Journal of Research, Innovation and Technologies
https://www.acadlore.com/journals/JORIT

Assessment of Safety and Privacy Challenges in ANN-Based English Handwriting Recognition

Jingyu Zhang[1*], Mengjiao Li[2]

[1] School of Management, Business Analytics, Universiti Sains Malaysia, 11700 Gelugor, Malaysia
[2] School of Foreign Language, English, Anyang Normal University, 455000 Anyang, China

[*] Correspondence: Jingyu Zhang (zjy923092@gmail.com)

Received: 11-05-2025 **Revised:** 01-26-2026 **Accepted:** 02-05-2026

Citation: Zhang, J. Y. & Li, M. J. (2026). Assessment of safety and privacy challenges in ANN-based English handwriting recognition. *J. Res. Innov. Technol.*, *5*(1), 136–148. https://doi.org/10.56578/jorit050109.

Abstract: Deep neural network-based English handwriting recognition has revolutionised the security or verification system of today; nevertheless, there are certain risks that are inevitable. This paper examines the status of the present technologies in the handwriting recognition systems, and more particularly, by the various deep neural network architectures. It also evaluated common cybersecurity risks such as data poisoning, model inversion, and adversarial attacks, which can be devastating to such systems, as well as common privacy and ethical issues. The potential regulatory compliance and mitigation measures that can be taken to avert these risks and hurdles are also addressed in detail, with requisite emphasis being made on the future outlook of a more secure handwriting recognition system.

Keywords: Deep neural networks; Artificial Neural Networks; Cybersecurity threats; English handwriting recognition; Privacy; Regulatory compliance

JEL Classification: O33, K11, L86

1. Introduction

1.1 Background

Recognising handwriting through automated systems is generally done through Artificial Neural Networks (ANNs), which find unique patterns in handwriting and then convert it to a digital form. In the past, the conventional approaches were rule-of-thumb and statistical models, such as the Hidden Markov models (HMMs) and the Support Vector Machines (SVMs), which needed to be trained by hand and were not very resilient to the transformations in writing patterns by people. With the emergence of deep learning, ANNs, in particular, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have altered this aspect by enabling end-to-end learning of raw pixel or stroke data. These algorithms, too, have been discovered to learn hierarchical representations automatically, which depict spatial and time-dependent dependence among handwritten characters, which are quite handy in the enhancement of the argument of accuracy and robustness.

As shown in Figure 1, Altwaijry & Al-Turaiki (2020) performed handwriting character recognition through a typical CNN architecture on the AHCD dataset, which produced exceptionally high accuracy over other models. Thus, ANN-based handwriting recognition has found a core application in automated document scanning, cheque clearing in banks, academic testing systems, and digital identity management. Being a form of human-computer interaction, it borders the area of cognitive perception and computational intelligence and results in the educational and business world, where natural writing is required.

1.2 Problem Statement

Even though the ANN-based handwriting recognition systems have been facing an increasing risk to cybersecurity and privacy, their accuracy is phenomenal. They include vulnerability to data poisoning, model

inversion, and adversarial attacks, which can interfere with the integrity of systems and the privacy of users. Moreover, the handwriting data is normally sensitive and personal, such that it can be vulnerable to an attack under the law of data protection. Therefore, in order to achieve safe, ethical, and legal implementation of these systems in reality, it is necessary to assess and minimise these risks, which is performed in this study.

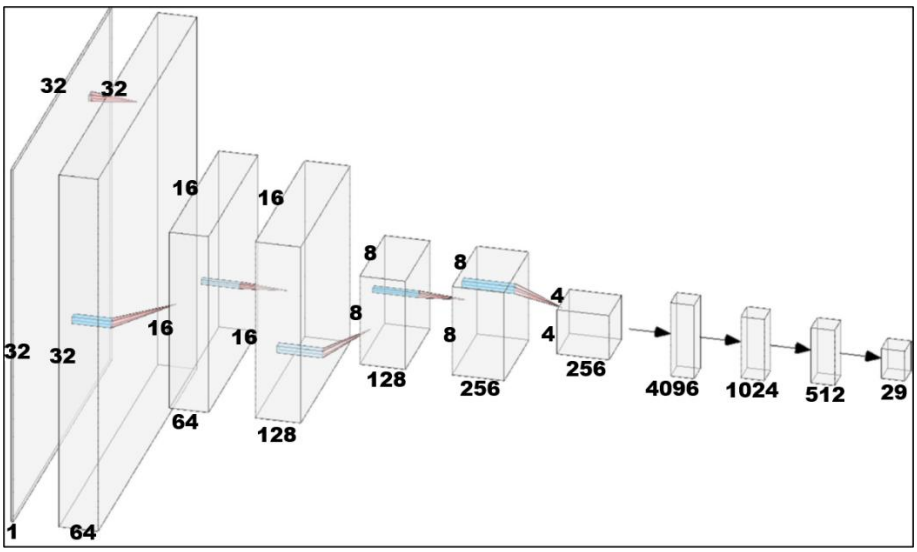

Figure 1. Deep Convolutional Neural Network (CNN) architecture for handwriting recognition (Altwaijry & Al-Turaiki, 2020)

1.3 Research Objectives

The objectives of this study were formed according to the problem statement, which are,

- To find and assess the risks of cybersecurity threats in ANN-based English handwriting recognition systems.
- To assess the ethical and privacy issues in handwriting or biometric data processing tasks like data collection, distribution, and storage.
- To recommend suitable strategies to overcome the issues in order to make handwriting recognition systems safe, while complying with standard regulatory frameworks.

1.4 Motivation

The world has seen that handwriting recognition has become a key technology in the automation of document handling, examination systems, and digital authentication due to the improved use of artificial intelligence in day-to-day operations. However, as the systems of this type were repeatedly used, which are built on ANNs and large volumes of personal handwriting, a new phenomenon of data security, the integrity of models, or user privacy has emerged. The driving force behind this is the realisation that, despite the high accuracy of ANN-based handwriting recognition, it has also exposed the user to the risks of data attacks, adversarial attacks, and abuse of biometric data. Even during minor compromise areas such as education, finances, and identity checking there is a probability of extreme ethical, legislative, and reputation repercussions. In addition to that, the existing literature has typically been focused on improving the accuracy of models, and not much attention has been paid to the mechanisms of cybersecurity resilience and privacy protection. The proposed study will fill that gap by performing a systematic review of the available vulnerabilities and commenting on privacy-focused design solutions and regulations. By so doing, it will contribute to the fact that there will be the emergence of trust and ethical handwriting recognition systems, which will not interfere with the rights of the individuals, yet will perform and can be applied in practice.

