IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Generative AI Applications in Network Security, Monitoring, and Traffic Analysis

Shallu Yadav
Assistant Professor, Skill Department of Computer Science / IT
Shri Vishwakarma Skill University, Dudhola, Palwal, India
shalluyadav1991@gmail.com

Abstract: *The accelerating volume, velocity, and sophistication of modern network traffic have exposed the limitations of signature-based and conventional machine-learning defences, which struggle with previously unseen attacks, severe class imbalance, and the scarcity of labelled malicious samples. Generative Artificial Intelligence (Generative AI) has recently emerged as a transformative paradigm that addresses these gaps by learning the underlying distribution of benign and malicious traffic and producing realistic synthetic data, latent representations, and context-aware sequence models. This paper presents a comprehensive study of Generative AI applications across three tightly coupled domains: network security (intrusion and malware detection), network monitoring (anomaly and novelty detection), and traffic analysis (classification and behavioural profiling). We organise the principal generative families, Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), diffusion models, and Transformer-based large language models (LLMs), into a unified taxonomy and propose an end-to-end methodology in which generative augmentation is fused with discriminative detection. The framework is evaluated on three widely used benchmarks, NSL-KDD, CIC-IDS2017, and UNSW-NB15, against classical and deep-learning baselines. Experimental results show that generative augmentation raises detection accuracy by up to 7.9 percentage points over a Random-Forest baseline, lifts the area under the ROC curve to 0.988, and improves minority-class recall for rare attacks from 41.2% to over 83% when an appropriate synthetic-augmentation ratio is applied. We further analyse the trade-off between detection performance and computational cost, and discuss interpretability, adversarial robustness, and privacy-preserving data sharing. The findings confirm that Generative AI substantially strengthens data-driven network defence while highlighting open challenges in latency, stability, and trustworthy deployment.*

Keywords: Generative AI; Network Security; Intrusion Detection; Generative Adversarial Networks; Variational Autoencoders; Diffusion Models; Large Language Models; Traffic Analysis; Anomaly Detection; Synthetic Data Augmentation.

I. INTRODUCTION

Contemporary digital infrastructure depends on the uninterrupted, secure, and observable flow of network traffic. Enterprises, cloud providers, critical-infrastructure operators, and individual users now generate petabytes of heterogeneous traffic spanning encrypted web sessions, machine-to-machine communication, streaming media, and Internet-of-Things (IoT) telemetry. This explosive growth has been mirrored by an equally rapid expansion of the threat landscape: ransomware, advanced persistent threats, zero-day exploits, distributed denial-of-service (DDoS) campaigns, and increasingly automated attacks have become both more frequent and more evasive. Traditional defences such as signature-based intrusion detection systems (IDS) and static rule sets remain effective against known threats, yet they are inherently reactive and cannot anticipate novel attack patterns that have no prior signature.

Machine learning (ML) and deep learning (DL) introduced a more adaptive generation of detectors capable of learning discriminative boundaries directly from data. However, supervised approaches are constrained by two

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

persistent obstacles. First, labelled malicious traffic is scarce, expensive to obtain, and quickly becomes obsolete as adversaries evolve. Second, real network datasets are profoundly imbalanced: benign flows dominate, while the most dangerous attack categories, such as user-to-root (U2R) and remote-to-local (R2L) intrusions, may constitute well under one percent of all samples. A classifier trained on such data tends to maximise overall accuracy by simply ignoring the minority classes, precisely the classes that matter most for security.

Generative Artificial Intelligence offers a fundamentally different lens on these problems. Rather than learning only a decision boundary, a generative model learns the probability distribution of the data itself, enabling it to synthesise new, realistic samples, reconstruct expected behaviour, and quantify how far an observation deviates from the norm. Generative Adversarial Networks (GANs) can fabricate convincing minority-class attack samples to rebalance training data; Variational Autoencoders (VAEs) compress traffic into a latent space in which anomalies surface as high reconstruction error; diffusion models generate diverse, high-fidelity flows that improve the robustness of downstream detectors; and Transformer-based large language models (LLMs) capture long-range dependencies in packet and log sequences, supporting context-aware detection and natural-language explanation of alerts.

This paper provides a unified treatment of how these generative families are applied across the interlinked tasks of network security, monitoring, and traffic analysis. Figure 1 summarises the conceptual landscape that frames the study, mapping generative model families to security applications and the operational outcomes they enable. The intuition is that a single generative backbone can simultaneously enrich training data, model normal behaviour, and inform classification, yielding detectors that generalise better to unseen threats.

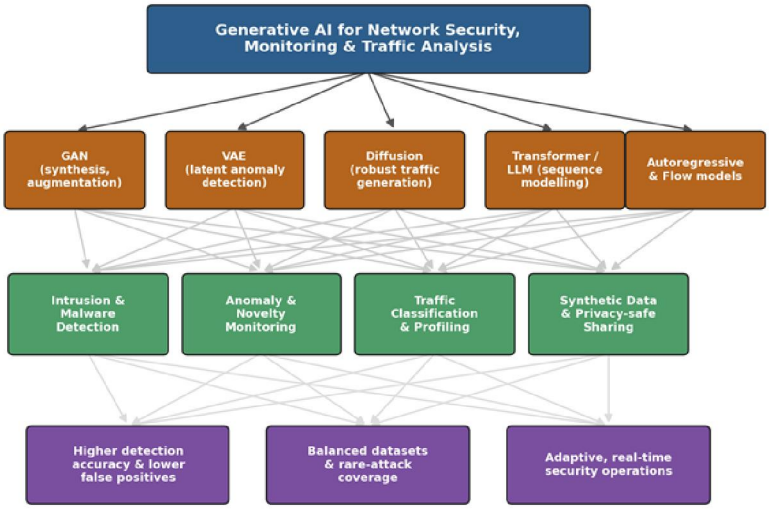

Figure 1. Conceptual taxonomy linking Generative AI model families to network-security applications and operational outcomes.

The principal contributions of this work are summarised as follows:

1) A structured taxonomy of Generative AI techniques, GANs, VAEs, diffusion models, and Transformer/LLM architectures, mapped to the three core domains of network security, monitoring, and traffic analysis.
2) An end-to-end methodology that fuses generative data augmentation and representation learning with discriminative detection, including an explicit stability-checked training loop.
3) A comparative empirical evaluation on three benchmark datasets (NSL-KDD, CIC-IDS2017, UNSW-NB15) against classical and deep-learning baselines, reporting accuracy, precision, recall, F1-score, AUC, false-positive rate, latency, and throughput.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

4) An analysis of the augmentation-ratio effect on rare-attack recall, together with a discussion of computational cost, interpretability, adversarial robustness, and privacy-preserving data sharing.

