Mesopotamian journal of Cybersecurity
Vol.4,No.2, **pp**. 129–142
DOI: https://doi.org/10.58496/MJCS/2024/0012; ISSN: 2958-6542
https://mesopotamian.press/journals/index.php/cybersecurity

Research Article

# Artificial Intelligence and Cybersecurity in Face Sale Contracts: Legal Issues and Frameworks

Lobna Abdalhusen Easa Al-saeedi[1], Doaa Fadhil Gatea Albo mohammed [2], Firas Jamal Shakir[2], Faris Kamil Hasan [2], Ghadeer Ghazi Shayea [3], Yahya Layth Khaleel [4, *], Mustafa Abdulfattah Habeeb [4]

[1] *College of Law, Universe of Baghdad, Iraq.*
[2] *College of Law, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq.*
[3] *Technical College, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq.*
[4] *Department of Computer Science, College of Computer Science and Mathematics, Tikrit University, Iraq.*

**ARTICLE INFO**

Article History
Received 05 Jun 2024
Accepted 08 Aug 2024
Published 30 Aug 2024

Keywords
Face Contracts
Artificial Intelligence
Cybersecurity
Deep Learning
Legal Obligations
Data Protection
Privacy Violations

**ABSTRACT**

The sale of facial features is a new modern contractual development that resulted from the fast transformations in technology, leading to legal, and ethical obligations. As the need rises for human faces to be used in robots, especially in relation to industries that necessitate direct human interaction, like hospitality and retail, the potential of Artificial Intelligence (AI) generated hyper realistic facial images poses legal and cybersecurity challenges. This paper examines the legal terrain that has developed in the sale of real and AI generated human facial features, and specifically the risks of identity fraud, data misuse and privacy violations. Deep learning (DL) algorithms are analyzed for their ability to detect AI generated faces in order to potentially function as an AI safety in face sale agreement to allow the authenticity and protecting data. In addition, it examines the legal mechanisms surrounding consent, liability and data protection and suggests changes to help accommodate the complexity of AI. This paper proposes a framework by which AI tools can be integrated into the evolution of cybersecurity strategies, to mitigate risks and ensure compliance with such new legal standards and contribute to discussing the ethical and secure use of AI in Face sale contracts.

## 1. INTRODUCTION

Artificial intelligence (AI) has seen a rapid evolution and has now brought major changes to the way cybersecurity, law and digital identity works. The generation of hyper realisitic facial images via AI technology such as deep learning (DL) is one of most profound developments. In areas such as identity verification, personal data protection and cybersecurity, these AI generated faces are indistinguishable from real human faces and hugely problematic [1]. The Face sale contracts, in which both real and AI generated facial data are bought, sold, and licensed, are rising, raising novel legal and ethical questions about who owns rights to the faces of those represented in online and augmented realities. Because media, entertainment, and marketing have seen an exponential increase in demand for digital representations, these contracts have been adopted. Yet these risks come with the twin risk of identity fraud, data misuse, privacy violations [2].

In the context of face sale contracts, the risks are twofold: the potential for this data to be misused, coupled with the fact that it is so difficult to distinguish real and AI generated facial data. Facial synthesis using the aid of AI systems, more especially those based on DL, has become increasingly realistic, thus making the identification difficult whether a face is real or not. All this represents a new frontier for identity fraud, where AI generated faces might be used to commit such offences as impersonation, use of someone's likeness without their permission or for use in illegal transactions [3].

*Corresponding author. Email: yahya@tu.edu.iq

This study uses three models the: ResNet50, efficient Net, and a Convolutional Neural Network (CNN) to detect and differentiate the real and AI synthesized facial images. These models establish a technical basis on which the face sale contract cybersecurity measures can be improved to prevent fraud and to verify the authenticity of the facial data that they contain.

As AI entails using faces in some times, the legal systems around the world struggle to deal with the complexity introduced by their use [4]. Current laws regarding privacy, especially as they pertain to data protection and contracting, are often ill adapted to the problems posed by this cutting edge technology. Particularly pressing are issues of consent, liability and ownership of facial data [5]. For instance, what happens when a real person's likeness is synthe-sized by AI without their conset and then sold or used in a contract? When AI generated faces are used in fraud, or other malicious purposes, who is responsible? These questions raise the need for a full body legal policy that covers the sales of facial data as well as the protection of those parties involved against any misuse or identity theft.