1.5 Review Methodology

The systematic literature review (SLR) process explored popular databases like MDPI, Springer Nature Link, ACM Digital Library, Arxiv, IEEE Xplore and Google Scholar to obtain ANN-related articles for English handwriting recognition that are searched through key strings like ‘ANN and English handwriting recognition’, ‘ethical AI and handwriting biometrics’ and ‘CNN handwriting and adversarial attacks’ as examples. The inclusion criteria are papers published within the last 10 years that are in peer-reviewed journals in English and conferences related to ANN-based handwriting recognition. This 10-year window is more relevant than the conventional 5–8-year traditional timeframe, due to the slow evolution of core English handwriting datasets (like MNIST, EMNIST,

and IAM), limited recent security-focused literature, and to better understand the threat evolution through longitudinal coverage. In the screening papers with non-ANN methods, non-English scripts and no focus on handwriting recognition are excluded. The screening process is performed in steps, like duplicate removal in initial research, title and abstract screening, and then screening by full text to keep papers that use different methods or datasets. Initially, 50 studies were collected, from which 40 remained after duplicate removal, through screening of the title and abstract, 15 papers were removed, and in the full-text review process, another 10 studies were excluded to comprise the final collection of 15 studies for the review. Dual reviewer screening was applied, and through consensus, disagreements were resolved to select a final number of 15 relevant studies for the review. The inter-rater reliability between two independent reviewers is $\kappa = 0.78$, indicating a substantial agreement between them during the screening process. Specific datasets like MNIST or EMNIST and IAM are prioritised as they have English alphanumeric characters and real curve English handwriting, respectively.

2. Artificial Neural Network-Based Handwriting Recognition Systems

2.1 Artificial Neural Network Architectures in Handwriting Recognition

ANNs have revolutionised the field of handwriting recognition due to the capability of the system to learn complex patterns with the aid of data, without necessarily having to implement manual feature engineering. The most widespread architectures that are employed are CNNs and RNNs since they complement each other and can process both visual and sequential data. CNNs are particularly effective at offline handwriting recognition in cases where an image of written text is used as input. The model acquires hierarchical features (i.e., edges, strokes, and letter shapes) automatically because of their stratified nature. Convolution and pooling operations can identify spatial relationships and provide a substantial level of noise, slants, and sizing sensitivity to handwriting, which CNNs can accomplish. The most popular CNN-based models, such as LeNet, VGGNet, and ResNet, have been shown to be great models in character and word classification.

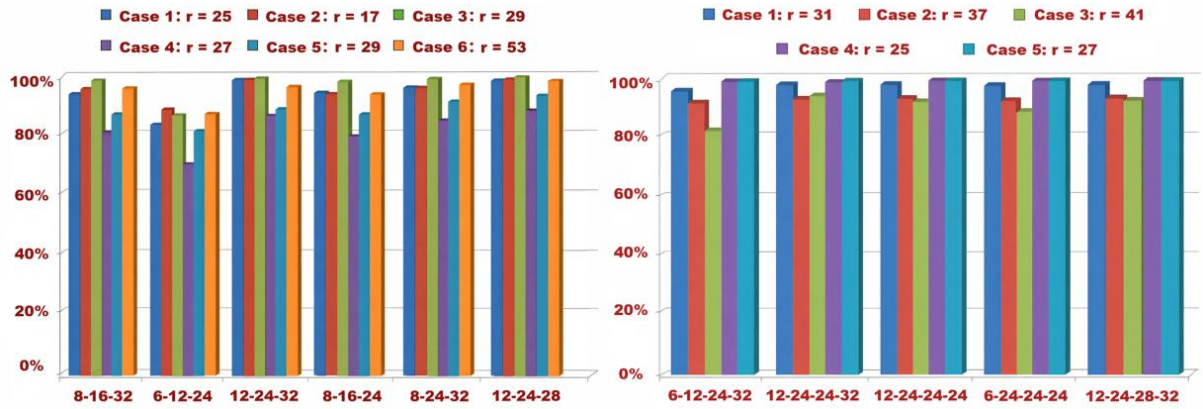

Figure 2. Percentage accuracies by Convolutional Neural Network (CNN) for different receptive fields
Note: 3-layer CNN (Left), 4-layer CNN (Right) (Ahlawat et al., 2020)

Ahlawat et al. (2020) experimented with different CNN architectures with necessary hyperparameter tuning to achieve a high recognition rate of about 99.89% when the Adam optimiser is used for training for the MNIST dataset. As shown in Figure 2, the recognition accuracy bars are slightly better with the 3-layer CNN model, irrespective of the receptive fields.

Nonetheless, RNNs, the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are more efficient in online handwriting recognition, where the platform is fed with time sequences with pen strokes. Since the RNNs can be trained on temporal relationships and context-based relationships between characters, they are also applicable in continuous text recognition. Hybrid systems based on CNNs to compute features and RNNs to model sequences, such as CNN-LSTM systems, have become typical in end-to-end handwriting recognition systems.

Sonkar & Kumar (2025) studied various CNN-RNN architectures for handwriting identification and presented a general process flow diagram for end-to-end handwriting recognition through the hybrid model, as given in Figure 3. They have found that hybrid models like the above can deliver over 90% accuracy in scene text detection, and the performance can be further improved by employing advanced techniques such as stochastic learning and multi-scale feature fusion. Meanwhile, attention mechanisms and networks of transformers have become more and more important in performance improvement in the recent past, as they allow models to focus on important spatial

or temporal locations in the input data. Not only are these architectures more interpretable, but they are also more recognition-capable. Such deep ANN structures have enabled major advancements in the capability to distinguish between different handwriting styles across people, hence creating a new benchmark of human recognition of the written text.

