recall is 94.14%, 95.34% and 94.04%

In figure 3 present an evaluation of our BERT model's performance, assessing its precision, accuracy, and F1 score. This analysis provides a detailed understanding of the model effectiveness. Figure 4 shows sample conversation.

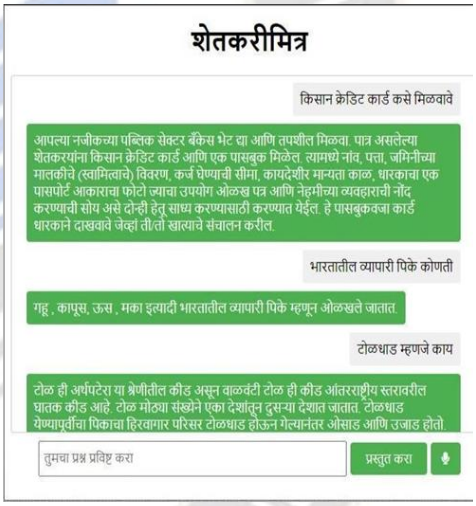

Figure 4: Sample Conversation

B. *Synthetic Image Generation*

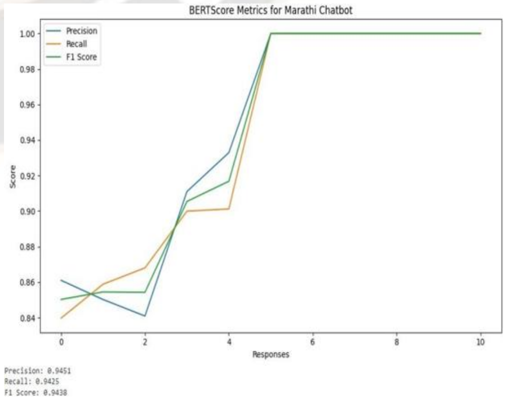

Figure 3: Bert Evaluation Metrics

The quality and diversity of the dataset are essential for creating efficient algorithms in the field of crop leaf analysis that are used for disease detection, yield prediction, and plant health assessment. A significant advancement in artificial intelligence, Generative Adversarial Networks (GANs) are a potent tool for creating synthetic visuals that closely resemble real-world data [29]. Using GANs to their full potential in this situation presents a fresh and promising solution to the lack of varied and high-quality crop leaf photos, which will eventually increase the precision and applicability of machine learning models in the field of agriculture.

The algorithm shown in Figure 5 is used for training GAN network. Sample synthetic images generated by this algorithm at various stages of training are shown in Figure 4. The synthetic images generated using above GAN algorithm are infused in the original dataset as follows, and new augmented dataset is used for further analysis. It's difficult to assess the effectiveness of generative models. There is need an empirical metric to objectively quantify the model quality and compare it to other models, even though outputs may be seen visually.

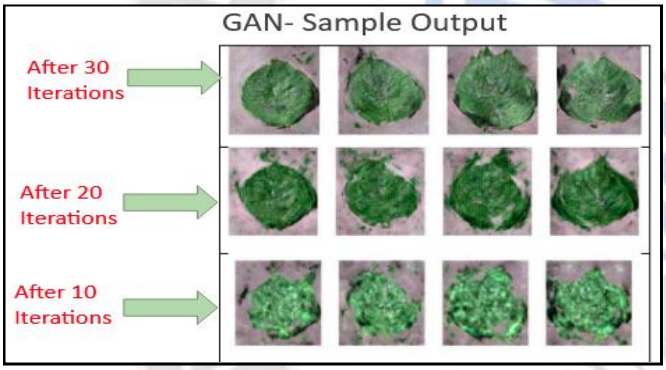

Figure 4: Images generated at various stages of GAN

Algorithm GAN Training for generating synthetic images

Input- Real images from crop disease Dataset

Output- Synthetic crop disease Images

for 100 number of training iterations do

for k steps do /# k=1;

Sample minibatch of m noise samples{z(1),...z(m)} from noise prior pg(z)

Sample minibatch of m samples {x(1),...x(m)} from data generating distribution pdata(x)

Update the discriminator by ascending its stochastic gradient:

$$\nabla_\theta \frac{1}{m}\sum_{i=1}^{m}\left[\log D_\theta\left(x^{(i)}\right)+\log\left(1-D_\theta\left(G_\phi\left(z^{(i)}\right)\right)\right)\right]$$

end for

Sample mini batch of m noise samples{z(1),...z(m)} from noise prior pg(z)

Update the Generator by ascending its stochastic gradient:

$$\nabla_\phi \frac{1}{m}\sum_{i=1}^{m}\left[\log\left(D_\theta\left(G_\phi\left(z^{(i)}\right)\right)\right)\right]$$

end for

end Algorithm

Figure 5: Algorithm for Training Generative Adversarial Network

IV. CONCLUSION

Marathi chatbot prototype shows how the creation and application of a Generative AI framework can significantly benefit the agricultural industry. This study reveals how Generative AI may significantly help farmers by offering localized, timely, and useful information in their own language. The prototype demonstrates the technology's abilityto handle the particular issues that Marathi-speaking farmersconfront, such as crop management and insect control. Second experiment image generation with GAN used for augmenting crop disease data. Research must continue to improve frameworks broaden its capabilities, and take on some significant difficulties to fully realize the potential of Generative AI in agriculture. This technology can be used to develop predictive models for crop yields, improve irrigation strategies, and identify the most appropriate crop strains for specific environmental conditions. The integration of generative AI in agriculture holds the potential to enhance farming operations in terms of efficiency, productivity, and sustainability. Generative, AI bring about significant concerns, including the potential for misleading content as well as undermining the credibility of scientific investigations. So it has to be used carefully under the guidance of experts.