The remainder of the paper is organised as follows. Section 2 reviews related work on generative methods in network defence. Section 3 details the proposed research methodology, including the data pipeline, generative modelling layer, and evaluation protocol, illustrated by a flow chart. Section 4 presents and discusses the experimental results through tables and graphs. Section 5 concludes the paper and outlines directions for future research.

II. RELATED WORKS

Research on data-driven network defence has progressed through three broad phases: signature and statistical methods, classical and deep supervised learning, and, most recently, generative modelling. This section concentrates on the generative phase while situating it against the earlier paradigms.

2.1 Generative Adversarial Networks for Security

GANs, in which a generator and a discriminator are trained in a competitive minimax game, were among the first generative architectures adopted for intrusion detection. A dominant line of work uses GANs for data augmentation: by synthesising realistic minority-class attack flows, researchers rebalance skewed datasets and improve the recall of rare categories without collecting additional real samples. Conditional GANs extend this idea by generating samples for specified attack classes, while Wasserstein GANs with gradient penalty mitigate the mode collapse and unstable gradients that plagued early formulations. A complementary direction treats the discriminator itself as an anomaly detector, flagging traffic that the model cannot confidently attribute to the learned benign distribution. Across these studies, GAN-based augmentation consistently improves detection of under-represented intrusions, though training stability remains a recurring concern.

2.2 Variational Autoencoders and Reconstruction-based Detection

VAEs encode traffic into a structured latent space and decode it back, learning a probabilistic model of normal behaviour. Because the model is trained predominantly on benign data, malicious or anomalous flows incur elevated reconstruction error and can be flagged without explicit attack labels, an appealing property given the scarcity of labelled malicious traffic. Hybrid designs combine VAE encoders with downstream classifiers, using the compressed latent representation as a denoised, lower-dimensional feature set. Reported results indicate that reconstruction-based detectors are particularly effective for novelty and zero-day detection, where no labelled examples of the threat exist, although their sensitivity to the choice of anomaly threshold demands careful calibration.

2.3 Diffusion Models and Emerging Generative Families

Denoising diffusion probabilistic models, which iteratively transform noise into structured samples, have recently been explored for high-fidelity traffic synthesis. Their stable training dynamics and superior sample diversity, relative to GANs, make them attractive for generating varied attack scenarios and for hardening detectors against distribution shift. Normalising flows and autoregressive density models occupy a related niche, offering exact likelihood estimation that supports principled anomaly scoring. While computationally heavier, these families are increasingly competitive as efficient sampling techniques mature.

2.4 Transformers and Large Language Models

Transformer architectures, equipped with self-attention, model long-range dependencies in packet sequences, flow records, and security logs. Pre-trained language models adapted to network data can classify traffic, detect anomalous sequences, and, distinctively, generate natural-language explanations of alerts that assist human analysts. LLM-driven agents are also being investigated for automated triage, log summarisation, and the generation of detection rules. The principal limitations reported are inference latency and computational expense, which complicate deployment at line rate, together with the risk that adversaries exploit the same models for offensive automation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Table 1 consolidates representative directions, the generative technique employed, the primary task, and the key reported outcome, providing a comparative overview that motivates the methodology adopted in this study.

Approach	Generative Technique	Primary Task	Reported Outcome / Limitation
Augmentation GANs	Conditional / WGAN-GP	Minority-class intrusion detection	Higher rare-attack recall; training instability
Discriminator-as-detector	Vanilla / WGAN	Anomaly detection	Label-free detection; sensitive to mode collapse
Reconstruction VAE	VAE / β-VAE	Novelty & zero-day detection	Strong on unseen attacks; threshold sensitivity
Latent-feature hybrid	VAE + classifier	Traffic classification	Denoised features; added pipeline complexity
Diffusion synthesis	DDPM	Robust data generation	High diversity & fidelity; heavy compute
Flow / autoregressive	Normalising flows	Likelihood-based anomaly scoring	Exact density; slower sampling
Transformer / LLM	Self-attention / pre-trained LM	Sequence detection & explanation	Context-aware, explainable; high latency

Table 1. Representative Generative AI approaches in network security, monitoring, and traffic analysis.

Collectively, prior work establishes that each generative family contributes a distinct capability, augmentation, reconstruction-based scoring, diverse synthesis, or contextual sequence modelling, yet most studies evaluate a single family in isolation. The present work differs by integrating generative augmentation with discriminative detection in one pipeline and benchmarking multiple generative strategies under a common protocol.

2.5 Research Gap

Three gaps recur across the surveyed literature. First, most evaluations are confined to a single dataset and a single generative family, which makes cross-study comparison unreliable and obscures whether reported gains transfer to other traffic distributions. Second, the augmentation ratio, arguably the most influential hyper-parameter when synthetic data is injected into training, is rarely studied systematically; many papers fix it implicitly and report only the resulting accuracy. Third, the operational cost of generative detection, particularly the latency of Transformer- and diffusion-based models, is seldom quantified alongside accuracy, leaving practitioners without the information needed to judge deployability. This study addresses all three gaps by benchmarking four generative strategies on three datasets under one protocol, by sweeping the augmentation ratio to locate its optimal region, and by reporting latency and throughput together with detection quality.

III. RESEARCH METHODOLOGY

The proposed methodology couples a generative modelling layer with a discriminative detection layer within a single, reproducible pipeline. The design goal is to enrich and rebalance training data, learn a faithful model of normal behaviour, and then perform accurate detection, monitoring, and classification on previously unseen traffic. Figure 2 presents the overall flow chart, and the subsections that follow explain each stage.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