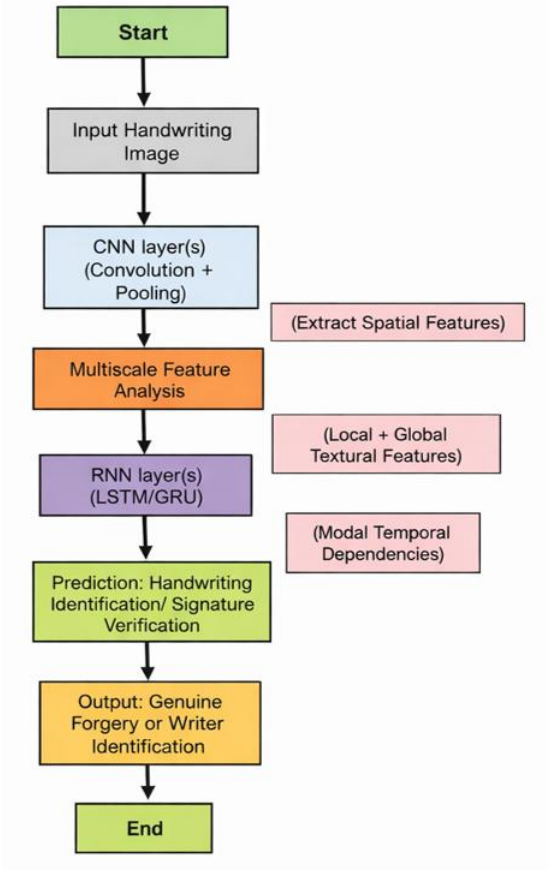

Figure 3. Process flowchart for Convolutional Neural Network-Recurrent Neural Network (CNN-RNN) hybrid framework for handwriting recognition (Sonkar & Kumar, 2025)

2.2 Performance Benchmarks in Handwriting Recognition

Performance benchmarking is a significant indicator of handwriting recognition, which quantifies the correctness of the models, generalisation, and practicality on a real-world basis. Over the past decade, deep learning models, namely CNNs, RNNs, and hybrid CNN-LSTM models, have improved in terms of accuracy to become recognised on standard datasets. The most used English handwriting benchmarks include the IAM Handwriting Database, EMNIST (Extended MNIST), and MNIST, which have provided a multitude of varied samples of written characters and words collected through the survey of thousands of individuals. The IAM data of the handwritten lines of the English text and samples of words have now become the standard way of testing the offline recognition systems. Connectionist Temporal Classification (CTC) loss models in models such as CNN-LSTM have achieved better word recognition rates that are improved over the earlier statistical and HMM-based methods. The self-distillation framework for CTC loss as proposed by Zhang et al. (2024) has achieved quite effective performance in text recognition for both English and Chinese benchmarks. Similarly, EMNIST can be viewed as an extension of the classic MNIST data, which includes not only numbers but also letters, and gives a wider range of benchmarks to recognise multi-class problems. Vaila et al. (2020) showed about 85% accuracy when biologically inspired methods, such as latency encoded spikes and STDP, are used with backpropagation on EMNIST classification. The results of the deep learning models trained on EMNIST often indicate very high rates of classification, and this proves that the current neural design performs well in acquiring the intricate character structures.

However, the benchmark performance tends to vary, according to the quality of the dataset, preprocessing approaches, and variants of handwriting. Controlled datasets are almost perfect, but they fail to use real-life tasks, like cheque processing or classroom testing, which introduce issues like uneven lighting, interspersions (cursive), and harsh backgrounds that reduce the accuracy of the model. Consequently, newer studies have been done on

cross-dataset validation and adversarial or distorted input robustness testing to establish the resilience of the actual model. In brief, the gradual increase in the benchmark scores shows that ANN-based handwriting recognition has reached maturity, and it addresses the importance of security-sensitive benchmarking. Besides the parameters of accuracy, the future assessments must include strong parameters of robustness and privacy, which will ensure the high reliability of implementation in real-life contexts concerning safety.

3. Cybersecurity Threats in Handwriting Recognition

3.1 Data Poisoning Attacks

The problem of data poisoning attacks is among the greatest cybersecurity risks to the handwriting recognition systems that are built on ANN. The attacks are made available when the malicious agents distort or corrupt the training set of data, which they use to generate the recognition model intentionally.

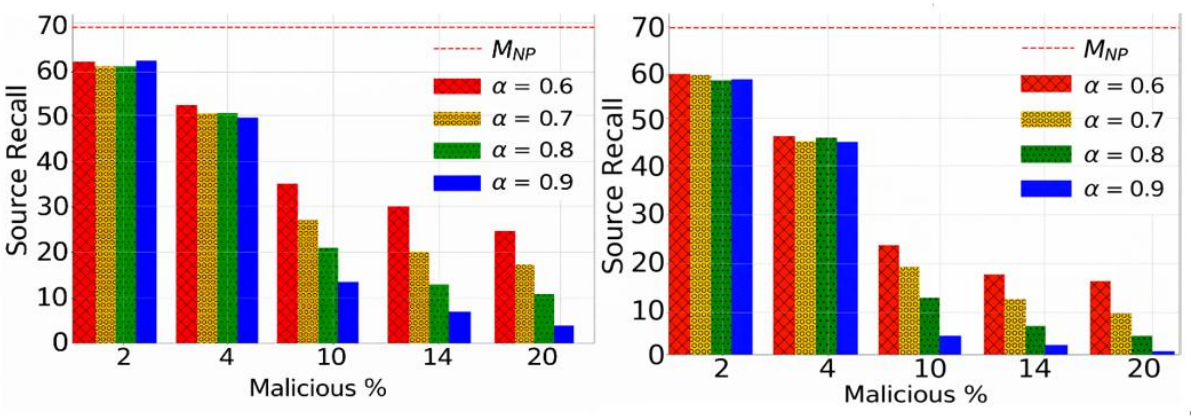

Figure 4. Availability of malicious participants on source class recall
Note: Left: CIFAR-10, Right: Fashion-MNIST (Tolpegin et al., 2020)

In Figure 4, even though CIFAR-10 and Fashion-MNIST are not English handwriting datasets, they are used here to present the general data poisoning dynamics, which are relevant for English handwriting recognition systems, as ANNs are vulnerable to training data manipulation. Tolpegin et al. (2020) investigated the effect of data poisoning on source class recall by studying two benchmark datasets, CIFAR-10 and Fashion-MNIST. The authors have found that federated learning systems are mainly vulnerable to label-flipping poison attacks, as Figure 4 clearly shows a reduction in source recall with the increase in poisoning effect α for both datasets. Taking into consideration that the deep learning algorithms depend heavily on the volumes of labelled samples of handwriting in the quest to learn patterns of characters and language structures, any manipulation of the small-scale data can send the performance and integrity of the model to ruin. The attackers in the typical data poisoning scenario include corrupt samples of handwriting or mislabelled samples of handwriting in the training set. An example of this would be they can modify some characters such that the model is similar to another label that is not accurate, such as the letter O written by hand can appear as Q. The model develops false associations after some time, resulting in systematic misclassification when it is implemented. More complex types of poisoning use perceptually inconsequential pixel-level distortion, which also generates visual images identical to the original handwriting. They can cause the unpredictability of the model, a decrease in accuracy, or even backdoors to be hidden.

