International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

A Comprehensive Survey on Generative Artificial Intelligence: Applications, Challenges and Future Directions

Dr. C. Uma[1], C. Kalaivani[2]

[1,2]*Assistant Professor, Department of Computer Technology and Information Technology, Kongu Arts and Science College (Autonomous), Erode-638107*

***Abstract*— Generative Artificial Intelligence (Generative AI) has developed from statistical generative modelling into a broad family of deep learning systems capable of producing text, images, audio, video, software code, and synthetic data. This survey reviews the major stages in the development of generative AI and examines the principal model families, including Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Transformer-based models, diffusion models, and Large Language Models (LLMs). The review compares these approaches in terms of their basic mechanism, strengths, limitations, and typical application areas. Recent literature indicates that generative models have moved beyond experimental content synthesis and are increasingly used in education, healthcare, software engineering, business, creative industries, cybersecurity, and scientific research. At the same time, concerns related to hallucination, bias, privacy, copyright, security, computational cost, and responsible use continue to limit dependable deployment. This paper synthesizes representative research contributions and identifies future directions including efficient model development, multimodal generation, retrieval-supported generation, explainability, privacy preservation, reliable evaluation, and responsible AI governance.**

***Keywords*—Generative Artificial Intelligence, Generative AI, GAN, VAE, Transformer, Diffusion Models, Large Language Models, Deep Learning, Multimodal AI**

I. Introduction

Artificial Intelligence has traditionally been applied to tasks such as classification, prediction, recognition, optimization, and decision support. Generative Artificial Intelligence extends this capability by learning statistical or semantic representations from existing data and producing new samples that follow important characteristics of the learned distribution. The rapid progress of deep learning has made generation possible at a scale and quality that was difficult to achieve with earlier statistical approaches [1], [2].

A major milestone was the introduction of Generative Adversarial Networks (GANs), in which a generator and discriminator are trained in an adversarial framework [3].

Variational Autoencoders (VAEs) provided another important probabilistic approach by learning continuous latent representations that can be sampled to produce new data [2]. The Transformer architecture subsequently changed sequence modelling by relying primarily on attention mechanisms rather than recurrent processing [4]. This architecture became the basis for many modern language and multimodal models.

Large-scale language models demonstrated that increasing model size, training data, and computational capacity could produce strong few-shot capabilities across a wide range of language tasks [5]. More recently, diffusion-based approaches have become important for high-quality image and other media generation [6], [7]. These developments have contributed to the broader foundation-model paradigm, in which a single large model can support many downstream applications [8].

The growing use of Generative AI has created opportunities in education, healthcare, software development, communication, design, entertainment, and scientific research. However, capability alone does not guarantee reliability. Research has highlighted concerns surrounding inherited data biases, misinformation, social impact, privacy, environmental cost, and misuse of large-scale models [8], [9].

This survey therefore focuses on the evolution of Generative AI, its major techniques, representative applications, comparative characteristics, limitations, and future directions. The objective is to organize the existing literature into a concise framework that can assist researchers in identifying suitable model families and important open problems.

II. Evolution Of Generative Artificial Intelligence

The development of generative AI has occurred through several stages, with each stage addressing limitations of earlier approaches.

2.1 Early Probabilistic Models

Early generative systems relied heavily on probability distributions, graphical models, and statistical assumptions.

These approaches were useful for modelling sequences and structured data, but their ability to represent highly complex, high-dimensional distributions was limited.

2.2 Variational Autoencoders

Kingma and Welling introduced the variational autoencoder framework, providing an efficient approach for learning latent-variable models with neural networks [2]. VAEs established an important connection between representation learning and generative modelling.

2.3 Generative Adversarial Networks

Goodfellow et al. introduced GANs in 2014 [3]. The framework uses two competing neural networks: a generator that produces synthetic samples and a discriminator that evaluates whether samples are real or generated. GANs became particularly influential in computer vision because of their ability to generate visually realistic outputs.

2.4 Transformer Architecture

The Transformer was introduced by Vaswani et al. in 2017 [4]. Its attention-based architecture removed the need for recurrence in the core sequence-processing mechanism and enabled greater parallelization. Transformers subsequently became central to modern natural language generation and multimodal modelling.

2.5 Large Language Models

Large language models demonstrated that scaling model parameters, training data, and computational resources can substantially improve general-purpose language performance. GPT-3, for example, demonstrated strong few-shot performance without task-specific parameter updates for each task [5]. This development helped establish the current generation of foundation-model research.