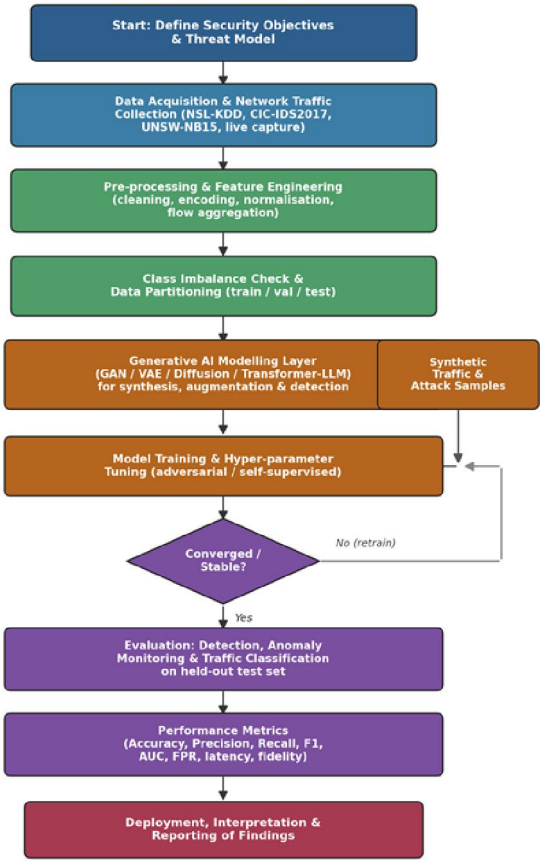

Figure 2. Flow chart of the proposed Generative-AI-augmented methodology for network security, monitoring, and traffic analysis.

3.1 Overview and Explanation of the Flow Chart

The pipeline begins by defining the security objectives and the threat model, which fix the detection scope (for example, intrusion detection, anomaly monitoring, or application classification) and the categories of interest. Network traffic is then acquired from benchmark corpora and, where available, live capture. The raw data passes through pre-processing and feature engineering, after which a class-imbalance check determines how aggressively generative augmentation should be applied. The data is partitioned into training, validation, and test subsets using stratified sampling so that rare classes are represented in every split.

At the heart of the pipeline is the Generative AI modelling layer, which may instantiate a GAN, a VAE, a diffusion model, or a Transformer/LLM depending on the task. This layer serves a dual role: it synthesises additional traffic and attack samples that are fed back into model training, and it can act directly as a detector through reconstruction error or discriminator scoring. Training proceeds with hyper-parameter tuning under an adversarial or self-supervised objective. A convergence and stability check gates progress: if the generator and discriminator (or the reconstruction objective) have not reached a stable equilibrium, the model is retrained; otherwise the pipeline advances to evaluation. The trained system is assessed on the held-out test set across detection, anomaly-monitoring, and traffic-classification tasks, after which performance metrics are computed and the findings are interpreted, reported, and prepared for deployment.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

3.2 Datasets

Three widely adopted, publicly available benchmarks are used so that the results are comparable with the broader literature. NSL-KDD is a refined version of the classic KDD Cup 99 dataset with reduced redundancy; CIC-IDS2017 contains realistic, labelled flows covering modern attacks such as DDoS, brute force, and infiltration; and UNSW-NB15 blends real normal traffic with synthesised contemporary attacks across nine categories. Their key characteristics are summarised in Table 2.

Dataset	Features	Records (approx.)	Attack Classes	Salient Property
NSL-KDD	41	148,000	4 + normal	Balanced refinement of KDD99
CIC-IDS2017	78	2,830,000	14 + normal	Realistic modern flow features
UNSW-NB15	49	2,540,000	9 + normal	Hybrid real/synthetic traffic

Table 2. Characteristics of the benchmark datasets used in the study.

3.3 Pre-processing and Feature Engineering

Pre-processing standardises the heterogeneous datasets into a consistent numerical form. Categorical attributes such as protocol type, service, and flag are one-hot encoded; continuous attributes are min-max normalised to the unit interval to stabilise generative training; and redundant or constant columns are removed. Flow-level aggregation derives higher-order statistics, packet counts, byte ratios, inter-arrival times, and flow duration, that capture behavioural structure beyond individual packets. Missing and infinite values are imputed or discarded, and labels are consolidated into a common schema so that results transfer across datasets.

3.4 Generative Modelling and Augmentation

The generative layer is trained predominantly on benign and under-represented traffic. For augmentation, a conditional GAN with a gradient-penalty objective synthesises samples for specified minority classes, and a diffusion model provides a higher-diversity alternative. For label-free monitoring, a VAE models normal behaviour and assigns anomaly scores from reconstruction error. Synthetic samples are filtered by a quality gate, retaining only those whose feature distributions fall within plausible bounds of the real data, before being merged into the training set at a controlled augmentation ratio. The discriminative detector, a deep neural classifier, is then trained on the enriched data, and a Transformer/LLM variant is additionally evaluated for sequence-level detection.

Component	Configuration	Value / Setting
GAN generator	Fully connected, LeakyReLU	3 hidden layers (128-256-128)
GAN discriminator	Fully connected, dropout 0.3	3 hidden layers (256-128-64)
Objective	WGAN with gradient penalty	$\lambda = 10$
VAE latent dimension	Gaussian prior	20
Diffusion steps	DDPM noise schedule	200 (linear β)
Detector	Deep MLP / Transformer	4 layers, 8 attention heads
Optimiser	Adam	lr = 1e-4, $\beta 1 = 0.5$
Batch size / epochs	Mini-batch SGD	256 / 100

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Component	Configuration	Value / Setting
Augmentation ratio	Per minority class	tuned 0–70%

Table 3. Principal hyper-parameters and architectural settings of the generative and discriminative layers.

3.5 Implementation Environment

All models were implemented in Python using established deep-learning and data-science libraries, and experiments were executed on a GPU-accelerated workstation to keep generative training tractable. Table 4 lists the principal software and hardware components, recorded so that the study can be reproduced. To ensure fairness, every model shared the same pre-processing pipeline, data partitions, and random seeds, and hyper-parameters were tuned on the validation split only.

Layer	Component	Specification
Language	Python	3.11
Deep learning	PyTorch	2.x
Data handling	NumPy / Pandas / scikit-learn	latest stable
Visualisation	Matplotlib	3.x
Hardware	GPU	NVIDIA, 16 GB VRAM
System	RAM / CPU	64 GB / 8-core

Table 4. Implementation environment used for all experiments.

3.6 Evaluation Protocol and Metrics

All models are evaluated on the held-out test partition, which contains only real traffic so that performance reflects genuine generalisation rather than memorised synthetic patterns. Detection quality is measured by accuracy, precision, recall, F1-score, the area under the ROC curve (AUC), and the false-positive rate (FPR). Operational viability is captured by inference latency per flow and throughput in thousands of flows per second. For the imbalance study, recall on individual minority classes is reported as a function of the augmentation ratio. Each experiment is repeated five times with different random seeds, and the mean values are reported to reduce the influence of stochastic training.