The point is even more terrifying in the matter of biometric or authentication systems. An infected handwriting recognition software on a financial institution or identity check system can either wrongly identify the user, or give fraud the green light, or disclose confidential information. In addition, the data used in training is typically gathered through some sort of shared repository or crowdsourcing, and thus, data integrity is an ongoing issue. Suspicious samples could be identified by model training beforehand, as well as methods such as data provenance tracking, outlier detection, and noisy-tolerant training algorithms. In brief, data integrity against manipulation is considered a primary security measure to guarantee credibility, reliability, and ethical principles of the handwriting recognition systems in highly sensitive security zones.

In the case of English handwriting recognition systems, data poisoning can corrupt training samples that exploit cursive character ambiguity, which can lead to cheque-cleaning fraud and signature manipulation. The poisoned samples precisely target similar-looking characters such as O/Q or I/l, causing misclassification in a systematic manner that can be exploited by fraudsters to bypass automatic credentials verification checking.

3.2 Model Inversion Attacks

ANN-based handwriting recognition systems may face a severe threat of privacy loss due to model inversion attacks to access trained models, which may be used to recreate or infer sensitive details about the data with which they were trained. Unlike data poisoning that disrupts the model learning process, model inversion aims at removing the private handwriting features or input representations of the model.

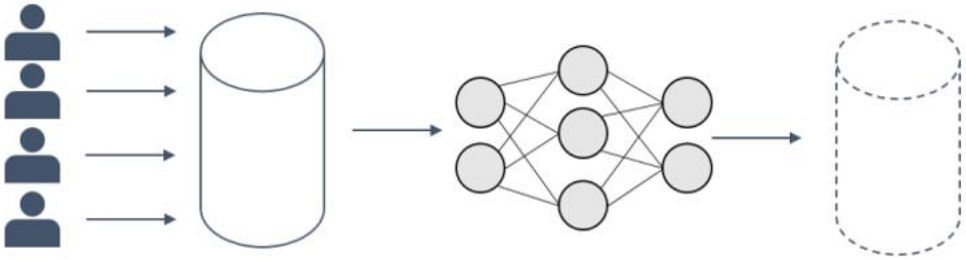

Figure 5. Model inversion attack framework (Song & Namiot, 2023)

The idea of model inversion attack is shown in Figure 5, which displays how trained classifiers are used to extract training data representation. In their study, Song & Namiot (2023) termed the model inversion attack a big threat to any machine learning applications, such as handwriting recognition, and especially warned against the ART attack, which is aimed at private data extraction. The attackers can use the output probabilities, gradient, or confidence score of the model to gradually replicate the original handwriting characters or reproduce the writing styles that belong to a specific user. In one of the typical situations, a challenger queries the recognition system on a few occasions and studies its responses to develop an impression of the most likely visual or structural patterns of the input data. To give an illustration, when a handwriting recognition model is trained on individual handwriting samples to identify them as either an authentication sample or a signature check, an attacker can reverse-engineer the personal handwriting samples to produce nearly identical handwriting images. This articulates a major concern about the confidentiality of the biometrics, as handwriting is a unique behavioural cue that can provide information about the individual or environmental details of people.

The effects are more than the invasion of privacy since model inversion may be employed to enable identity spoofing, intellectual property theft, and illegitimate surveillance. In an educational or financial system where identity is verified by writing, reconstructed samples can give the impression of impersonation or forged documents. Moreover, publicly available AI-based models that provide prediction APIs are particularly vulnerable since assailants can query them more than once without prior access to training data. To prevent such an attack, a number of privacy-saving tactics have been proposed. They include the principle of differential privacy, which introduces statistical noise on the outputs, federated learning, which makes sure that the information used remains decentralised, and access control systems, which limit the visibility of the model outputs. It can also prevent enemies from stealing sensitive data by adding encryption techniques and other safe implementations of the models. Vulnerability issues that have to be addressed to model inversion are therefore critical to maintain confidentiality and ethical handwriting recognition systems.

3.3 Adversarial Attacks

Adversarial attacks were also one of the threats to the cybersecurity of ANN-based handwriting recognition systems, based on the fact that the system is sensitive to input perturbations. These attacks mean introducing certain invisible noises or the distortion of images of handwriting and fooling the neural network to produce the wrong classification, yet the distorted handwriting is virtually identical to the human eye. Chakraborty et al. (2021) demonstrated that adversarial attacks need to be dealt with robust learning techniques to prevent malfunction of machine learning algorithms, especially for security-related applications, such as in handwriting recognition systems. In a nutshell, adversarial attacks show a fundamental weakness in deep learning models, which are founded on excessive reliance on high-dimensional patterns of features rather than real semantic interpretation of handwriting. Considering an example, when an intruder wants to confuse the model, he or she can just make minor changes to the pixel values on a picture of the printed word CAT to show the model that it is CAR or DOG, and that the human eye will not notice this. In applications with serious security requirements, such as bank cheques, checking exam papers, or digital signatures, such manipulations may lead to serious security and ethical consequences, including financial fraud or misjudgement. Adversarial attacks can be either white-box (where the attackers have knowledge of model parameters) or black-box (where model outputs are accessible), and therefore,

are highly adaptable and not easily identified.

3.4 Comparative Evaluation of Threats

The three types of threats are compared in Table 1 based on their technicality, impact severity, applicability of handwriting recognition, and mitigation techniques.

Table 1. Cybersecurity threats in Artificial Neural Network (ANN)-based English handwriting recognition

Threat Type	Technical Mechanism	Impact Severity	Applicability to English Handwriting Recognition	Representative Mitigation Techniques
Data Poisoning	Label flipping Backdoor injection	High	Exploits cursive ambiguity Cheque misclassification Signature misclassification	Data provenance checks Robust training Outlier detection
Model Inversion	Gradient-based Confidence-based	High	English signatures reconstruction Writer-specific stroke pattern generation	Differential privacy Output restriction Federated learning
Adversarial Attacks	Pixel-level Stroke-level	Medium	Case ambiguity slant deception Stroke-based deception	Adversarial training Defensive distillation Input normalisation

Table 1 summarises the comparison of the three stated cybersecurity threats, of which the impact severity is high for data poisoning and model inversion due to model corruption and direct privacy leakage issues, but for adversarial attack impact is medium, as with stroke-based or case ambiguity deception, only some targeted tasks can be achieved with high protection for confidential information. Being technologically different, the mitigation for the three threats is also different, with a focus on robust training and federated learning for data poisoning and model inversion threats, while for adversarial attacks, defensive distillation is one of the key solutions.