This paper focuses on the legal and cybersecurity issues arising in face sale contracts, especially circumstances under which real faced data and AI data could be distinguished. This paper proposes a framework that integrates current AI detection technology into face sale agreements that protects parties from fraud and unauthorized use by exploring current AI detection technology. The face sale contract validations can use AI based detection mechanisms specifically Deep Learning (DL) based mechanism to be strong deterrents in the authentication process [6].

Additionally in this paper, the law governing the face sale contract will be examined by also looking at factors such as consent, liability, and data protection laws. In addition, it will explore how the complexities introduced by AI can be accommodated by these frameworks through modification and even an extension to yield new legal provisions that would better secure these contracts. This study is a contribution to the growing ethics and security of use discourse for AI in the context where people's identity and personal data at the risk.

**This study seeks to address several key questions, including:**

- How can AI models used to detect and classify facial data as being real or made up of AI?
- When selling and using AI generated faces what kind of legal risks do they pose with, for example, identity fraud, privacy violation and misuse? What is the role played by consent and liability in rights sale contracts, and what solutions to these problems can existing legal frameworks provide?
- What can be done with existing data protection and contract laws to accommodate the problems arising from sales of AI generated faces in face sale contracts?
- What ethical implications exist in the manufacturing, selling, and use of AI generated facial data? Here is how can AI based detection mechanisms be incorporated into cyber security strategies to safeguard face sale contract and to stop fraudulent activities.
- What might the consequences be of not acting to moderate the utilization of AI created facial data in a timely and appropriate manner? applied to detect and distinguish between real and AI-generated facial data?
- What legal risks arise from the sale and use of AI-generated faces, particularly in terms of identity fraud, privacy violations, and unauthorized use?
- What role do consent and liability play in face sale contracts, and how can these issues be addressed within the current legal frameworks?
- How can existing data protection and contract laws be adapted or extended to account for the challenges posed by AI-generated faces in face sale contracts?

The study seeks to develop a complete framework answering these questions, examining the technical and legal complexities of face sale contracts. This paper aims to propose new regulatory strategies based on the integration of AI detection technologies against current legal frameworks so as to have not only security but also ethical use of AI generated facial data.

## 2. RELATED WORKS

Over the past few years the progress that has been made in artificial intelligence (AI) [7] and Deep Learning (DL) [8] techniques has allowed the creation of extremely realistic synthetic images, that present significant security and ethic and data authenticity issues.y. These challenges have been studied in several instances, and the methods of distinguishing between AI generated and real content have been proposed.

One study [9] takes the idea of AI created images that are similar but not identical to real photographs and runs with it, both creating a synthetic dataset that looks like CIFAR-10 but using latent diffusion. The classification of real and AI generated images was framed as a binary problem and a CNN was used to achieve 92.98 accuracy through extensive hyperparameter tuning. The study used explainable AI techniques like Gradient Class Activation Mapping (Grad-CAM)

and revealed that subtle imperfections in the background, rather than central image features, were the essential element of classifying images. It also contributed to future AI detection research through the making of available the CIFAKE dataset. Indeed, another study [10] explored the security and moral ramifications of AI generated written content, as part of their classifying AI generated versus actual images. This study extracted features using pre trained CNN models like SqueezeNet, InceptionV3 and VGG19 with a dataset of 975 images. Using ML models, such as Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN) and Support Vector Machines (SVM), we classified these features. The InceptionV3+ANN model achieved highest classification accuracy. But the study indicated that much larger datasets would be required to boost classification performance, as the size of the the dataset was small.

Another research [11] used DL techniques like ResNet50 model to detect real faces and fake faces. The researchers trained a 9,000 image dataset for 150 epochs and got over 100% training accuracy, 99.18% validation accuracy and 99% testing accuracy. This work highlights the power of DL models like ResNet50 in recognizing fake AI generated faces and prove the tremendous potential of neural networks used in face detection projects.

As these works highlight the growing necessity for reliable methods to recognize AI-generated content, the ethical and security challenges it brings a debate on the importance of introducing robust methods for recognizing AI generated content arises. Moreover, they also show that DL models like CNNs and neural networks can effectively solve the problem of real vs. AI generated image classification, which forms a basis for future work exploring ways to protect digital identities and data authenticity.

## 3. SALE CONTRACTS AND LEGAL ISSUES

A sale contract It is a legally binding agreement that makes a deal of exchange between goods or services between a buyer and a seller [12]. These contracts are important to the general clarity, lack of unfairness and enforceability of the contracted upon obligations. There are a number of fundamental elements which are necessary for a valid sale contract, offer and acceptance, a consideration legal capacity a lawful purpose and an agreement between the parties [13]. Offer and acceptance constitute the mutual agreement between the parties, consideration is the exchange of value ordinarily expressed in the form of a particular consideration. To enter into the agreement both parties have legal capacity, that is, they need to have sound mind and meet the legal age requirement. Additionally, the contract has to have a bona fide purpose, parties must act with consent and freely devoid of coercion or misrepresentation [14].

