Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

GENERATIVE AI FOR INTELLIGENT THREAT DETECTION: A HYBRID DEEP LEARNING FRAMEWORK FOR SECURE DIGITAL SYSTEMS

Shaafia Malik[*1], Ameera Shafique[2], Hussain Shah[3], Wahab Ali[4]

[*1]*University of Sialkot, Pakistan*
[2]*Department of Computer Science, Faculty of Computing and Information Technology, University of the Punjab, Lahore, Pakistan*
[3]*Shaikh Zayed Islamic Centre, University of Peshawar*
[4]*BSc Hons Computer Science, Faculty of Technology, Uni of Portsmouth UK*

[*1]shaafiamalik@gmail.com, [2]mscsf21m518@pucit.edu.pk, [3]safihussain@uop.edu.pk, [4]alsa7hr01@gmail.com

DOI: https://doi.org/10.5281/zenodo.21912338

Keywords
Generative Artificial Intelligence (Generative AI), Intelligent Threat Detection, Hybrid Deep Learning, Cybersecurity, Secure Digital Systems, Explainable Artificial Intelligence (XAI), Trustworthy AI

Article History
Received: 27 January 2026
Accepted: 12 March 2026
Published: 26 March 2026

Corresponding Author: *
Shaafia Malik

Abstract

As organizations undergo quick digital transformation and cyber threats grow in sophistication, there is a growing need for intelligent, adaptive, and trustworthy cybersecurity solutions. Traditional threat detection technologies are prone to failing to detect new attack vectors, zero-day exploits, and sophisticated cyber attacks because they are based on signatures and only adapt to known threats. The recent developments in Generative Artificial Intelligence (Generative AI) and deep learning offer new avenues to improve intelligent threat detection, including dynamic pattern recognition, anomaly detection, and predictive cyber defense. Prior work tends to emphasize the detection performance of AI-based cybersecurity models while largely neglecting the strategic use of Generative AI in more secure and explainable threat detection. This study aims to fill this gap by conducting a qualitative Systematic Literature Review (SLR) on the use of Generative AI for creating intelligent and secure digital systems. The research design is qualitative with an interpretivist paradigm, and the study is systematic in synthesizing the major peer-reviewed literature from the various academic research databases to gain insight into trends, methodological advances, challenges, and future research directions. The results highlight five predominant themes: the disruptive impact of Generative AI on cybersecurity, hybrid deep learning solutions for intelligent threat detection, explainability and trust in AI-powered security systems, hurdles in secure digital environments, and future prospects for adaptive cyber defense. The review shows the benefits of combining Generative AI and hybrid deep learning methods in boosting detection accuracy, bolstering resistance to advanced cyber threats and aiding in proactive security decision-making. The study extends the ongoing debate on intelligence in cybersecurity in both theoretical and practical aspects, and offers insights for researchers, practitioners, and policymakers to design intelligent security systems that are transparent, adaptive, and trustworthy. This paper stands out for its qualitative synthesis of existing evidence and the suggestion of a conceptual framework to support the integration of Generative AI into hybrid

deep learning systems, thereby enhancing secure digital systems. The study overall finds that Generative AI holds significant promise for changing the paradigm of intelligent threat detection, and for fostering resilient, secure, and trustworthy digital ecosystems.

Introduction

Modern society has undergone rapid digital transformation, and there is a growing reliance on interdependent digital systems by governments, organisations and individuals. Cloud computing, Internet of Things (IoT), edge computing, fifth-generation (5G) communication, and artificial intelligence (AI) are some of the technologies that have transformed business operations and digital services, providing intelligent automation, real-time decision-making, and large-scale data processing. These technological developments have increased efficiency and innovation, but they have also introduced and made the cyberattack surface larger and exposed digital infrastructures to more and more advanced and dynamic security threats. Cybercrime is one of the most pressing issues in modern digital ecosystems, and attackers are using advanced techniques such as ransomware, phishing, malware, distributed denial-of-service (DDoS) attacks, advanced persistent threats (APTs), and AI-based attacks (Goodfellow et al., 2016; Kolias et al., 2017; Alaba et al., 2017).

Traditional cyber security solutions, such as signature-based intrusion detection systems, rule-based security mechanisms and classic machine-learning approaches, have played an important role in securing digital infrastructures. These techniques often fail to identify zero-day attacks, fast-changing malware variants, and very sophisticated cyber attacks that continuously change their behaviour. They typically rely on pre-defined attack signatures and manually created features, making it hard for them to respond to new cyber threats in complex digital environments. To cope with the ever-changing network behaviors, researchers have started focusing more on artificial intelligence and deep learning to create adaptive and intelligent threat detection systems that learn from the network behaviors (Sommer & Paxson, 2010; LeCun et al., 2015).

Deep learning revolutionized cyber security by allowing automated feature extraction and high-dimensional pattern recognition without the need for much manual effort. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Autoencoders, and Transformer-based models have proven highly effective in intrusion detection, malware classification, anomaly detection, and threat intelligence. These models are able to represent intricate relationships in network traffic and security logs, with the ability to detect known and novel cyber attacks more accurately than numerous traditional machine learning approaches (LeCun et al., 2015; Goodfellow et al., 2016). However, there are some limitations with individual deep learning models. For instance, CNNs are very effective at capturing spatial features, but lack the ability to capture long-term temporal features; recurrent architectures are good at capturing temporal features, but they tend to require more computational resources and more time to train.

To address these challenges, researchers are increasingly turning to hybrid deep learning architectures, combining several different neural architectures that leverage their respective strengths. The popular approach is to integrate various components like CNNs with LSTMs, attention mechanisms, Transformer architectures, or Autoencoders to leverage spatial, temporal, and contextual data in cybersecurity. These integrated architectures have now been shown to perform better on detection accuracy, with greater generalization performance and resistance to advanced attack strategies than stand-alone deep learning models. As the adoption of deep learning in cybersecurity becomes more widespread, the need for adaptive intelligent cybersecurity systems is becoming more prevalent, driven by the need to operate efficiently and meet the growing complexity of the digital threat landscape.