The variation in the reported effectiveness of attacks in previous studies can be mostly explained by the nature of the data sets and the architecture of the models. Attacks tested on controlled data like MNIST tend to claim greater success rates because of clean backgrounds and isolated characters, whilst EMNIST brings greater variety to classes, and IAM involves actual cursive English handwriting, which results in lesser attack transferability. Equally, CNN-based models that are trained on the fixed character images are more sensitive to pixel-level perturbations than the hybrid CNN-RNN models, which utilize sequence-based dependencies as well as context-based information. Hence, the effectiveness of poisoning and adversarial attacks on simple benchmark datasets does not imply their effectiveness for real-world English Handwriting data, as vulnerabilities depend on the dataset and architecture.

Another source of inconsistency between studies is due to the methodological shortcomings in the threat assessment. No standard benchmark exists at the moment to evaluate cybersecurity threats in the field of English handwriting recognition, which results in incoherent evaluation procedures and incomparable outcomes. Most of the investigations focus on classification accuracy as the main measure without considering privacy-related measures like biometric reconstruction fidelity or information leakage in case of model inversion. Also, adversarial attacks usually take the generic conditions of image recognition without considering English-specific properties like cursive connectivity or case transitions. These differences inhibit the ability of cross-studies to be interpreted and highlight the importance of having single security-conscious benchmarking models that are specific to English handwriting recognition.

4. Privacy and Ethical Challenges

4.1 Data Sensitivity Issues

Personal and sensitive data contains not only behavioural but also physiological peculiarities of a specific person. Compared to the generic texts or images information, the handwriting contains invisible biometric features such as pressure pattern, sequence of what is being typed, writing rhythm, and separation between letters, among others, which might be used as a positive type of identification. It is therefore the case that handwriting recognition systems are highly sensitive systems as far as privacy is concerned, particularly in an instance such as in training such systems using ANNs, where the process of training them is founded upon extensive data sets. The information contained in handwriting can be directly linked to the identity of an individual, and any type of infringement of this information may be a possible violation of personal identifiable information (PII).

This risk is augmented when this data is collected and processed without any informed consent or when it is collected without a high level of anonymisation. In other cases, the handwriting samples may have some form of contextual information, e.g., names, addresses, or any other indicator of writing style, which can be utilised to

complete profiling or make an identity inference. Besides that, the fact that the handwriting records are centrally stored in the cloud servers incurs the risks of unauthorised access, leakage, or intentional use. In case of compromise, the attackers can manage to reconstruct single handwriting styles, which can then be used to forge, impersonate, or misuse digital transactions. Morally, the collection of handwriting and analysis is expected to be done within the confines of transparency, fairness, and accountability. They should inform the users how their information will be stored, shared, and used in the model training. Hence, there is a need for good encryption standards, anonymisation protocols, and a limited access policy on data. The value of handwriting identification as a biometric identity cannot be underestimated as far as technical defence is concerned, yet it must also be viewed as morally responsible and morally upright towards the protection of human rights by applying AI-based handwriting recognition systems.

4.2 Personal Data Exposure Risks

The privacy concern that has been reported to be among the biggest in ANN-based handwriting recognition systems is the possibility of revealing personal information. Writing samples tend to be tapped and include personal data of actual people, i.e., names, addresses, or signatures. The sample can be stolen, copied, or misused by hackers or other unauthorised users if such samples are stored or shared negligently. Being able to forge signatures or identity theft, or even financial fraud, using the leaked samples is possible due to the uniqueness of handwriting to each individual. Many models of handwriting recognition are developed based on the huge datasets that are stored in the cloud or shared by researchers and organisations. The attackers have access to these datasets through weak passwords, unencrypted files, or system vulnerabilities unless these are well secured. Even more complex machine learning algorithms can, in specific circumstances, re-identify a person by matching their handwriting to the ones present in the known set. This makes anonymisation or elimination of names meaningless in ensuring actual privacy.

It is even more disastrous in cases where handwriting is used in authentication measures such as online tests, e-signatures, or authentication of access. The information may reveal hidden information about user behaviour in their writing, in instances of leakage of model outputs or embedded storage. Not only is this a violation of their privacy, but it also destroys the confidence of AI systems. Some of the data protection measures that the developers and organisations should use to reduce the probability of data exposure include data encryption, cloud storage, and access control. The sensitive handwriting information can also be saved through restricting the sharing of data, monitoring the systems, and applying the techniques of differential privacy. Most importantly, the users should be adequately informed about how their information on handwriting will be gathered, stored, and utilised. With the assistance of good data protection habits, the handwriting recognition systems can be maintained in a secure, reliable, and privacy-invasive condition.

4.3 Ethical Implications

An identification of handwriting by using ANNs is linked to several ethical issues, including privacy, fairness, consent, and accountability. Since handwriting is a form of individual expression and it might work as a biometric identifier, its use without the consent and safety required can lead to ethical violations. Much handwriting data is stored on students, employees, or web users who might not fully understand the manner in which their data will be used. When the consent is not clear or free, then this is against the universal ethical standard of autonomy and informed consent. The other ethical problem is data ownership and control. Once the handwriting samples are collected and trained on the AI models, an individual is, as a rule, denied the right to choose what to do with their information, to filter it, reuse it, or sell it to a third party. This gives a disproportionate power distance between the people considered in the data sets and the organisations behind the development and implementation of handwriting recognition systems. Besides, they may also be trained on biases accidentally owing to input data on such systems. In the case of other samples, which are from different age groups or cultures, the system might not be very productive, as the results can be unfair and discriminatory.

Regarding ethics, it is the responsibility of organisations that develop such systems to ensure that AI decisions are clear, impartial, and comprehensible. It must enable users to understand the utilisation of their handwriting and must enable them to access their data in some form of withdrawal. One needs to adhere to the existing ethical consequences and codes of conduct, such as those designed by the European Commission and its Ethics Guidelines in AI or the national ethics boards in research. Concisely, handwriting recognition should be morally designed to ensure that the balance between innovation and responsibility is restored. The systems are supposed to be created to preserve the rights of individuals, ensure fairness, and promote trust so that AI technologies can be applied in such a way that they do not undermine society and sacrifice human dignity and privacy.

5. Regulatory and Compliance Considerations

5.1 Data Protection Frameworks

The data protection systems are relevant in the protection of the information of handwriting when used in AI and handwriting recognition systems. Those models define both legal and ethical standards that the process of gathering, retaining, and transferring personal and biometric data needs to adhere to.