A lot of legal issues arise from the execution or interpretation of sale contracts. Breach of contract is a common problem wherein one party fails to discharge his obligations [15]. Damages, specific performance and rescission of the contract are among remedies for breach of contract. Working challenges are frequent misrepresentation and fraud as false statements made in negotiations can result in the nullification of a contract, or a claim for compensation from misrepresentation. Complications arise when there are disputes on title and ownership and determining when title and ownership transfer leads to liability claims. In addition, warranties, whether expressed or implied, are commonly included in sale contracts, which warranties the product quality in case the failing of which are disputed. A second point of law is related to force majeure clauses that are intended to protect against unforeseen events, such as natural disasters, which would prevent one or both parties performing under the contract.

Throughout the history, the regulations of sale contracts are determined by different fundamental rules and consumer protection rules, which aimed to encourage fair practices and defending the buy’s interests. Across borders, there are instruments, such as the United Nations Convention on Contracts for the International Sale of Goods (CISG) [16], for use in facilitating cross border transactions. Many jurisdictions adopt laws domestically, such as the Sale of Goods Act, legislating valid rights and obligations of buyers and sellers in respect of goods under terms of quality, fitness and conformity. In short, these laws help keep the trust in the commercial transactions. Consumer protection laws also have the job of keeping a check on unfair practices and evaluating to action against sellers when there is a defect or misrepresentation.

Sale contracts are drafted in many instances and they contain some resolution mechanisms in situations of dispute. It is common to do the negotiation and mediation that brings the parties to agree a resolution not through formal legal process. However, an arbitration affords the participants a better structured but more private option, as the arbitrator's decision is binding on the users.

The emergence of e-commerce and electronically transacted business has brought new legal, including about online contract law, electronic signatures and jurisdiction. With the growth of the digital era, courts and lawmakers are becoming more focussed to update and modify the existing legal frameworks to the problems on privacy, cybersecurity and consumer protection. These developments underscore the fact that sale contracts are developing, and contract drafting should result in documents that are flexible and in the recognition of the changeable commercial environment.

Over the last few years, a novel dimension of sale contract emerged with the sales of rights to use human faces in robotics and artificial intelligence. This issue, powered by the growth of AI, is in the medium of licensing for the likeness of face

for the humanoid robot, virtual assistant and avatars of AI. The sale of facial rights is a tricky exercise in legal and ethical considerations, legal rights and public privacy. The rights themselves are granted by individuals, and must be given in the form of informed and explicit consent, and the specifics of use in them must be clearly defined in the contract to avoid misuse. There are also questions about ownership: Who or what owns the facial data, and who gets to use it? Beyond that, the General Data Protection Regulation (GDPR) in the European Union mandates strict use of personal data, and biometric data in particular [17]. These regulations prevent legal liabilities for contracts facilitating the sale of facial rights, thus contracts offering the sale of facial rights must be in conformity with these regulations. Additionally, these agreements involve identity manipulation concerns, which are ethical and creating transparancy and responsible usage essential. As AI technologies grow, the legal framework being developed to govern the sale of facial rights and rights more generally will evolve so that they are capable of keeping pace with innovation and yet protecting individuals' rights and freedoms.

## 4. ETHICAL IMPLICATIONS OF AI IN FACE SALE CONTRACTS

Artificial Intelligence (AI) is one of the most evolving technology of the 21st century that traveled all the way from the horizon of research labs to revolutionizing multiple industries and influencing daily life [18], [19]. AI is the means to build computer systems that model human intelligence in solving tasks considered to require intelligence [20]. In that sense, the idea of AI is nothing new, but it did take the advent of the advancements made lately in areas of machine learning, deep learning (DL) and computational power to practically enable the use of these technologies on a large scale [21]–[23]. Due to our not so distant past and the reached understanding of the basics of the knowledge systems theory, during this century, AI has already become a necessary component of many fields such as healthcare, finance, transportation, and retail and has become an unsurpassed tool for precision, efficiency, and inventiveness [24]. But with these benefits come huge legal and ethical problems, particularly with contracts and sales where the influence of AI is mounting [25]–[27].