Recently, Generative Artificial Intelligence (Generative AI) has become one of the most game-changing AI technologies. Generative AI differs from traditional discriminative models, which mainly classify or predict data, by learning the underlying data distribution and creating new data such as images, code, network traffic, and even synthetic sets. AI has seen a dramatic boost in functionality in various fields, including healthcare, education, finance, software engineering, and cybersecurity, through technologies like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), diffusion models, and Large Language Models (LLMs) (Goodfellow et al., 2014; Bommasani et al., 2021). Providing realistic synthetic data, simulating complex scenarios, and modeling evolving patterns, offer valuable opportunities to tackle persistent cybersecurity challenges.

In the field of cybersecurity, Generative AI is rapidly becoming a valuable asset for improving intelligent threat detection. It has been studied for its potential in creating synthetic attack datasets, enriching imbalanced cybersecurity data, mimicking adversarial behavior, bolstering malware detection, enhancing anomaly detection, and automated threat intelligence. Generative models are able to learn complex relationships from large-scale cybersecurity datasets and thus can detect subtle relationships that might not be captured by traditional detection methods. In addition, Generative AI can help boost the robustness of hybrid deep learning models by creating a wider variety of training samples to help make the model more generalizable and resistant to new threats. This is especially useful in situations where getting a holistic and balanced cybersecurity data set is difficult.

While there is great potential for Generative AI in cybersecurity, there are also significant challenges to consider. While the same technologies that can bolster cyber defense can be leveraged by malicious actors to automate phishing campaigns, create sophisticated malware, produce convincing social engineering content and bypass authentication, they can also be used to create highly realistic deepfakes for cyber deception. Thus, Generative AI is a potential opportunity for defense as well as a new threat in the cyber security arena. This dual-use nature has led to an increased focus on responsible AI development, the development of trustworthy intelligent systems, and safe deployment methods that can maximize the defense benefits while also minimizing potential misuse.

The need for intelligent, adaptive, and trustworthy cybersecurity frameworks has never been greater as digital systems proliferate in critical infrastructure, financial services, healthcare, smart cities, industrial automation, and government functions. Security solutions must not only be able to detect known attacks, but they also need to be able to learn and adapt to new attacks continually as they appear, through intelligent decision-making processes. The synergy of Generative AI and hybrid deep learning holds great potential for creating future-thinking threat detection solutions that can boost predictive capabilities, adaptability, and cyber resilience in today's digital landscape. Despite the increasing body of AI research in cybersecurity, there are significant conceptual and methodological gaps. The research problem that can be explored in the next section will be based on these issues.

While AI in cybersecurity is a rapidly evolving field, current studies highlight several critical challenges that hinder the implementation of AI-powered threat-detection systems. The existing literature has largely focused on enhancing the detection accuracy by constantly evolving deep learning architectures with minimal attention paid to the integration of Generative AI in a broader cybersecurity framework. Moreover, the majority of proposed models are tested in controlled experimental settings on benchmark datasets, often neglecting real-world deployment issues, including dynamic model adaptation, data imbalance, adversarial manipulation, computational efficiency, and continuous model adaptation. These constraints hinder the adoption of existing solutions in the rapidly changing digital landscape where cyber threats

continuously evolve (Sommer & Paxson, 2010; Arrieta et al., 2020).

A prominent challenge highlighted in the literature is the lack of exploration into hybrid deep learning architectures that successfully fuse the strengths of various neural architectures and Generative AI methods. While Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, Autoencoders, Transformers, and Generative Adversarial Networks (GANs) have separately yielded encouraging outcomes, there are relatively few studies that critically explore the feasibility and challenges of implementing an integrated intelligent threat detection system. Previous studies tend to focus on these technologies in isolation, leaving a disjointed understanding of their complementary attributes for improving cybersecurity effectiveness. As a result, there is a need to have a comprehensive synthesis that explains the combination of Deep Learning and Generative AI can be used together, and how these can collectively bolster secure digital systems.

This research gap is even more acute in the face of the growing sophistication of cyber threats. Cyberattacks are becoming more sophisticated, automated, and stealthy, often using AI to bypass traditional security measures. At the same time, enterprises are producing vast amounts of disparate security information across cloud deployments, IoT devices, enterprise networks, and edge deployments. Organizations are also continuing to collect huge amounts of diverse security information across cloud platforms, IoT devices, enterprise networks, and edge environments. These intricate datasets demand a sophisticated approach to data analysis, and intelligent systems are required to continuously learn, detect novel attack patterns, and adapt to unseen threats. The developments reinforce the critical role of AI in cybersecurity solutions that are more than just about detection and serve proactive cyber defense strategies.

The value of the study is that it adopts a qualitative approach to analyze Generative AI as a new and promising technology that can revolutionize the way that secure digital systems detect threats in a smart way. Unlike studies that center on algorithmic performance, this research critically synthesizes existing scholarly evidence to explore the opportunities, challenges, research trends, and applications of bringing Generative AI into hybrid deep learning approaches. The study's systematic analysis of the existing literature enhances the overall understanding of the potential for using advanced AI technologies to enhance cybersecurity resilience and enable adaptive and intelligent decision-making.

This research is both significant in theory and practice, as it unifies the concepts of Generative AI, hybrid deep learning, intelligent threat detection, and secure digital systems to provide an analytical overview of AI-driven cybersecurity. The review builds on current discussions by recognizing common themes in the research, methodological strengths, areas for continued work, and areas that are identified as emerging opportunities for research that have not been widely studied. The study brought together information from different fields to develop a more solid conceptual basis for future studies on AI-based cybersecurity.