2.6 Diffusion and Score-Based Models

Diffusion approaches introduced a different route to generative modelling by learning processes that transform noisy samples into structured data. Ho et al. demonstrated high-quality image synthesis using denoising diffusion probabilistic models [6]. Score-based generative modelling provided another related framework for estimating gradients of data distributions and sampling through stochastic processes [7]. Subsequent diffusion research expanded these methods to image, video, audio, scientific, and multimodal applications [10], [11].

2.7 Foundation and Multimodal Models

The foundation-model paradigm describes large models trained on broad datasets and adapted to many downstream tasks [8].

Modern Generative AI increasingly combines language, vision, audio, and video capabilities. This transition from single-modality generators toward multimodal systems represents an important current direction.

III. Major Techniques In Generative AI

3.1 Generative Adversarial Networks

GANs consist of a generator and a discriminator trained through an adversarial objective [3]. The generator attempts to produce samples that resemble real data, while the discriminator attempts to distinguish real samples from generated samples. GANs have been widely studied for image synthesis, image translation, super-resolution, and synthetic data generation. Their major strengths include realistic outputs and flexible conditioning. However, unstable optimization and mode collapse can make training difficult.

3.2 Variational Autoencoders

VAEs combine an encoder-decoder architecture with probabilistic latent-variable modelling [2]. The encoder maps input observations into a latent distribution, and the decoder reconstructs or generates observations from sampled latent variables. VAEs are useful for representation learning, reconstruction, anomaly detection, and controlled generation. Their main limitation is that generated outputs may be less sharp than outputs produced by some adversarial or diffusion models.

3.3 Transformer-Based Models

Transformers use attention mechanisms to model relationships among sequence elements [4]. They are highly scalable and have become fundamental to language generation and increasingly to multimodal generation. Applications include translation, summarization, question answering, dialogue, code generation, and content creation. Their major challenges include computational requirements, data dependence, and the need for careful evaluation of generated content.

3.4 Diffusion Models

Diffusion models learn a generative process in which noise is progressively transformed into structured data [6], [10]. They have achieved strong results in image generation and have been extended to video, audio, molecular design, and other domains [10]. Their major strengths are generation quality and diversity, while sampling cost and computational requirements remain important limitations.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

3.5 Large Language Models

LLMs are generally Transformer-based models trained on very large text corpora. Their capabilities include text completion, question answering, summarization, translation, dialogue, reasoning support, and code generation [5]. Their broad generality makes them useful as foundation models, but concerns about factual reliability, bias, data provenance, privacy, and misuse remain important [8], [9].

IV. Applications Of Generative AI

4.1 Education

Generative AI can assist in preparing learning materials, generating practice questions, summarizing content, explaining concepts, and supporting personalized learning. However, educational deployment requires attention to factual accuracy, assessment integrity, learner privacy, and appropriate human supervision [8].

4.2 Healthcare

Generative models can support synthetic data generation, medical research, image synthesis, drug discovery, and clinical documentation. In sensitive healthcare settings, privacy, validation, and human oversight are essential because incorrect generated information may have serious consequences.

4.3 Software Development

LLMs can generate code, explain programs, produce documentation, support debugging, and assist with software development activities. The benefits are substantial, but generated code must be reviewed for correctness, security, maintainability, and licensing concerns.

4.4 Content Creation

Generative systems can produce text, images, music, speech, and video. This has applications in advertising, media, entertainment, design, and digital marketing. Diffusion models and Transformer-based systems have contributed strongly to this area [5], [6], [10].

4.5 Business and Marketing

Generative AI can support report preparation, customer communication, product descriptions, advertising content, data summarization, and personalization. Foundation models provide a common platform that can be adapted to multiple organizational tasks [8].

4.6 Cybersecurity

Generative AI can help create synthetic datasets, support security analysis, and assist in threat intelligence. At the same time, generative systems may be misused for phishing, impersonation, automated misinformation, and other malicious activities. Security evaluation and responsible deployment are therefore necessary.

4.7 Scientific Research

Generative models are being explored for scientific simulation, molecular design, synthetic datasets, literature assistance, and multimodal scientific analysis. Their value depends on rigorous validation because plausible generated outputs are not necessarily scientifically correct.

V. Comparison Of Existing Generative AI Techniques

The comparison below summarizes representative model families and their commonly reported characteristics in the literature [1], [3], [4], [6], [8], [10].