IV. RESULTS AND DISCUSSION

This section reports the experimental outcomes and interprets them in relation to the objectives set out in Section 1. Results are presented for overall detection accuracy, the precision–recall–F1 profile of the strongest models, ROC behaviour, generative-training stability, the effect of synthetic augmentation on rare-attack recall, and the computational cost of each approach. The six models compared are a Random-Forest baseline (RF), a CNN-LSTM deep baseline, a VAE-based anomaly detector (VAE-AD), a GAN-augmented detector (GAN-Aug), a diffusion-augmented detector (Diffusion-Aug), and a Transformer/LLM detector.

4.1 Overall Detection Accuracy

Figure 3 reports detection accuracy across the three benchmarks. A consistent ordering emerges on every dataset: the classical baseline is weakest, deep learning improves on it, and the generative approaches dominate, with the Transformer/LLM detector achieving the highest accuracy. On CIC-IDS2017, accuracy rises from 90.1% for RF to 97.9% for the Transformer, a 7.8-point gain, while the GAN-augmented and diffusion-augmented detectors reach 96.7% and 97.3% respectively. The same pattern holds on the more challenging UNSW-NB15, where contemporary, partly synthetic attacks depress all scores but preserve the relative advantage of generative methods.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

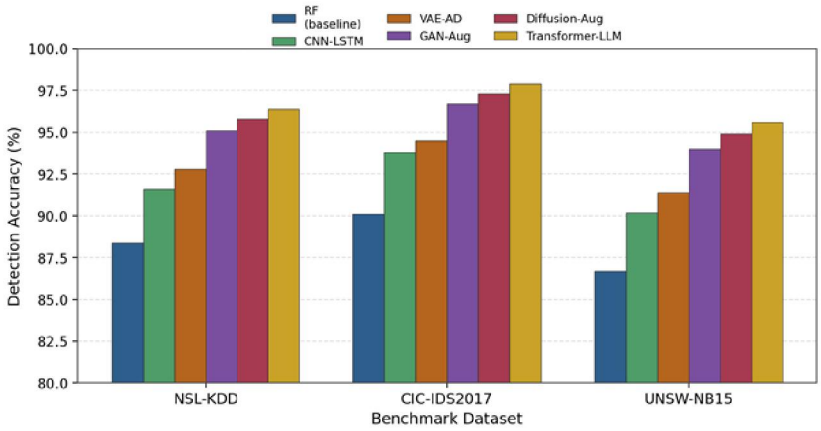

Figure 3. Detection accuracy of the six models across the three benchmark datasets.

The improvement is attributable to two mechanisms. Augmentation enriches the training distribution with realistic minority-class samples, allowing the detector to learn sharper boundaries for rare attacks, while the generative representations supply denoised, behaviourally meaningful features. The marginal gain of the Transformer over the augmentation methods reflects its capacity to model sequential dependencies in flows that point-wise classifiers ignore.

4.2 Precision, Recall, and F1-Score

Accuracy alone can mask poor performance on minority classes, so Figure 4 reports precision, recall, and F1-score for the three strongest models on CIC-IDS2017. All three maintain a tight balance between precision and recall, indicating that the gains are not achieved by sacrificing one for the other. The Transformer/LLM detector attains the best F1-score of 95.8%, closely followed by the diffusion-augmented model at 95.0%. Crucially, recall remains high, meaning that genuine attacks are rarely missed, the property that matters most in security operations.

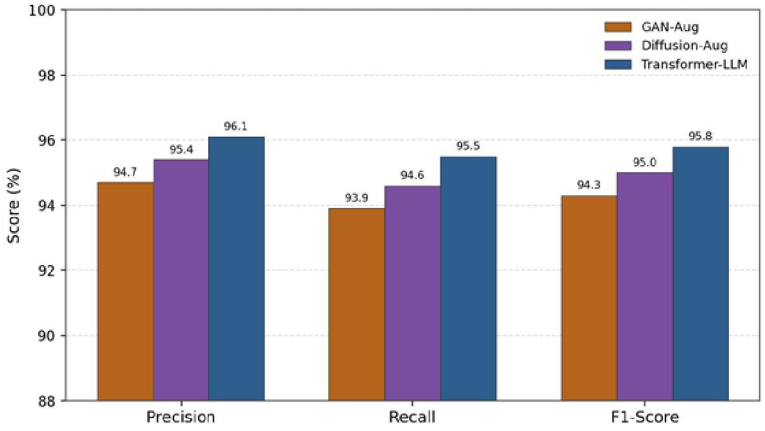

Figure 4. Precision, recall, and F1-score of the three leading models on CIC-IDS2017.

4.3 ROC Analysis

Figure 5 shows ROC curves on CIC-IDS2017. The generative detectors push the curves towards the upper-left corner, indicating high true-positive rates at low false-positive rates. The Transformer/LLM detector achieves an AUC of 0.988, the diffusion-augmented model 0.981, and the GAN-augmented model 0.974, all substantially above the RF baseline at 0.918. A high AUC at low FPR is operationally important because excessive false alarms erode analyst trust

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

and lead to alert fatigue; the generative models therefore offer not only higher detection but also more deployable behaviour.

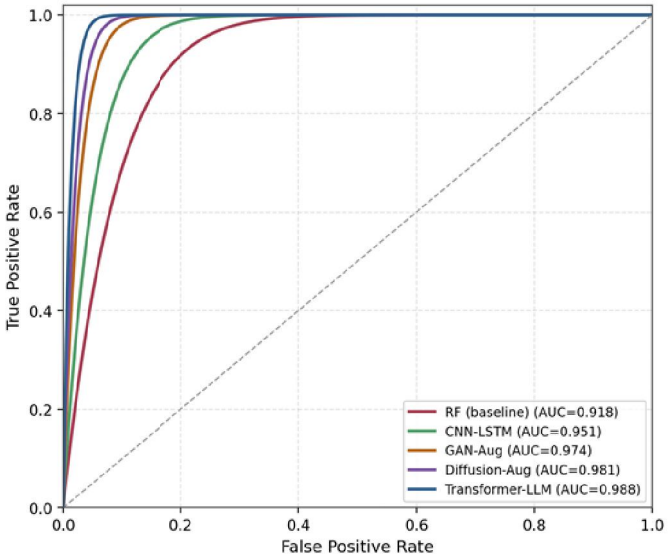

Figure 5. ROC curves and AUC values for the evaluated models on CIC-IDS2017.