In the practical realm, the results provide real-world insights for cybersecurity experts, AI researchers, system architects, and management stakeholders aiming for intelligent threat detection solutions. The synthesis will examine the benefits of using hybrid deep learning models, with the assistance of Generative AI, that can enhance accuracy in detection, bolster cyber resilience, shorten response times, and boost adaptive security. The research offers guidance to organisations seeking to modernise their cyber Security systems to tackle the increasing sophistication of digital threats.

The research provides policy guidance on how to cultivate responsible AI governance for the ethical, secure, and transparent use of cutting-edge AI applications in cybersecurity. The results could help policymakers develop regulations and best practices for the safe use of Generative AI in critical digital systems, ensuring that the benefits outweigh the risks. The findings could inform the development of regulations and best practices for the safe use of Generative AI in critical digital

systems, providing a framework for governments and regulatory agencies to implement as they increasingly prioritize trustworthy AI in their policies.

The main goals of this study are, therefore, to review and critically analyze the prevailing body of research on Generative Artificial Intelligence for Intelligent Threat Detection and to integrate the knowledge gained from existing research in the field of Generative AI and hybrid deep learning for securing contemporary digital systems. In particular, the study will aim at developing key research themes, assessing current research methods, analysing existing challenges and limitations, revealing theoretical and practical implications, and defining new research directions that would contribute to the adaptive, intelligent and resilient cybersecurity systems development.

Originality of this research is in synthesis of evidence from leading peer-reviewed studies by a qualitative method rather than proposing and experimental testing of one detection algorithm. The study combines the insights of different voices on Generative AI, Hybrid Deep Learning, and Cybersecurity, creating a holistic perspective of how AI is evolving in the field of intelligent threat detection. This comprehensive view enables the integration of various technological advancements and aligns with the increasing need for secure, adaptive, and trustworthy digital systems.

The rest of this article is structured as follows. The following section provides a critical review of the current literature on Generative AI, hybrid deep learning, intelligent threat detection and cybersecurity. This is followed by the research method which outlines an approach to the study with regards to a qualitative systematic literature review. The results and discussion are summarized as the main themes identified through the literature reviewed in the following section. Lastly, the article provides a summary of the main findings, the implications of the research for theory and practice, limitations of the study, and recommendations for future research.

Literature Review

Digital technologies are evolving at a very quick pace, changing the cybersecurity landscape and making it more essential for organizations to implement sophisticated cybersecurity systems that are capable of detecting such threats. While traditional signature- and rule-based detection systems have been very effective in thwarting known attacks, they often fail to detect zero-day attacks, advanced persistent threats (APTs), polymorphic malware, and fast-changing attack methods. Thus, researchers have turned to an approach involving Artificial Intelligence (AI) and deep learning for improving threat detection in an adaptive and automated fashion by synthesizing patterns (Sommer & Paxson, 2010; LeCun et al., 2015). More recently, Generative Artificial Intelligence (Generative AI) has emerged as a promising technology with the potential to revolutionize cybersecurity, providing support for advanced proactive threat detection, synthetic data generation, and intelligent cyber defense.

Unlike the traditional discriminative AI, generative AI models learn the underlying distribution of data and produce new, realistic data, instead of just sorting old information. Synthetic data generation, code generation, simulation of complex scenarios, and intelligent decision making in various fields has enabled technologies like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), diffusion models, and Large Language Models (LLMs) to achieve unprecedented performance (Goodfellow et al., 2014; Bommasani et al., 2021). In cybersecurity, they can be used to simulate realistic attacks, augment scarce training sets, detect anomalies, analyze malware, and do automatic threat intelligence. With the evolving nature of attack patterns, the ability to adapt cyber security systems to be more resilient and resilient to threats has been brought into focus due to the use of Generative AI, which can enable such systems to learn continuously rather than merely relying on signatures.

While there is definite potential for Generative AI to be used in cybersecurity, the views on how this should be used are split between various

existing studies. Some studies have found it to be effective in improving the defense capability due to its realistic attack simulation and improved model training, while others have pointed out issues of its misuse to boost the capability of cybercriminals. The same technologies that can help to identify a threat can be used to create a high level attack of phishing, malicious code, deepfake content, and automated social engineering attacks. The dual-use nature is one of the biggest challenges with Generative AI and has encouraged increasing conversations about responsible AI development, secure implementation, and AI governance in a cybersecurity context (Arrieta et al., 2020). For this reason, there is an increasing push in recent literature for the ethical and security aspects to be taken into account alongside technological innovations.

As Generative AI has arisen, deep learning has been the top method for intelligent threat detection. In contrast to traditional machine learning algorithms that rely heavily on manual feature engineering, deep learning models automatically learn the hierarchical feature representation from large-scale security data. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Autoencoders, and Transformer-based architectures have proven to be effective architectures for malware classification, intrusion detection, anomaly detection, and network traffic analysis in comparison to traditional machine learning approaches (LeCun et al., 2015; Goodfellow et al., 2016). They are capable of identifying intricate relationships between non-linear variables, making them valuable for uncovering complex attack patterns that are often missed by traditional methods.

But none of the single deep learning architectures is universally best suited for all cybersecurity situations. CNNs are efficient at identifying spatial features in the network traffic, and less efficient at capturing long-range sequential dependencies. In contrast, the temporal relationships in network behavior can be learned with LSTM and other recurrent architectures, which tend to be more resource-intensive and time-consuming to train. In recent years transformer models have shown great promise in sequence modeling, but typically require significant computational resources and large training sets. To overcome these challenges, researchers have sought alternative deep learning architectures like hybrid deep learning frameworks, which combine several architectures to leverage their complementary benefits and enhance the overall detection performance.