TABLE I
COMPARISON OF EXISTING GENERATIVE AI TECHNIQUES

Technique	Basic Principle	Typical Data	Strengths	Limitations	Ref.
VAE	Probabilistic latent-variable learning	Images, structured data	Stable training; useful latent space	Can produce less sharp outputs	[2]
GAN	Adversarial generator-discriminator training	Images, signals	Highly realistic samples	Training instability; mode collapse	[3]
Transformer	Attention-based sequence modelling	Text, code, multimodal data	Scalable; versatile	High compute and data requirements	[4]
LLM	Large-scale Transformer language modelling	Text and code	Broad general-purpose capability	Hallucination, bias, resource cost	[5], [8], [9]
Diffusion	Iterative denoising / score-based generation	Images, video, audio, scientific data	High-quality and diverse outputs	Sampling and compute cost	[6], [7], [10]

VI. Challenges And Research Issues

Despite rapid progress, several issues remain open. First, generated content may be fluent or visually convincing while still being factually incorrect. Second, models can reproduce or amplify biases present in training data. Third, training data provenance raises privacy, copyright, and intellectual-property questions.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

Fourth, large models can require substantial computational and energy resources. Fifth, malicious actors can use generative systems to produce misleading or harmful content. Foundation-model research emphasizes that capabilities and risks can emerge from scale and that defects in a widely reused model may propagate into downstream applications [8]. Research on language-model risks has also emphasized the importance of considering dataset documentation, environmental cost, bias, and accountability [9]. These issues indicate that progress should be measured not only by generation quality but also by reliability, safety, transparency, and societal impact.

VII. Future Directions

Future research should focus on efficient and reliable generative systems. Smaller domain-specific models and efficient inference methods may reduce computational cost. Retrieval-supported generation can help connect generative models with external sources of verified information. Multimodal systems are expected to integrate language, vision, audio, and video more effectively. Research is also needed on explainability, provenance tracking, privacy-preserving training, robust evaluation, watermarking, model security, and human oversight. In addition, interdisciplinary research involving computer science, education, law, ethics, and domain specialists will be important for responsible deployment of foundation and generative models [8].

VIII. Conclusion

Generative Artificial Intelligence has progressed from probabilistic generative methods to sophisticated models capable of producing diverse forms of digital content. VAEs, GANs, Transformers, diffusion models, and LLMs represent important stages in this development and each provides different advantages for particular tasks. The literature indicates that Generative AI has significant potential in education, healthcare, software engineering, business, creative industries, cybersecurity, and scientific research.

At the same time, technical and societal challenges prevent Generative AI from being treated as an automatically reliable source of information. Hallucination, bias, privacy, copyright, security, computational cost, and misuse require continued research and careful governance. Future progress should therefore combine improved model capability with reliable evaluation, transparency, efficiency, privacy protection, and human oversight. A balanced approach that considers both technological opportunity and responsible use will be essential for the long-term development of Generative AI.

IX. Comparison Of Existing Generative Ai Techniques

The comparison shows that the techniques are complementary rather than directly interchangeable. VAEs are attractive when latent representations and relatively stable probabilistic learning are important. GANs remain influential for realistic sample generation but require careful training. Transformers provide the dominant architectural foundation for many language and multimodal systems, while diffusion methods have become particularly strong for high-quality visual generation [3], [4], [6], [10]. LLMs extend Transformer-based generation to broad general-purpose tasks, but their reliability and social risks require additional safeguards [5], [8], [9].

REFERENCES

[1] S. Bengesi, H. El-Sayed, M. K. Sarker, Y. Houkpati, J. Irungu, and T. Oladunni, "Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers," arXiv:2311.10242, 2023.

[2] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in International Conference on Learning Representations (ICLR), 2014, arXiv:1312.6114.

[3] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in Advances in Neural Information Processing Systems, vol. 27, 2014.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[5] T. B. Brown et al., "Language Models are Few-Shot Learners," in Advances in Neural Information Processing Systems, vol. 33, 2020.

[6] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," in Advances in Neural Information Processing Systems, vol. 33, 2020.

[7] Y. Song and S. Ermon, "Generative Modeling by Estimating Gradients of the Data Distribution," in Advances in Neural Information Processing Systems, vol. 32, 2019, pp. 11895–11907.

[8] R. Bommasani et al., "On the Opportunities and Risks of Foundation Models," arXiv:2108.07258, 2021.

[9] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021, pp. 610–623.

[10] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion Models: A Comprehensive Survey of Methods and Applications," arXiv:2209.00796, 2022.

[11] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, "A Survey on Generative Diffusion Model," arXiv:2209.02646, 2022.