Table 2. Security incidents and compliance rate for different governance framework models (Julakanti et al., 2025)

Model	Security Incidents (Per Year)	Regulatory Compliance Rate (%)
Centralized	3	95
Decentralized	7	88
Hybrid	4	92

A case study conducted by Julakanti et al. (2025) in Table 2 shows that among three types of governance frameworks for data protection, most security incidents happened in decentralised frameworks, and their compliance rate is also lowest, while for centralised frameworks, the compliance rate is highest and its incidents are also lowest, indicating centralised frameworks are most secure.

One of the topmost priority laws in the European Union was the General Data Protection Regulation (GDPR), according to which handwriting and other biometric data are considered personal sensitive data. Under GDPR, organisations must seek direct permission when gathering the handwriting information of users, and they must explain how they would use such information. Other legislations, such as the California Consumer Privacy Act (CCPA) in America, are related in this way and safeguard user rights by providing them with the opportunity to access, update, or delete information stored about them. Such models force businesses to use effective security measures like encryption and restricted access in an attempt to prevent unauthorised usage. The compliance with these rules is not an obligatory law, and it also leads to the establishment of trust and openness between organisations and users. Under ANN-based handwriting recognition systems, compliance with these structures would assist in a responsible and ethical way of running the personal handwriting information to reduce the chances of abuse, identity theft, or data leakage.

Table 3. Regulatory compliance comparison for English handwriting recognition systems

Regulation	Classification of Handwriting Data	Consent Requirements	Data Subject Rights	Penalties for Non-Compliance	Practical Implications for Developers
General Data Protection Regulation (GDPR) (EU)	Biometric Sensitive personal data	Explicit Informed consent	Access Erasure Portability	Penalty on Global turnover	Privacy-by-design Encryption Audit logs
General Data Protection Regulation (CCPA) (USA)	Personal data Biometric identifiers	Opt-out Notice-based	Access Deletion	Civil penalties per violation	Data minimisation User deletion workflows
Digital Personal Data Protection (DPDP) Act (India)	Personal data Sensitive	Explicit consent	Access Correction Erasure	Financial penalties	Secure storage Consent management
Office of the Australian Information Commissioner (OAIC) Guidelines (Australia)	Sensitive Personal information	Informed consent	Access Correction	Regulatory enforcement actions	Transparency Controlled data sharing

Table 3 compares different regulatory compliance frameworks across different regions, such as the EU, the USA, India, and Australia. Among these, the penalty for non-compliance with consent requirements and data subject rights is highest in the EU, as a direct penalty on the organisation is imposed as a percentage of global turnover, whereas for other countries, indirect penalties are used as civil or financial penalties and regulatory enforcement actions.

5.2 Data Controller Obligations

A data controller is the person or organisation that decides on the mode and purpose of collecting and using personal information, such as handwriting. Data controllers should be able to unambiguously safeguard the interests of individuals and ensure that their personal data is safe and processed correctly, in addition to being fair

according to the privacy laws, including the GDPR. Finck (2021) emphasised that when data controllers lack physical control over personal data that the software utilises for performing any business task, then the risk to the rights of data subjects is increased. Thus, they need to be taught to seek informed consent from the users prior to accessing handwriting data and explain the purpose in a straightforward manner. Other people should not know the information unless it is for the purpose given. Data controllers are expected to make sure that the data available in handwriting is accurate, secure, and up to date as well. This involves the encryption system, authorised control to authorised staff, and regular analysis of security threats. They ought to notify the authorities and people who are exposed in case of any data breach. In addition, a user can gain access, make changes, or delete his or her data. These rights should be exercised by the data controllers. In compliance with these requirements, data controllers ensure that ANN-based systems that can be deployed in identifying handwriting address privacy, provide transparency, and comply with international data protection rules.

5.3 Regulatory Implications

There exist regulatory liabilities and consequences of using handwriting data in AI-based systems. It is possible to consider handwriting as a type of biometric information, and, in that case, it is liable to the standards of the high laws of privacy, such as the GDPR and other country-specific acts. Sirur et al. (2018) found that GDPR compliance is mostly accepted by large companies within the given deadlines for its simple regulatory norms. These laws govern organisations to comply with the laws, as well as fairness and transparency in the development and use of handwriting recognition models. A failure by an organisation to ensure that its information is secured or abused may lead to an organisation paying huge fines, prosecution, and its image will be tarnished. The third alternative is that regulators may apply restrictions with the help of AI systems, which do not meet privacy or security demands. Considering that a handwriting recognition system would be disrespectful to the privacy laws, provided that it retains individual data without the knowledge of the user, or that it is accessible to a third party. These rules also enforce privacy-by-design solutions, i.e., security must not be added onto the system later on. Overall, the regulatory implication indicates the relevance of compliance with the data protection policies and ethical application of AI, since it is necessary to consider the fact that handwriting recognition technologies should be used in a safe, fair, and ethical manner that cannot conflict with any of the global privacy standards.

5.4 Compliance Challenges

ANN-based handwriting recognition systems may prove to be a difficult undertaking in terms of compliance since this type of technology is typically utilised based on extensive and highly complex datasets that are accumulated by many different sources. The most important thing is to obtain a written consent of the entire population whose handwriting data is accessed, especially when the information is received on shared or public databases. It is also not easy to anonymise handwriting samples such that removal of identification occurs even in the absence of names or labels, as writing pattern is still used to show individual identity in several organisations. The other concern is connected to the adherence to different global privacy regulations, such as GDPR in Europe or CCPA in the United States, that have other demands in terms of data processing and user privacy. Huang (2025) highlighted the main difference in needs of the two regulation frameworks as GDPR demands accountability, transparency, and individual autonomy, while CCPA prioritises more on the choice of consumers and commercial viability. The threat of accidental leakage or unauthorised use of the data also increases during the maintenance of data safety during cloud storage, model training, and data sharing across institutions.

Secondly, organisations need to update their AI policy and systems on a regular basis to conform to the new regulations and ethical standards of AI use. Complete compliance may be harder because of the lack of awareness or technical knowledge. Therefore, the solutions to those problems are to invest in data management systems, staff training, and privacy-sensitive technologies that will ensure that handwriting recognition systems are safe, transparent, and legally accountable.

6. Mitigation and Defence Mechanisms

Different ANN architecture presents different vulnerabilities, which lead to common privacy and ethical risks, and thus, mitigation layers are important when models are used for handwriting recognition to ensure relevant regulatory constraints are followed.