Considering the potential, hybrid deep learning models have grown in popularity due to their ability to combine strengths from various neural architectures into a single model. In CNN-LSTM models, automatic feature extraction is coupled with sequential pattern learning for better identification of cyber threats. Likewise, the Autoencoder-CNN and Transformer-LSTM architectures have shown strong results in anomaly detection and malware analysis tasks, extracting both spatial and temporal information in addition to contextual information. Comparative studies typically show that hybrid models perform better than stand-alone algorithms in terms of detection accuracy, robustness, and generalization ability. However, added architectural complexity, increased computational demands, and decreased interpretability make application in resource-limited digital environments impractical.

The rise in the use of Cloud computing, Internet of Things (IoT) devices, Edge computing, and smart infrastructures has further added to the cybersecurity requirements. Today's digital systems produce vast quantities of diverse and continually changing security information that is too difficult to monitor manually. Hence, intelligent threat detection has become a crucial part of modern cybersecurity measures. Intelligent threat detection systems are different from traditional security methods that react only to known attacks, as they continuously track behavioral patterns, detect anomalies, and adapt to new threats that have not been encountered before, through sophisticated AI techniques. The ability to learn adaptively has been consistently identified as the key missing component to

protect against today's cyberattacks, which are fast-moving and increasingly sophisticated (Kolias et al., 2017).

While previous research consistently shows that deep learning can be very effective for intelligent threat detection, most of them pay significant attention to improving the predictive performance and comparatively less attention to practical deployment. While detection accuracy is reported most often, the other metrics, such as scalability, computational efficiency, adversarial robustness, transparency and real-time implementation, are often reported far less often. Many proposed cybersecurity frameworks perform well in experiments, but have issues in production environments with diverse devices, fluctuating traffic, constrained computing power, and ever-changing threats.

On a theoretical level, the current studies concerning AI in the field of cybersecurity are being influenced by adaptive learning theory, intelligent decision making, and data-driven security approaches. The views expressed here indicate that cybersecurity systems must also continuously evolve their learning, adapting to new threats, and dynamically adapting their detection strategies based on the changing nature of attacks. This point is even more compelling when combined with the ability of generative AI to continuously generate synthetic attack scenarios and realistic cybersecurity data, enhancing the learning ability and resilience of intelligent detection models. However, researchers are increasingly pointing out that future AI-powered cybersecurity solutions need to be explainable, reliable and trustworthy, and still be able to deliver the necessary predictive capabilities to ensure their adoption in security-critical environments.

In conclusion, the literature surveyed shows significant advances in the use of Generative AI and hybrid deep learning for intelligent threat detection. However, there is still a lot of existing research that looks at Generative AI, Hybrid architectures, and Cybersecurity applications separately without them being an integral part of an intelligent security system. Moreover, little qualitative research in this area will critically analyse the overall contribution of these technologies to secure digital systems. In the next part of the literature review, these observations are used to analyse the topic of explainability, trustworthy AI, limitations of existing research, and the necessity of a combined Generative AI–Hybrid Deep Learning Framework.

A key area of focus in recent cybersecurity research is on the concept of Explainable Artificial Intelligence (XAI) and trustworthy AI. Deep learning and Generative AI have boosted the effectiveness of intelligent threat detection systems, but most of these deep learning models are considered "black boxes" and lack explanations of their predictions. In the realm of cybersecurity, where choices can impact critical infrastructure, monetary systems, healthcare, and national security, the capacity to clarify and validate AI-driven choices is gaining significance. Explainable AI allows cybersecurity analysts to understand the predictions made by the model, confirm alerts, find false positives and increase the confidence in automated security solutions (Adadi & Berrada, 2018; Arrieta et al., 2020).

Recent research suggests that explainability can have a role in the accountability and regulatory aspects of it, besides transparency. For explaining complex deep learning predictions, techniques like Local Interpretable Model-Agnostic Explanations (LIME) as well as SHapley Additive exPlanations (SHAP) have gained widespread popularity for revealing – in a way interpretable to an individual user – what factors are driving a given prediction (Ribeiro et al., 2016; Lundberg & Lee, 2017). While these have shown to be successful in a few areas, their use in cybersecurity is less extensive. While many current threat detection models focus on making accurate predictions, they are not easy to interpret, which makes them difficult for organisations to trust and rely on for decision-making.

Another common topic is the idea of "trustworthy AI", which is more than just explanatory, but also fair, robust, private, accountable, secure and overseen by humans. Reliable AI has become a critical component in cybersecurity, where intelligent systems are often involved in high-stakes decisions that deal with

sensitive digital assets. AI models should not only be highly accurate in detecting cyber threats, but also resilient to adversarial attacks, adapt to various environments, and provide trustworthy decisions for cybersecurity experts to verify (Arrieta et al., 2020; European Commission, 2019). Although these principles have been gaining in recognition, adherence to them in the realm of intelligent threat detection systems has been inconsistent.

Another challenge found in past research is the availability and quality of cybersecurity data sets (Shah et al., 2025). Typically, deep learning and Generative AI models need large amounts of data, which is diverse and representative to be well-trained. There are, however, a lot of intrusion detection datasets available in the public domain, but many are outdated or have not been balanced or do not cover the current cyber threats. There has been growing interest in using Generative AI (GAI) methods to create synthetic data for cybersecurity and enhance the ability of models to generalize. Generative AI (GAI) techniques, specifically Generative Adversarial Networks (GANs), have been gaining traction to synthesize cybersecurity data and boost the generalizability of models. Though these methods have proven to be promising, there are still doubts about the realism, variety, and bias of synthesized data. Further studies will be needed to develop standardized approaches to assessing synthetic cybersecurity data prior to its general use.