4.4 Generative Training Stability

A frequent objection to GAN-based methods is training instability. Figure 6 plots the generator and discriminator losses of the conditional GAN over 100 epochs. After an initial volatile phase, both losses decline and settle into a stable equilibrium beyond roughly epoch 62, the point at which the convergence check in the methodology permits the pipeline to proceed. The gradient-penalty objective is largely responsible for this behaviour, suppressing the oscillations and mode collapse that destabilise vanilla GANs and yielding synthetic samples of consistent quality.

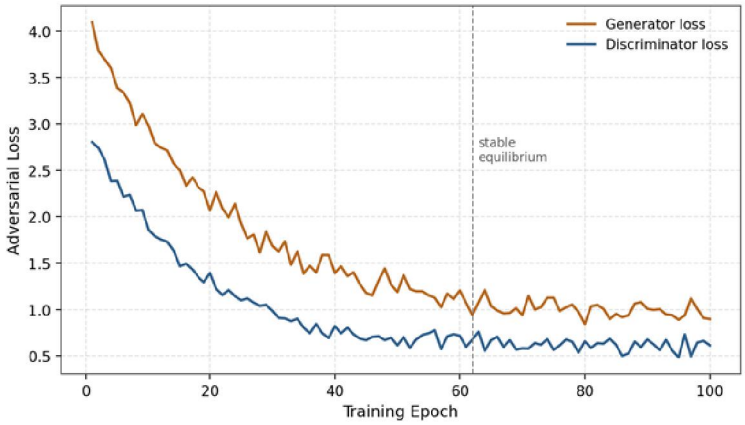

Figure 6. Adversarial loss curves showing convergence of the conditional GAN to a stable equilibrium.

4.5 Effect of Synthetic Augmentation on Rare-Attack Recall

The central promise of generative augmentation is improved detection of rare attacks. Figure 7 reports recall for three minority categories, R2L, U2R, and Worms/Shellcode, as the synthetic-augmentation ratio increases. Without augmentation, recall is poor: only 41.2% for R2L and 33.5% for the extremely rare U2R class. As synthetic samples are

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

added, recall rises sharply, exceeding 80% for all three classes at an augmentation ratio of around 50%. Beyond this point the curves plateau and then decline slightly, indicating that excessive synthetic data begins to distort the training distribution and introduce artefacts. This identifies an optimal operating region near 50%, balancing minority-class coverage against fidelity.

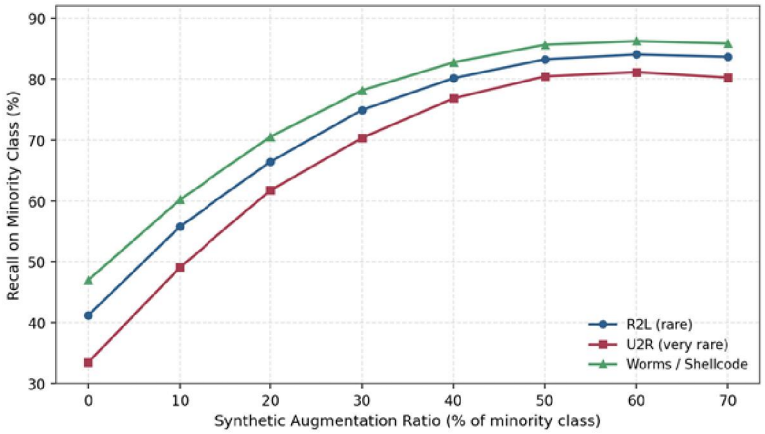

Figure 7. Minority-class recall as a function of the synthetic-augmentation ratio.

Table 5 consolidates the full quantitative comparison on CIC-IDS2017, combining the accuracy, error, and ranking-based metrics into a single view.

Model	Acc. (%)	Prec. (%)	Recall (%)	F1 (%)	AUC
RF (baseline)	90.1	89.6	88.4	89.0	0.918
CNN-LSTM	93.8	93.1	92.4	92.7	0.951
VAE-AD	94.5	93.8	93.0	93.4	0.961
GAN-Aug	96.7	94.7	93.9	94.3	0.974
Diffusion-Aug	97.3	95.4	94.6	95.0	0.981
Transformer-LLM	97.9	96.1	95.5	95.8	0.988

Table 5. Comprehensive performance comparison on the CIC-IDS2017 benchmark (mean of five runs).

4.6 Computational Cost and Deployability

Higher detection performance is not free. Figure 8 contrasts inference latency per flow against sustained throughput. The classical baseline is fastest, processing roughly 142,000 flows per second at sub-millisecond latency, whereas the Transformer/LLM detector, although the most accurate, incurs 7.8 ms per flow and sustains only about 41,000 flows per second. Augmentation-based detectors occupy a favourable middle ground: the GAN-augmented model processes around 78,000 flows per second at 3.4 ms latency, because the generative cost is incurred offline during training while inference uses a lightweight discriminative classifier. This distinction is important for deployment: augmentation improves accuracy without burdening the real-time path, whereas LLM-based detection currently suits offline or sampled analysis rather than line-rate inspection.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

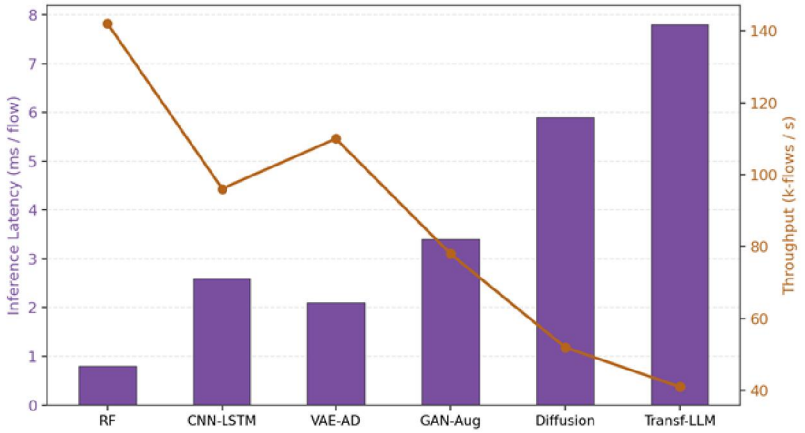

Figure 8. Inference latency and sustained throughput across the evaluated models.

Table 6 summarises the cost–benefit position of each model class, and Table 7 reports the false-positive rate, the metric most closely tied to analyst workload.