Figure 6 shows that architectures like CNNs and RNNs have common vulnerabilities like adversarial attacks, overfitting, temporal manipulation, and complex hybrid architectures mostly have explainability issues. These lead to privacy and ethical risks like data breaches, bias, and discrimination for all models, while a lack of transparency is mostly faced for hybrid models, as they are difficult to understand. Hence, models are recommended to pass through common mitigation layers at data-level, model-level, and deployment-level in order to ensure the inference follows necessary regulatory constraints for English handwriting recognition in the respective region.

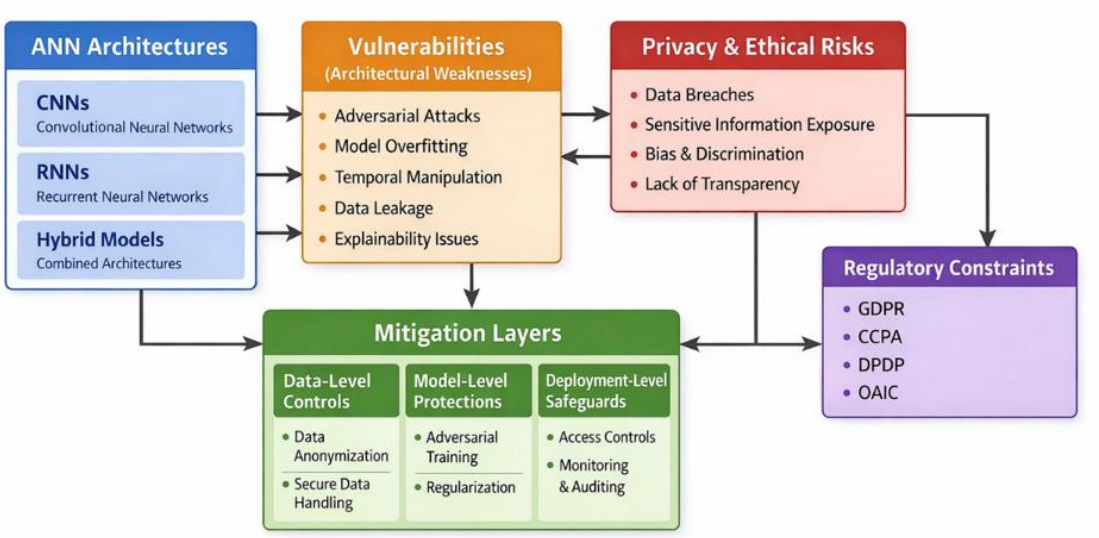

Figure 6. Concept map of vulnerabilities, privacy risks, regulatory constraints, and mitigation layers for different Artificial Neural Network (ANN) architectures in English handwriting recognition

6.1 Data Security Measures

ANN-based recognition systems are very sensitive, and hence the handwriting data requires a strong data security policy that would help to deter theft, misuse, or tampering of this type of data. The information in handwriting can contain personal and biometric information, and after this, it is important to take care of it in its lifecycle, which includes the period of its capture, storage, and processing of the data. One of the measures is encryption of the data, which states that files cannot be read even in case of theft of the files, as they will require the corresponding decryption key. Ananya et al. (2023) mentioned the advantages of the AES data encryption algorithm and its evolution over time from various studies to ensure data security. Besides, unauthorised access can be prevented by using safe data storage systems (guaranteed servers and limited access to the cloud). The other important measure is access control, whereby the data can only be accessed and updated by the authorised individuals, such as the system administrators or researchers. Cyberattacks are also prevented by routine security audits, firewalls, and an Intrusion detection system. Organisations should also take advantage of data anonymisation and tokenisation to eliminate or hide personal identifiers in a sample of handwriting before model training. Finally, information backups and disaster recovery strategies are also commendable steps towards having the information returned safely in the event of a system breakdown or intrusion. Taken together, these measures will allow organisations to reduce the risk of a data breach by a large margin and ensure the privacy and integrity of handwriting information wherever it is used in AI recognition.

6.2 Model Robustness Techniques

The ANN-based handwriting recognition systems are made more robust to cyberattacks, and errors are the objective of the methods of model robustness. One of them is adversarial training, where one party is trained to recognise normal handwriting, and the other party is trained to recognise deliberately corrupted handwriting, in the hope that the model will learn to respond adversely to attacks. Bai et al. (2021) pointed out certain limitations of adversarial training, such as min-max optimisation and overfitting. To prevent overfitting, which allows the model to act consistently on the unknown data, the regularisation methods of dropout and weight decay are applied. The other method is defensive distillation, as the edges of the choices of the model are diffused to make it less sensitive to the slightest input variation that occurs in the adversarial attacks. Papernot & McDaniel (2016) performed different experiments with defensive distillation and found that it can mitigate adversarial samples by using the fast gradient descent sign method in combination with the Jacobian-based iterative approach. The risk of fooling a model may also be countered by using such methods as ensemble learning, in which the models are employed in order to reach a group decision. In addition, before recognition, input preprocessing methods may eliminate malicious distortions such as noise filtering or image normalisation. The integration of these ways will make the model stronger and more accurate in recognition systems of handwriting in the natural environment.

Particularly for English handwriting recognition, models must be robust enough to deal with cursive connectivity, ambiguity between lower and uppercase letters, and stroke variations for various types of slants. Adversarial models need to be trained with examples of character joins, case transitions, and variations due to

different slant types to reduce the overall classification error across a diverse range of English handwriting styles.

6.3 Secure Model Deployment

Secure model deployment is a significant procedure of making sure that the handwriting recognition systems designed based on ANNs are not vulnerable to unauthorised access, tampering, and misuse of the system once they are introduced to practice. Its models are prone to a vulnerability in the process of deployment, as they are likely to be linked with the external networks, the cloud services, or the devices that people have. To curb these risks, organisations are recommended to deploy in safe environments, which have encrypted communication systems such as HTTPS or VPNs, in an attempt to discourage the interception of data.

Access control and authentication are another important step for deploying the model in a secure manner. The trusted interfaces or applications are only supposed to access the deployed model through the secured APIs or limited interfaces. Its fluctuating functionality or even its invasion during the initial stages can be detected with the help of regular security patches and model monitoring. The protection of intellectual property other than that can also be performed by using model watermarking or digital signatures; it may be employed to authenticate ownership of models and mark illegal copies. The other measure that ensures minimal impact in case of breaches is the fact that the model is not interconnected with the other systems. Finally, accountability should be present with audit logs, on which all the access and changes will be recorded. By so doing, the developers would be in a position to ensure that the handwriting recognition model is safe, in line with privacy, and resistant to cyber-attacks throughout its life cycle.