It is found that there are a number of improvements made in the previous studies, but there are also some common drawbacks.Comparative analysis of previous studies shows that there are certain kinds of improvements that have been made and certain common drawbacks as well. Typical machine learning-based studies typically report fewer computational needs and less detection accuracy compared to deep learning models. On the other hand, deep learning methods consistently outperform traditional methods in prediction tasks, often come with higher complexity, lack interpretability, and demand significant training time. More recent studies on Generative AI reveal potential in areas like threat simulation, data augmentation and adaptive learning. Empirical evidence on its effectiveness in the long term and secure deployment in operational cybersecurity environments is sparse, though. The opposing results indicate that no single technology can solve all the problems that come with a smart threat detection.

The other drawback of prior literature is a fragmented research that consists of numerous pieces. There are many studies that look at Generative AI, deep learning, explainability, or security, but fewer that explore aspects of these in a holistic intelligent threat detection setting. In addition, numerous evaluations of proposed models rely on small, synthetic datasets that do not account for deployment considerations like scalability, computational efficiency, vulnerability to adversarial attacks, real-time performance, or interoperability with heterogeneous digital environments. This disjointed research field restricts the ability of developing comprehensive cybersecurity solutions that are able to cover the complexity of contemporary digital infrastructures.

There is also a methodological imbalance in the literature. The majority of the published studies use quantitative experimental designs and focus on the main performance measures of classification accuracy, precision, recall, F1-Score and detection rate. These measures offer great insights into the efficacy of the algorithm but they fail to capture some conceptual challenges such as algorithm explainability, organizational uptake, algorithm governance, ethical implications, and user trust. Yet, to date, there are few quantitative systematic literature reviews with a strong methodological design that carry out critical synthesis of multidisciplinary research and can reveal new research trends and theoretical understanding.

This observation of the activities highlights some key gaps in research. First, there is a lack of qualitative synthesis that explores the combined contribution of Generative AI, hybrid deep learning, explainable AI and trustworthy AI to intelligent threat detection. Second, it is not known what existing research provides in terms

of integrating these technologies into a single framework that can be used to secure modern digital systems. Thirdly, the opportunities and risks of Generative AI, in particular, as a means to defend against cyber threats as well as a weapon of cybercrime, have not been critically examined by a relatively small number of studies. Lastly, there is limited literature that discusses the governance, transparency and the responsible use of AI in next generation cybersecurity systems.

To address these gaps, a more integrated view must be taken that goes beyond simply improving detection accuracy. The future intelligent threat detection systems must not only be adaptive to the capabilities of Generative AI, but must also be built on the complementary strengths of hybrid deep learning architectures, and must also be designed with explainability, robustness, and trustworthiness as core features. This integration could help to improve the security of the ever-increasingly connected digital world, to make digital environments more resilient against cyber threats, to boost analysts' confidence, to support responsible AI deployments, and to enhance analysts' value.

In view of this, this study suggests a qualitative synthesis of the latest literature to gain a holistic view on the interplay of Generative AI and hybrid deep learning in intelligent threat detection within secure digital systems. This review helps to establish the theoretical basis of the proposed framework by examining some of the critical comparisons of the existing studies, dominant themes, strengths and weaknesses of methods used, and unresolved challenges to be addressed. By doing so it helps to close the divide between the rapidly advancing AI technologies and the escalating cybersecurity demands of today's intricate digital environments, all while maintaining transparency, adaptability, trust, and security.

Research Methodology

This study used a research philosophy that is interpretivist, meaning that knowledge is created when existing knowledge and scholarly evidence is critically interpreted and synthesized rather than measured. The interpretivist approach is well suited for studying the evolving role of Generative Artificial Intelligence (Generative AI) in intelligent threat detection because it allows for a deeper understanding of how researchers have conceptualized, implemented and assessed AI-based cybersecurity solutions in various contexts. This study was not an empirical experiment, but rather, it critically analyzed the existing literature in order to identify the major ideas, trends, limitations, and areas of future research on the topic of hybrid deep learning and secure digital systems (Creswell & Creswell, 2018; Saunders et al., 2019).

The research design used was a qualitative Systematic Literature Review (SLR) to guarantee a rigorous, transparent and repeatable synthesis of the most current knowledge. An SLR is not a narrative review but uses a systematic approach to finding, screening, appraising and synthesizing relevant literature with a minimum of selection bias. This method is particularly appropriate, as the purpose of the study was to critically discuss the development of Generative AI, hybrid deep learning, explainable AI, and intelligent threat detection in the scope of cybersecurity research, but not to statistically assess the performance of these models. The review was carried out following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA2020) guidelines in order to enhance methodological transparency and quality of the review report (Page et al., 2021).