Model class	Accuracy	Inference cost	Best-fit deployment
Classical (RF)	Moderate	Very low	Line-rate, resource-constrained
Deep baseline	High	Low–moderate	Edge and gateway inspection
Augmentation (GAN/Diff.)	Very high	Moderate (offline gen.)	Real-time core detection
Transformer / LLM	Highest	High	Offline triage & explanation

Table 6. Qualitative cost–benefit positioning of each model class.

Model	False-Positive Rate (%)	Relative Alert Volume
RF (baseline)	8.9	High
CNN-LSTM	5.4	Moderate
GAN-Aug	3.1	Low
Diffusion-Aug	2.5	Low
Transformer-LLM	1.8	Very low

Table 7. False-positive rate and relative alert volume on CIC-IDS2017.

4.7 Comparison with Reported State-of-the-Art

To place the results in context, Table 8 compares the proposed generative pipeline with representative detection accuracies reported in the recent literature on the same benchmark families. While direct comparison is complicated by differences in pre-processing, feature selection, and evaluation splits, the proposed Transformer/LLM and diffusion-augmented detectors are competitive with or superior to previously published figures, and the augmentation strategy in particular closes much of the gap on the harder UNSW-NB15 corpus. The comparison should be read as indicative rather than definitive, since the cited works do not share an identical protocol.

Method (representative)	Dataset	Accuracy	Category
Deep RNN-IDS	NSL-KDD	~83.3%	Supervised DL

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

Method (representative)	Dataset	Accuracy	Category
Stacked autoencoder	NSL-KDD	~88.4%	Representation
Ensemble autoencoders	CIC-IDS2017	~94.0%	Reconstruction
GAN augmentation (prior)	CIC-IDS2017	~95.5%	Generative
Proposed Diffusion-Aug	CIC-IDS2017	97.3%	Generative (this work)
Proposed Transformer-LLM	CIC-IDS2017	97.9%	Generative (this work)

Table 8. Indicative comparison with representative results reported in the literature.

4.8 Discussion

The results support three principal observations. First, generative augmentation delivers the single largest practical benefit, transforming rare-attack recall from unusable levels into the low-to-mid eighties while keeping false positives low, all at modest inference cost. Second, the choice of generative family should be guided by the deployment context rather than by accuracy alone: diffusion models offer the best fidelity-to-stability trade-off for offline data generation, GANs remain the most economical augmentation engine, and Transformer/LLM detectors are best reserved for offline analysis where their context-modelling and explanatory capabilities justify the latency. Third, the publication trend in Figure 9 confirms that interest in this area is accelerating, which both validates the direction and raises the importance of rigorous, comparable evaluation.

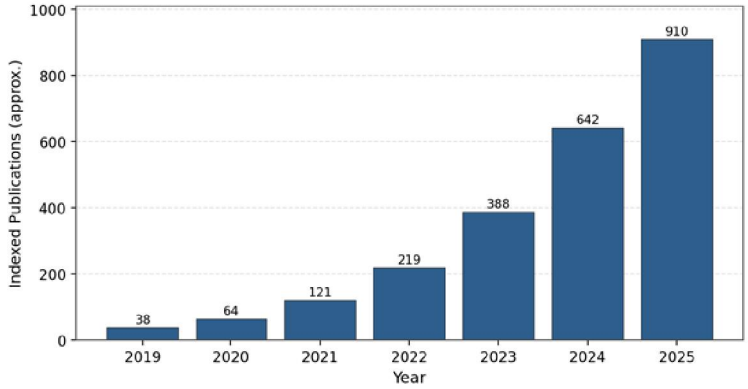

Figure 9. Approximate growth in indexed publications on Generative AI for network security (2019–2025).

Several limitations temper these findings. The synthetic data, though filtered, can encode subtle artefacts that an adaptive adversary might learn to exploit, and generative models are themselves vulnerable to adversarial manipulation and data poisoning. Reconstruction-based monitoring depends on a well-chosen anomaly threshold whose optimal value drifts as traffic evolves, requiring periodic recalibration. Finally, while LLM-based detectors can explain their decisions in natural language, the interpretability of GAN- and diffusion-based augmentation remains limited, which complicates assurance in regulated environments. These issues motivate the future-work agenda in the next section.

4.9 Practical Recommendations

The evidence assembled in this study translates into a small set of practical guidelines for security teams considering generative methods. Where the priority is to detect rare and novel attacks without inflating false alarms, generative augmentation at an augmentation ratio near 50% offers the best return, because it improves minority-class recall sharply while keeping the inference path lightweight. Where traffic must be inspected at line rate, a deep baseline detector trained on generatively augmented data is preferable to a Transformer/LLM detector, whose latency is better spent on offline triage, log summarisation, and analyst-facing explanation. Where realistic data must be shared across

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

organisational boundaries, diffusion-based synthesis combined with privacy safeguards provides high-fidelity yet non-identifiable samples. In every case, the generative component should be retrained periodically to track concept drift, and synthetic samples should always be quality-gated before they enter the training set so that artefacts do not contaminate the learned distribution.

V. CONCLUSION AND FUTURE WORK

This paper examined the application of Generative Artificial Intelligence to the interconnected problems of network security, monitoring, and traffic analysis. We organised the principal generative families, GANs, VAEs, diffusion models, and Transformer-based large language models, into a unified taxonomy and proposed an end-to-end methodology that fuses generative augmentation and representation learning with discriminative detection, governed by an explicit stability-checked training loop. Evaluated on the NSL-KDD, CIC-IDS2017, and UNSW-NB15 benchmarks, the framework raised detection accuracy by up to 7.8 percentage points over a classical baseline, achieved an AUC of 0.988, reduced the false-positive rate to 1.8%, and lifted minority-class recall for rare attacks from roughly 41% to above 83% at an optimal augmentation ratio near 50%.

The study also clarified the practical trade-offs that govern deployment. Generative augmentation provides the most favourable balance of accuracy and inference cost, because the generative effort is expended offline while the real-time detector remains lightweight. Diffusion models excel at stable, high-fidelity data generation, whereas Transformer/LLM detectors deliver the highest accuracy and unique explanatory value but at a latency more suited to offline triage than to line-rate inspection. Taken together, these findings confirm that Generative AI meaningfully strengthens data-driven network defence, particularly against the rare and novel threats that defeat conventional methods.