REFERENCES

[1] Mechri, A., P. Lys, and F. Cachia. Productivity and efficiency measurement in agriculture: literature review and gaps analysis. technical report series GO-19-2017, 2017.

[2] Rada, Nicholas, and David Schimmelpfennig. "Propellers of Agricultural Productivity in India." Economic Research Report ERR-203, US Department of Agriculture 52 (2015).

[3] Olayiwola, O. O., P. K. Awasthi, and N. K. Raghuwanshi. "Total factor productivity of major crops in central India." Russian Journal of Agricultural and Socio-Economic Sciences 37, no. 1 (2015).

[4] Ekbom, A. (1998). Some Determinants To Agricultural Productivity - an application to the Kenyan highlands 1.

[5] Tsaurai, K. (2022). Macroeconomic Determinants of Agricultural Sector Growth in Upper Middle-Income Countries: Is Financial

Development Relevant?. Acta Universitatis Danubius. Œconomica, 18(1), 59-77.

[6] Shah, M. I., Zakari, A., Kumar, S., Abbas, S., & Sheraz, M. (2022). Quantifying the effect of waste on soil health in European Union: what are the roles of technology, natural capital, and institutional quality?. Environmental Science and Pollution Research, 29(48), 73227-73240.

[7] R (2018), Agricultural Policies in India, OECD Food and Agricultural Reviews, OECD Publishing, Paris. https://doi.org/10.1787/9789264302334-en.

[8] Lele, Uma, and SambuddhaGoswami. "The fourth industrial revolution, agricultural and rural innovation, and implications for public policy and investments: a case of India." Agricultural Economics 48, no. S1 (2017): 87-100.

[9] Ben Ayed, R., & Hanana, M. (2021). Artificial intelligence to improve the food and agriculture sector. Journal of Food Quality, 2021, 1-7.

[10] Zhang, Q., Liu, Y., Gong, C., Chen, Y., & Yu, H. (2020). Applications of deep learning for dense scenes analysis in agriculture: A review. Sensors, 20(5), 1520.

[11] Santos, L., Santos, F. N., Oliveira, P. M., & Shinde, P. (2020). Deep learning applications in agriculture: A short review. In Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics, Volume 1 (pp. 139-151). Springer International Publishing.

[12] Cravero, A., Pardo, S., Sepúlveda, S., & Muñoz, L. (2022). Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review. Agronomy, 12(3), 748.

[13] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and electronics in agriculture, 147, 70-90.

[14] Osinga, S. A., Paudel, D., Mouzakitis, S. A., & Athanasiadis, I. N. (2022). Big data in agriculture: Between opportunity and solution. Agricultural Systems, 195, 103298.

[15] Mamat, N., Othman, M. F., Abdoulghafor, R., Belhaouari, S. B., Mamat, N., & Mohd Hussein, S. F. (2022). Advanced Technology in Agriculture Industry by Implementing Image Annotation Technique and Deep Learning Approach: A Review. Agriculture, 12(7), 1033.

[16] Cheng, G., Yang, C., Yao, X., Guo, L., & Han, J. (2018). When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs. IEEE transactions on geoscience and remote sensing, 56(5), 2811-2821.

[17] Park, Sung-Wook, et al. "Review on generative adversarial networks: focusing on computer vision and its applications." Electronics 10.10 (2021): 1216.

[18] Wang, L., Chen, W., Yang, W., Bi, F., & Yu, F. R. (2020). A state-of-the-art review on image synthesis with generative adversarial networks. IEEE Access, 8, 63514-63537.

[19] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144.

[20] Goodfellow, I. (2016). Nips 2016 tutorial: Generative adversarial networks. arXiv preprint arXiv:1701.00160.

[21] Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018, April). Synthetic data augmentation using GAN for improved liver lesion classification. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 289-293). IEEE.

[22] Zhang, P., Wu, Q., & Xu, J. (2019, July). VN-GAN: identity-preserved variation normalizing GAN for gait recognition. In 2019 international joint conference on neural networks (IJCNN) (pp. 1-8). IEEE.

[23] Sixt, L., Wild, B., & Landgraf, T. (2018). Rendergan: Generating realistic labeled data. Frontiers in Robotics and AI, 5, 66.

[24] Quintana, M., Schiavon, S., Tham, K. W., & Miller, C. (2020, November). Balancing thermal comfort datasets: We GAN, but should we?. In Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (pp. 120-129).

[25] Kim, S. W., Zhou, Y., Philion, J., Torralba, A., & Fidler, S. (2020). Learning to simulate dynamic environments with gamegan. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1231-1240).

[26] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in plant science, 10, 621.

[27] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 1419.

[28] Bhalshankar, Satish V., and Ratnadeep R. Deshmukh. "MARGEN: Marathi Question Answering Generative Conversation Model." International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022). Atlantis Press, 2023.

[29] Tran, N. T., Tran, V. H., Nguyen, N. B., Nguyen, T. K., & Cheung, N. M. (2021). On data augmentation for gan training. IEEE Transactions on Image Processing, 30, 1882-1897.