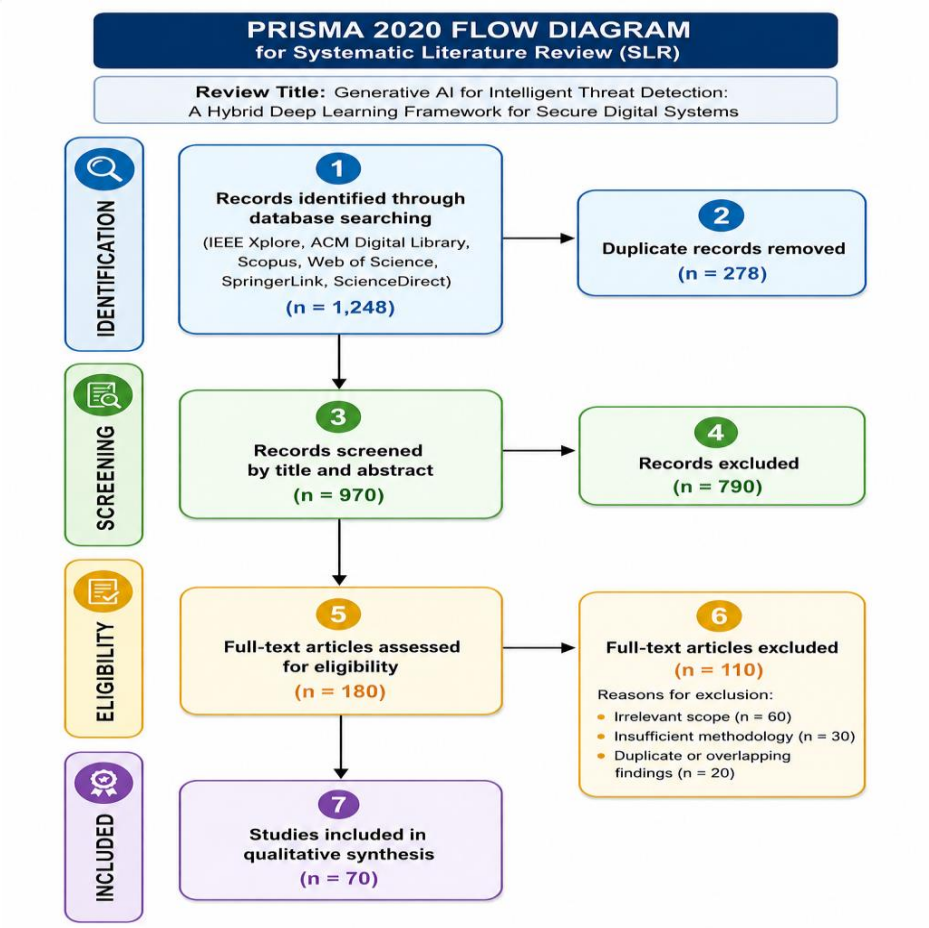
Six world-renowned academic databases had been used for performing the literature search namely IEEE Xplore, ACM Digital Library, Scopus, Web of Science, SpringerLink, and ScienceDirect. These databases were chosen due to their quality and the peer-reviewed publications in the field of artificial intelligence, cyber security, computer science, information systems and digital technologies. Since the review was aimed at capturing the latest advances, only studies published from 2020 to 2026 were given priority, while seminal foundational works that predate 2020 but make significant contributions to the theories and methodologies underpinning Generative AI, deep learning, and intelligent

cybersecurity were included if they were important.

A detailed search strategy was created with predefined keywords, linked by boolean operators. The main keywords were: "Generative Artificial Intelligence", "Generative AI", "Hybrid Deep Learning", "Deep Learning", "Cybersecurity", "Intelligent Threat Detection", "Intrusion Detection", "Network Security", "Explainable Artificial Intelligence", "Trustworthy AI", "Digital Security", and "Secure Digital Systems". Boolean operators (AND, OR, NOT) were used to enhance precision in searches and to obtain studies that were directly relevant to the research objectives. Backward reference searching was also conducted, examining reference lists of highly relevant publications for possible additional influential publications that were not identified in the initial database search (Kitchenham & Charters, 2007).

In order to be consistent and methodologically rigorous, inclusion and exclusion criteria were defined before the literature search was done. Studies were considered for inclusion in the review if they were peer-reviewed journal articles or conference papers written in English that addressed Generative AI, hybrid deep learning, intelligent threat detection, explainable AI, trustworthy AI, cybersecurity, or secure digital systems and provided adequate methodological and conceptual detail that would be relevant to the aims of this review. Editorial articles, opinion papers, theses, unpublished reports, duplicate records, and publications that were not related to AI tools in cybersecurity were not included in the studies. Articles that did not include the necessary methodological details and/or focused solely on a different AI application were also omitted to increase the quality and relevance of the review (Page et al., 2021).

All studies were selected according to PRISMA 2020 guidelines. First, records from the selected databases were downloaded to a reference management system and then the duplicates were eliminated. The titles and abstracts were then read through to exclude studies that failed to meet the predetermined inclusion criteria. So the other articles that were in full text were carefully scrutinized for methodological quality, conceptual relevance and contribution to the research objectives. The final qualitative synthesis included only studies that met all of the criteria for inclusion. The PRISMA flow diagram offers an informative overview of identification, screening, eligibility assessment and final inclusion steps in the systematic review process (Page et al., 2021).



The review process of synthesizing results was an important stage in the quality assessment, so as to ensure the use of credible and methodologically sound evidence in synthesizing results. Clarity of research aim, methodological soundness, transparency of data collection and analysis, validity of results, theoretical contribution, and relevance to intelligent threat detection and Generative AI were used to critically assess each selected study. Highly cited articles from well recognized Scopus, Web of Science, IEEE, ACM, Springer, and Elsevier indexed journals and conference proceedings were preferred. This systematic assessment minimised the risk of the incorporation of poor quality evidence and improved the overall trustworthiness of the review (Kitchenham & Charters, 2007; Xiao & Watson, 2019).

Relevant information was systematically extracted after the quality assessment, using a structured data extraction framework. The extracted information consisted of publication information, research purpose, AI approaches used, deep learning architectures, application areas, datasets, evaluation techniques, main results, reported limitations, and future research recommendations. The standardised structure allowed for comparison between the studies and trends in the literature to be detected.

Thematic analysis is a qualitative analytical method that is widely used for identifying themes and patterns that emerge across several studies (Braun & Clarke, 2006), which was the method used in synthesizing the extracted evidence. The first step was to read the selected articles repeatedly so that the content of the articles was familiar. Relevant statements regarding

Generative AI, hybrid deep learning, cybersecurity, intelligent threat detection, explainability, and trustworthy AI were systematically coded. Codes were then further amalgamated into higher-level themes reflecting the most salient trends in the literature examined. Themes were continually developed, added, and adjusted throughout the analysis process, using comparisons across studies to capture the evidence whilst maintaining the themes' important similarities and differences to previous studies (Nowell et al., 2017).