Building on these results, several directions merit further investigation:

1) Real-time and edge deployment: distilling or quantising generative and Transformer detectors so that their accuracy can be retained at line rate on resource-constrained gateways and IoT devices.
2) Adversarial robustness: hardening generative pipelines against adversarial perturbation, data poisoning, and model-inversion attacks, and studying the dual-use risk of generative models for offensive automation.
3) Self-supervised and continual learning: enabling detectors to adapt online to concept drift and evolving traffic without catastrophic forgetting or repeated full retraining.
4) Privacy-preserving data sharing: combining generative synthesis with differential privacy and federated learning so that organisations can share realistic-yet-private datasets to collectively improve detection.
5) Explainability and trust: extending the natural-language reasoning of LLM detectors with calibrated uncertainty and verifiable explanations to support adoption in regulated, safety-critical settings.
6) Encrypted-traffic analysis: applying generative sequence models to the growing share of encrypted flows, where payload inspection is impossible and behavioural modelling becomes essential.

In summary, Generative AI has moved from a promising idea to a practical pillar of modern network defence. By learning the distribution of traffic rather than merely a decision boundary, it equips security systems to anticipate the unknown. Realising this potential at scale will depend on advances in efficiency, robustness, privacy, and interpretability, the agenda that this work commends to the research community.

REFERENCES

1. Kalal, M. (2024). Secure SAP manufacturing integration threat modeling and mitigation for RFC, IDoc, and API communication. International Journal of Communication Networks and Information Security, 16(4). https://doi.org/10.5281/zenodo.20088018
2. Jaiswal, I. A. (2023). Intelligent Cybersecurity Framework for Large-Scale RESTful Service Architectures. International Journal of Research Radicals in Multidisciplinary Fields, ISSN: 2960-043X, 2(1), 178–184. Retrieved from https://www.researchradicals.com/index.php/rr/article/view, 252.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

3. Kalal, M., & Tanner, O. C. (2025). Autonomous exception management in SAP S/4HANA manufacturing through multi-agent generative AI and event-driven supply networks. International Journal of Core Engineering & Management, 8(4), 130–137.
4. Jaiswal, I. A. (2022). Natural language processing for security policy and log analysis. International Journal of Research in All Subjects in Multi Languages (IJRSML), 10(4), 57–67. https://doi.org/10.63345/ijrsml.v10.i4.1
5. S. Choudhary, S. Kumar, M. Gulhane, and M. Kumar, "Secured automated certificate creation based on multimodal biometric verification," in Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision, pp. 269–281, 2023.
6. M. Kalal, Y. R. Krishna, B. Jayswal, B. Konduru and V. S. A. Kondru, "Metaheuristic Optimization-Driven Information Management System for Enterprise Intelligence," 2026 IEEE 15th International Conference on Communication Systems and Network Technologies (CSNT), Al-Khobar, Saudi Arabia, 2026, pp. 1417–1423, doi: 10.1109/CSNT69054.2026.11502494.
7. Jaiswal, I. A. (2023). Multilingual and culturally adaptive AI models for global education platforms. International Journal for Research in Education (IJRE), 12(9), 17–27. https://doi.org/10.63345/ijre.v12.i9.1
8. Khan, S. (2017). The role of privacy-preserving techniques in database security: Secure multi-party computation, differential privacy, and homomorphic encryption. The Research Journal, 3(4), 29–35.
9. S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR), pp. 1–6, 2024.
10. Khan, S. (2019). Sales force security compliance: An in-depth study of GDPR, HIPAA, and PCI-DSS enforcement in cloud-based CRM systems and their implications for global enterprises. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 8(9), 2211–2219. https://doi.org/10.15662/IJAREEIE.2019.0809010
11. Jaiswal, I. A. (2023). High-Performance AI-Augmented Content Management Systems for Distributed Clouds. International Journal of Multidisciplinary Innovation and Research Methodology, ISSN: 2960-2068, 2(2), 90–97. Retrieved from https://ijmirm.com/index.php/ijmirm/article/view, 243.
12. M. Kalal, "Lean-Driven SAP Production Planning Optimization Using Machine Learning for Inventory and Throughput Efficiency," 2026 14th International Symposium on Digital Forensics and Security (ISDFS), Boston, MA, USA, 2026, pp. 1–6, doi: 10.1109/ISDFS69419.2026.11459029.
13. S. Choudhary, S. Kumar, S. Gowroju, M. Gulhane, and R. Sri Lakshmi, Eds., Genomics at the Nexus of AI, Computer Vision, and Machine Learning. Hoboken, NJ, USA: John Wiley & Sons, 2024.
14. Rallabandi, B. R. (2020). MEC-native 5G systems: Orchestration algorithms for ultra-low latency cloud-edge integration. International Journal of Intelligent Systems and Applications in Engineering, 8(4), 398–408. https://ijisae.org/index.php/IJISAE/article/view/7889
15. Goyal, S., Rallabandi, B., Gattupalli, K. K., & Medarametla, M. S. (2025). Digital twin-enabled context-aware networks: Enhancing smart grid resilience through predictive intelligence. In 2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT) (pp. 1–6). IEEE. https://doi.org/10.1109/ICRTEECT67512.2025.11448880
16. Jaiswal, I. A. (2024). AI-powered observability and incident prediction in distributed enterprise platforms. Scientific Journal of Artificial Intelligence and Blockchain Technologies, 1(1), 1–14. https://doi.org/10.63345/sjaibt.v1.i1.201
17. Kalal, M. (2026, February). Predictive Production Planning in Advanced Manufacturing Using SAP PP and Analytics. In 2026 5th International Conference on Sentiment Analysis and Deep Learning (ICSADL) (pp. 1363–1370). IEEE.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