7. Conclusion

7.1 Summary

This specific study has managed to come up with a review on cybersecurity, privacy, and ethical issues of ANN-based English handwriting recognition systems. The attacks, such as data poisoning, model inversion, and adversarial attacks, were used to explicate the integrity and privacy of data problems. The paper has also examined the legal criteria like GDPR and CCPA and has found that a company must be accountable, transparent, and ethical. It lastly suggested mitigation measures, such as encryption, use of good models, and safe deployment, to enhance the reliability and trustworthiness of the systems in the educational, financial, and identity verification systems.

7.2 Limitations

The study is more of a theoretical survey based on the available literature, and there is no experimental testing or performance testing of the mitigation measures. It concentrated upon the English handwriting orders and was incapable of generalisation across other languages and writings. Furthermore, the fact that the actual implementation data was not taken into account because of the privacy issues was accepted in the study. The future requirements demand empirical studies on the scrutinised defence strategies that utilise bigger and more diversified databases and cross-platform security studies to increase reliability in a real-life scenario.

Moreover, the results of this review can be affected by the bias of data sets since most of the studies referenced are based on commonly used benchmarks like MNIST, EMNIST, and IAM, which might not be a true reflection of English handwriting variation in the real world. The study is also constrained to the fact that it uses secondary literature to carry out the analysis, as opposed to empirical experimentation to directly prove the effectiveness of security and privacy mitigation measures in practical deployment environments.

7.3 Future Directions

Priority should be made in further research involving privacy-sensitive learning algorithms (federated encryption and homomorphic encryption) should be done in the future to reduce the chances of direct data exposure. Preventing privacy leakage from an ANN-based handwriting recognition system should be the utmost priority here due to frequent model inversion attacks, which present an open challenge for future work. Here, Explainable AI (XAI) may contribute to the increase in transparency and confidence in decision-making. Moreover, the security-oriented benchmarking model and AI governance schemes will be able to facilitate safer usage in the critical areas. The ANN-based handwriting recognition would require inter-disciplinary cooperation among the developers of AI, ethics, and policymakers to develop balanced systems that will integrate performance, equity, and adherence.

Author Contributions

Conceptualization, J.Y.Z. and M.J.L.; methodology, J.Y.Z.; validation, J.Y.Z. and M.J.L.; formal analysis,

J.Y.Z.; investigation, J.Y.Z.; resources, J.Y.Z.; writing—original draft preparation, J.Y.Z.; writing—review and editing, J.Y.Z. and M.J.L.; visualization, J.Y.Z.; supervision, M.J.L.; project administration, J.Y.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

Ahlawat, S., Choudhary, A., Nayyar, A., Singh, S., & Yoon, B. (2020). Improved handwritten digit recognition using convolutional neural networks (CNN). *Sensors*, *20*(12), 3344. https://doi.org/10.3390/s20123344.

Altwaijry, N. & Al-Turaiki, I. (2020). Arabic handwriting recognition system using convolutional neural network. *Neural Comput. Appl.*, *33*(7), 2249–2261. https://doi.org/10.1007/s00521-020-05070-8.

Ananya, B. L., Nikhitha, V., Arjun, S., & Gowda, N. C. (2023). Survey of applications, advantages, and comparisons of AES encryption algorithm with other standards. *Int. J. Comput. Learn. Intell.*, *2*(2), 87–98. https://doi.org/10.5281/ZENODO.7921019.

Bai, T., Luo, J., Zhao, J., Wen, B., & Wang, Q. (2021). Recent advances in adversarial training for adversarial robustness. *arXiv Preprint*, arXiv:2102.01356. https://doi.org/10.48550/ARXIV.2102.01356.

Chakraborty, A., Alam, M., Dey, V., Chattopadhyay, A., & Mukhopadhyay, D. (2021). A survey on adversarial attacks and defences. *CAAI Trans. Intell. Technol.*, *6*(1), 25–45. https://doi.org/10.1049/cit2.12028.

Finck, M. (2021). Cobwebs of control: The two imaginations of the data controller in EU law. *Int. Data Priv. Law.*, *11*(4), 333–347. https://doi.org/10.1093/idpl/ipab017.

Huang, M. (2025). Digital privacy in the age of surveillance: A comparative study of GDPR and CCPA. *OTS Can. J.*, *4*(7), 65–74. https://doi.org/10.58840/1t99rb13.

Julakanti, S. R., KiranmayeeSattiraju, N. S., & Julakanti, R. (2025). Data protection through governance frameworks. *arXiv Preprint*, arXiv:2502.10404. https://doi.org/10.48550/ARXIV.2502.10404.

Papernot, N. & McDaniel, P. D. (2016). On the effectiveness of defensive distillation. *arXiv Preprint*, arXiv:1607.05113. https://doi.org/10.48550/ARXIV.1607.05113.

Sirur, S., Nurse, J. R., & Webb, H. (2018). Are we there yet? Understanding the challenges faced in complying with the General Data Protection Regulation (GDPR). In *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security* (pp. 88–95). https://doi.org/10.1145/3267357.3267368.

Song, J. & Namiot, D. (2023). A survey of the implementations of model inversion attacks. In *Communications in Computer and Information Science* (Vol. 1748, pp. 3–16). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-30648-8_1.

Sonkar, C. K. & Kumar, V. (2025). Review of artificial intelligence methods in handwriting identification using CNN-RNN for textural features. In *2025 International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI)* (pp. 648–654). https://doi.org/10.1109/ic3ecsbhi63591.2025.10990766.

Tolpegin, V., Truex, S., Gursoy, M. E., & Liu, L. (2020). Data poisoning attacks against federated learning systems. In *Lecture Notes in Computer Science* (pp. 480–501). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-58951-6_24.

Vaila, R., Chiasson, J., & Saxena, V. (2020). A deep unsupervised feature learning spiking neural network with binarized classification layers for the EMNIST classification. *IEEE Trans. Emerg. Top. Comput. Intell.*, *6*(1), 124–135. https://doi.org/10.1109/tetci.2020.3035164.

Zhang, Z., Lu, N., Liao, M., Huang, Y., Li, C., Wang, M., & Peng, W. (2024). Self-distillation regularized connectionist temporal classification loss for text recognition: A simple yet effective approach. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, Issue 7, pp. 7441–7449). https://doi.org/10.1609/aaai.v38i7.28575.