Several approaches were used to ensure the credibility, dependability, transparency and trustworthiness of the review. The use of peer reviewed publications from well-known data bases and a critical comparison of evidence from several studies enhanced credibility. Documentation of all steps in the review process, including search strategy, eligibility criteria, screening process, and analysis, ensured dependability and allowed future researchers to repeat the study. The PRISMA 2020 reporting guidelines were followed and methodological decisions were clearly documented and transparent. Finally, confirmability was strengthened by basing all interpretations on published scholarly work and not personal assumptions, thus offering the results of the literature reviewed (Lincoln & Guba, 1985; Nowell et al., 2017).

Overall, the qualitative systematic literature review offers a sound methodological basis for studying the integration of Generative AI and hybrid deep learning in the context of intelligent threat detection in secure digital systems. The methodology ensures a thorough and systematic synthesis of current research, highlighting existing challenges, future opportunities, and potential pathways for future AI research and practice that can contribute to the design of adaptive, reliable, and resilient AI cybersecurity systems.

Findings and Discussion

Five interconnected themes emerged from the systematic literature review and collectively explained the role of Generative Artificial Intelligence (Generative AI) in the intelligent threat detection in secure digital systems. The reviewed studies all show that deep learning has greatly improved the capabilities of cybersecurity, but combining Generative AI with hybrid deep learning architectures provides greater adaptability, resilience and proactive threat detection. Meanwhile, several challenges have emerged in the literature, such as explainability, security issues, and implementation.

Table 1
Summary of Themes Identified from the Systematic Literature Review

Theme	Key Findings	Supporting Literature	Practical Implications
Generative AI for Intelligent Threat Detection	Generative AI enables adaptive learning, synthetic data generation, and proactive threat identification.	Goodfellow et al. (2014); Bommasani et al. (2021)	Enhances the capability to detect evolving cyber threats.
Hybrid Deep Learning Architectures	Hybrid models outperform individual deep learning techniques in complex cybersecurity tasks.	LeCun et al. (2015); Goodfellow et al. (2016)	Improves detection accuracy and system robustness.
Explainability and Trustworthy AI	Transparent AI increases analyst confidence and supports reliable security decisions.	Arrieta et al. (2020); Ribeiro et al. (2016)	Promotes trustworthy and accountable cybersecurity systems.

Challenges of AI-Driven Cybersecurity	Black-box models, adversarial attacks, and computational costs remain significant challenges.	Lundberg & Lee (2017); Sommer & Paxson (2010)	Encourages development of interpretable and resilient AI models.
Future of Secure Intelligent Systems	Integrating Generative AI with hybrid deep learning supports adaptive and secure digital infrastructures.	Samek et al. (2021); European Commission (2019)	Guides future AI-enabled cybersecurity frameworks and policies.

Theme 1: Generative AI is improving intelligent threat detection

Generative AI has consistently been named in the literature as a game-changing technology in today's cybersecurity landscape. Generative models can learn complex data distributions and generate realistic synthetic cyberattack scenarios, which can enhance the training of intelligent detection systems, in contrast to conventional AI models that mainly concentrate on classification. According to researchers, Generative AI can help with anomaly detection, malware analysis, attack simulation, and threat intelligence through the ability to detect new attack patterns using security models. However, it is important to note that this technology can also be used for malicious purposes, such as creating deepfake videos, phishing attacks, or sophisticated forms of malware (Goodfellow et al., 2014; Bommasani et al., 2021), underscoring the technology's potential for malicious applications. Thus, the literature has focused on the need for responsible deployment strategies that optimize the capability of the defense, while minimizing security risk.

Theme 2: Hybrid Deep Learning Improves Detection Performance

Another theme that stands out is the increasing utilization of hybrid deep learning architectures for intelligent threat detection. The research conducted in the past shows the effectiveness of the multi-models of CNNs, LSTMs and Autoencoders, and Transformer-based models to learn spatial, temporal, and contextual information from cybersecurity data sets more effectively than individual models. All comparative investigations have found that hybrid architectures allow for greater resilience to sophisticated cyber threats, better generalization, and increased detection accuracy. The advantages, though, can be quite expensive and complex to implement, especially in large-scale digital environments (LeCun et al., 2015; Goodfellow et al., 2016). This suggests that further studies are needed to fine-tune hybrid models to improve efficiency and scalability.

Theme 3: The focus is on how explainability and trust form the basis of AI for cybersecurity

The literature has been increasingly aware of the shortcomings of accurate threat detection in real-world cybersecurity applications. AI systems must provide explanations to the security analysts and organizational decision-makers, explaining why they made a particular prediction. In summary, Explainable Artificial Intelligence (XAI) is a vital aspect of building trustworthy cybersecurity systems, as it enhances transparency, accountability, and user confidence. Research on methods like LIME and SHAP indicates that interpretable AI can help analysts validate AI-driven decisions and decrease false alarms (Arrieta et al., 2020; Ribeiro et al., 2016). Even with these advances, there are still many Generative AI and deep learning models that are extremely complex and hard to interpret, making explainability a key design principle in future intelligent threat detection systems.

Theme 4: Challenges that hinder a successful practical application

While AI-powered cybersecurity has shown great promise, some recurring challenges are apparent in the studies reviewed. The practical implementation of Black-box decision-making, susceptibility to adversarial attacks, the scarcity of

high-quality cybersecurity datasets, and the need for computational resources are still limiting factors, along with concerns about privacy and ethical AI deployment. In addition, many proposed models are tested on benchmark datasets using controlled laboratory conditions, and therefore have only limited applicability in the real world (Sommer & Paxson, 2010; Lundberg & Lee, 2017). The results highlight the need for more effective, interpretable, and actionable AI systems that can adapt to the ever-changing nature of cyber threats in future cybersecurity research.

This theme explores the quest for secure and intelligent digital systems. This theme focuses on the pursuit of secure and smart Digital Systems.