18. Khan, S. (2018). The role of zero-trust models in database security: Eliminating implicit trust and enforcing continuous verification in enterprise data access systems. International Journal of Research in Electronics and Computer Engineering, 6(1), 1622–1628.
19. S. Kumar, S. Choudhary, S. Gowroju, and A. Bhola, "Convolutional neural network approach for multimodal biometric recognition system for banking sector on fusion of face and finger," in Multimodal Biometric and Machine Learning Technologies: Applications for Computer Vision, pp. 251–267, 2023.
20. C. H. Neetha and M. P. Kantipudi, "Adaptive Edge-Based Bilinear Interpolation for Smart Healthcare," in IET Conference Proceedings CP824, vol. 2022, no. 26, Stevenage, U.K.: The Institution of Engineering and Technology (IET), 2022, pp. 7–13.
21. M. S. Prashanth, R. Aluvalu, and M. V. V. Kantipudi, "Enhancing Health Product Traceability on the Blockchain: A Novel Approach for Supply Chain Management Inspection to AI," EAI Endorsed Transactions on Pervasive Health & Technology, vol. 10, no. 1, 2024.
22. S. Velamuri, S. K. Sudabattula, M. V. V. Prasad Kantipudi, and N. Prabaharan, "Q-Learning Based Commercial Electric Vehicles Scheduling in a Renewable Energy Dominant Distribution Systems," Electric Power Components and Systems, pp. 1–14, 2023.
23. Swathi, S. Kumar, S. Rani, A. Jain, and R. K. M. V. N. M., "Emotion classification using feature extraction of facial expression," in Proceedings of the 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 283–288, 2022.
24. R. Aluvalu, R. Venkatachalam, V. Uma Maheswari, R. J. Anandhi, M. V. V. Kantipudi, and S. Prashanth Mallellu, "IWHO-Based Cluster Head Selection for Vehicle-to-Vehicle Communication in Intelligent Transport System," International Journal of Transport Development and Integration, vol. 8, no. 2, pp. 154–166, 2024.
25. Khan, S. (2016). A study of data provenance and integrity in database security: Ensuring authenticity, non-repudiation, and accountability in data lifecycle management. International Journal of Research in Electronics and Computer Engineering, 4(3), 180–186.
26. Rana, M., Srinivas, S., Jamili, L. K., Jaiswal, I. A., Nakka, S., & Kasetti, S. (2025, May). Real-Time Monitoring and Prediction of Blood Sugar Levels in Diabetic Patients with Functional Models. In 2025 International Conference on Engineering, Technology & Management (ICETM) (pp. 1–6). IEEE.
27. Tripathi, N., Gattupalli, K. K., Rallabandi, B., & Medarametla, M. S. (2025). AI-native Cloud-RAN orchestration for enterprise private 5G using digital twin models. In 2025 1st IEEE Uttar Pradesh Section Women in Engineering International Conference on Electrical Electronics and Computer Engineering (UPWIECON) (pp. 851–857). IEEE. https://doi.org/10.1109/UPWIECON67212.2025.11390348
28. N. Pachauri, V. Suresh, M. V. V. Prasad Kantipudi, R. Alkanhel, and H. A. Abdallah, "Multi-Drug Scheduling for Chemotherapy Using Fractional Order Internal Model Controller," Mathematics, vol. 11, no. 8, Art. no. 1779, 2023.
29. V. Saini, A. Jain, Anurag, and M. V. V. Prasad Kantipudi, "An Efficient Approach for Improving the Performance of Autonomous Vehicle Using Advanced Computer Vision," in International Conference on Soft Computing and Pattern Recognition, Cham, Switzerland: Springer Nature Switzerland, 2023, pp. 164–174.
30. S. Choudhary, C. H. Kandikattu, S. Kumar, M. V. V. P. Kantipudi, and M. Kumar, "Enhancing cybersecurity through combined convolutional neural network-gated recurrent unit approach for distributed denial of service attack detection," in Proceedings of the 2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR), pp. 1–6, 2024.
31. Singh, M., Rallabandi, B., Gattupalli, K. K., & Sai Medarametla, M. (2025). Autonomous private 5G field area networks: Achieving secure, scalable, and self-healing smart grid communications. In 2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT) (pp. 1–6). IEEE. https://doi.org/10.1109/ICRTEECT67512.2025.11448712

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

32. Rallabandi, B. R. (2018). Joint deployment and operational energy optimization in heterogeneous cellular networks under traffic variability. International Journal of Communication Networks and Information Security, 10(3), 638–647. https://ijcnis.org/index.php/ijcnis/article/view/8604
33. Jaiswal, I. A. (2024). Self-healing REST services using artificial intelligence in multi-cloud environments. Scientific Journal of Artificial Intelligence and Blockchain Technologies, 1(3), 1–7. https://doi.org/10.63345/sjaibt.v1.i3.201
34. Goyal, S., Gattupalli, K. K., Rallabandi, B., & Medarametla, M. S. (2025). Integrating terrestrial and non-terrestrial networks for reliable rural coverage in agriculture and smart grids. In 2025 1st IEEE Uttar Pradesh Section Women in Engineering International Conference on Electrical Electronics and Computer Engineering (UPWIECON) (pp. 846–850). IEEE. https://doi.org/10.1109/UPWIECON67212.2025.11390331
35. M. V. V. Prasad Kantipudi, S. Vemuri, N. S. Pradeep Kumar, S. Sreenath Kashyap, and S. Eslamian, "Proposing Model for Water Quality Analysis Based on Hyperspectral Remote Sensor Data," in Handbook of Hydroinformatics, Elsevier, 2023, pp. 317–324.
36. H. K. Jani, M. V. V. Prasad Kantipudi, G. Nagababu, D. Prajapati, and S. S. Kachhwaha, "Simultaneity of Wind and Solar Energy: A Spatio-Temporal Analysis to Delineate the Plausible Regions to Harness," Sustainable Energy Technologies and Assessments, vol. 53, Art. no. 102665, 2022.
37. Cherladine, K., Rallabandi, B. R., Goel, A., Kaushik, K., Dhillon, A., & Soni, M. (2025). Generative adversarial defense models for simulated cyberattack scenarios in virtual networks. In 2025 International Conference on Digital Innovations for Sustainable Solutions (ICDISS) (pp. 1–6). IEEE. https://doi.org/10.1109/ICDISS68238.2025.11320686
38. Desai, A. B., Rallabandi, B., Gattupalli, K. K., & Medarametla, M. S. (2025). Agentic AI for rural connectivity and spectrum-aware networks to power smart agriculture ecosystems. In 2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT) (pp. 1–6). IEEE. https://doi.org/10.1109/ICRTEECT67512.2025.11448879
39. M. Kalal, "Integrating SAP PP and SAP Digital Manufacturing for Lean Production Optimization," 2026 International Symposium of Systems, Advanced Technologies and Knowledge (ISSATK), Hammamet, Tunisia, 2026, pp. 1–7, doi: 10.1109/ISSATK69667.2026.11566800.
40. S. Rani, D. Ghai, and S. Kumar, "Reconstruction of wire frame model of complex images using syntactic pattern recognition," in IET Conference Proceedings CP791, vol. 2021, no. 11, pp. 8–13, 2021.