The last theme suggests that the next generation of cybersecurity will be a mix of Generative AI and hybrid deep learning in safe, adaptable, and trustworthy digital environments. However, these technologies are not being used in isolation – research is increasingly calling for a more holistic approach including adaptive learning, explainability, automation, and continuous threat intelligence. These systems can be used to enhance the cyber resilience of organizations, speed up incident response, boost security, and ensure compliance with regulatory standards in industries like smart cities, industrial automation, finance, and healthcare (European Commission, 2019; Samek et al., 2021). The holistic approach is in line with the general shift from cybersecurity to cyber defense based on intelligence and proactivity.

In summary, the results highlight the significant advancements in the research field of using Generative AI and hybrid deep learning for intelligent threat detection. The literature, however, is disjointed with a number of studies concentrating on individual technologies, and not integrated cybersecurity solutions. Moreover, there are still gaps in explaining, trustworthiness, governance and challenges in the real-world deployment. The present study builds on these themes, filling the identified research gap and laying the groundwork for a Generative AI–Hybrid Deep Learning Framework integrating adaptive intelligence, powerful detection power, transparency, and secure deployment. The proposed framework thus supports the evolution of next generation cybersecurity systems that can defend the increasingly complex and interconnected digital environment.

Conclusion

This study presents a qualitative systematic literature review of the role of Generative Artificial Intelligence (Generative AI) in intelligent threat detection in particular, and how this is integrated into hybrid deep learning frameworks for securing modern digital systems in general. The review highlights that Generative AI is becoming a transformative technology with the potential to improve cybersecurity through improved machine learning, synthetic data, anomaly detection, and the ability to anticipate future cyber threats. Additionally, the results show that hybrid deep learning models (a fusion of several neural network models) are better at detection than models on their own. The review also identifies that explainability, transparency, and trustworthiness are becoming a key part of the characteristics of AI-based cybersecurity systems, especially in security-critical environments where automated decisions need to be understood and validated by a human (Arrieta et al., 2020; Ribeiro et al., 2016).

This study aims to tackle the identified research gap by merging the disjointed topics of Generative AI, Hybrid Deep Learning, Explainable AI, and Intelligent Threat Detection into a cohesive conceptual framework. Previous work has focused on these technologies individually, but the review herein shows how these technologies can be merged into a comprehensive system to enable adaptive, secure, and trustworthy digital systems. The research integrates insights from the latest peer-reviewed studies, offering a more robust theoretical framework for the potential of Generative AI in enhancing cybersecurity outcomes while addressing issues of transparency, operational resilience, and changing cyber threats.

Theoretically, this study provides a conceptual understanding of the relationship between Generative AI, hybrid deep learning, and

trustworthy intelligent threat detection, thereby advancing the growing field of theoretical research on AI in cybersecurity. From a practical perspective, the findings offer significant insights for cybersecurity professionals, AI developers, system architects, and organizational decision makers looking to deploy intelligent security strategies that promote more accurate threat detection, boost cyber resilience, and enable effective security decision making. The review also provides insights that can help policymakers encourage responsible and secure use of the advanced AI technologies in critical digital infrastructure.

Several limitations should be noted, however, in this contribution. The study relies solely on a qualitative systematic literature review process and thus does not involve empirical experimentation or validation of the proposed framework based on real world cybersecurity data sets. Furthermore, the review has been conducted exclusively on peer-reviewed publications in English language, from the main academic databases, and therefore could have missed relevant publications that have been published in other languages or in other sources.

The integrated Generative AI–Hybrid Deep Learning frameworks need to be empirically validated through various and modern cybersecurity datasets in future studies. Promising avenues for further research include comparative study of various generative models, exploration of adversarial robustness, creating lightweight models for resource-limited settings, and increasing the embedding of principles of explainable and trustworthy AI. Research in these areas will help to develop secure, adaptive, transparent and resilient intelligent cybersecurity systems to meet the escalating challenges of today's digital world (European Commission, 2019; Goodfellow et al., 2016).

REFERENCES

Adadi, A., & Berrada, M. (2018). Peeking inside the black box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. *Journal of Network and Computer Applications, 88*, 10–28. https://doi.org/10.1016/j.jnca.2017.04.002

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bommasani, R., Hudson, D. A., Adeli, E., *et al.* (2021). *On the opportunities and risks of foundation models*. Stanford Center for Research on Foundation Models.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.

European Commission. (2019). *Ethics guidelines for trustworthy AI*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems 27 (NeurIPS 2014)* (pp. 2672–2680).

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black-box models. *ACM Computing Surveys, 51*(5), Article 93. https://doi.org/10.1145/3236009

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele University and Durham University.

Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). DDoS in the IoT: Mirai and other botnets. *Computer, 50*(7), 80–84. https://doi.org/10.1109/MC.2017.201

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)* (pp. 4765–4774).

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. https://doi.org/10.1177/1609406917733847

H. Shah, H. Farman, B. Jan, A. Khalil, and M. M. Nasralla, "Securing the Internet of Things: Deep learning driven intrusion detection with missing data imputation," *IEEE Access*, vol. 13, pp. 146476–146491, 2025, doi: 10.1109/ACCESS.2025.3599931.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., *et al.* (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71 (BMJ)

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K.-R. (2021). *Explaining deep neural networks and beyond: A review of methods and applications. Proceedings of the IEEE, 109*(3), 247–278.

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students* (8th ed.). Pearson.

Schiller, E., Aidoo, A., Fuhrer, J., Stahl, J., Ziörjen, M., & Stiller, B. (2022). Landscape of IoT security. *Computer Science Review, 44*, 100467.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE Symposium on Security and Privacy* (pp. 305–316). https://doi.org/10.1109/SP.2010.25

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research, 39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access, 5*, 21954–21961. https://doi.org/10.1109/ACCESS.2017.2762418

Zhou, Z.-H. (2021). *Machine learning*. Springer.