The development of the AI has greatly evolved from simple automation to complex, self learning systems which can handle huge verse of data and make decisions to a minimum human intervention [28]. The early AI systems were basically rule based, that is, they could process incoming information following fixed rules, but were unable to learn something new. When came machine learning, this is where AI systems started to learn from data and find out patterns and learning from experience. More impetus to AI's capabilities was provided by deep learning — a subset of machine learning used to teach machines how to learn the way humans learn, by using neural networks to simulate human brain functioning [29], [30]. Because today's AI systems can not only analyze complex datasets to deliver insights and recommendations but can do it in real time, they are essential to industries where making quick, data driven decisions are paramount [31].

Sales and contracts is one of the sectors where AI has made a huge difference [32]. AI in sales is used to tailor customer experience, optimize the pricing strategy and simplify the customer service through the usage of the chatbots and virtual assistant [33], [34]. These tools allow businesses to communicate with customers better, funneling through unique solutions that are tailored to individuals preferences and behaviors [35]. For example, an e-commerce platform exploits the use of an AI algorithm that suggests products to a customer depending on their browsing history and past purchases to improve the chance of getting a conversion. Just like a customer inquiry round the clock can be done by AI-powered chat bots to optimize service quality without a drop in operational cost [36]. And in all this, AI became an integral part of the sales processes of the businesses [37].

AI is also changing the legal and contractual side of sales beyond customer interactions [38]. Today, several tools that let the AI draft, review, and manage the contract; automate the work that once took a lot of human effort [39]. The tools can help analyze the large volumes of legal documents, highlight the key clauses and can also warn about potential risks that takes away the time and cost from contract management. AI can create security through sales contract agreements, so that terms and conditions of all sales are fair, enforceable, and in compliance with the laws, giving an extra layer of security for buyers and sellers [40], [41]. In addition, AI systems can monitor contract performance in real time, sending alerts of either breaches or deviations from agreed upon terms [42]. This proactive approach will prevent disputes, and will make possible for all parties to fulfill the contractual obligations.

Yet the integration of AI into the sales contracts and legal processes poses far reaching security and ethical issues [27], [43]. The main concern is on liability. Traditionally in contract law, liability is imposed on the parties to the agreement. When an AI system does the drafting or enforcement of a contract, who tends to be accountable when something goes wrong? For example, imagine an AI tool inserts an error into a contract that results in financial loss or some legal dispute and who is to blame, the AI developer, the company using the tool, or the parties to the contract. The ambiguity adds further complication to a legal landscape whose lines are already becoming blurred by the use of AI in contractual agreements.

Liability concerns are not the only issues surrounding AI driven contracts: accountability, transparency and fairness are also important [44], [45]. In legal contexts, this lack of transparency can be a problem in allowing parties to know what the rationale might be for certain contractual terms, or decisions. An example AI tool would be something that changes the price of a product automatically according to a customer's purchasing behaviour, but the customer might not realize that

the price change was dictated by the AI tool. The lack of transparency poses a risk that AI could be used to leverage consumers, or to engineer contract language in favor of one side over another.
So, several ethical problems arise as a result of the rise of AI technologies in face sale contracts. All these technologies present new opportunities but also contain a range of complex risks. The ethics of the bias and fairness, privacy and consent, transparency and accountability, and environmental sustainability are this section's subject matters.

- **Bias and Fairness:** Facial AI systems are biased because they are trained on datasets that are painted with the same societal inequalities. For instance, facial recognition models have been shown to perform worse on some racial compared with others because they have not been trained on enough data which represents those groups. In the face sale contracts context, biased algorithms might result in unfairness towards (and/or wrongful denial of services to) some groups. If overlooked, these biases keep things as they are, or worse. Fairness means developers use diverse datasets, regular audits to detect and mitigate algorithmic bias in order to promote fairer outcomes.

- **Privacy and Consent:** There is a real problem with the sale of both real, and, yes, AI generated facial data, as well as its use without proper consent when it comes to individuals' faces. AI generated synthetic faces can look just like real people, making it difficult to tell the difference between real and artificial. It opens the door for someone to gain financially from the linking of their likeness without their knowledge. Furthermore, such privacy laws as GDPR [17] require explicit consent prior to collecting or using personal data — but privacy laws around AI generated content are often ambiguous. On the business side, face sale contracts need to be laden with clear consent mechanisms that give clear details about how the data or likeness of the person will be used, stored, and shared.

- **Transparency and Accountability:** AI systems that are used in face sale contracts are often 'black boxes' where we do not understand how they decide what to choose. Because of this lack of transparency, there can be a lack of trust between the parties they involve, because it doesn't make clear whether the facial data they use is real or AI generated. In addition, if AI generated faces are exploited for identity theft or fraud, chargeability for the actions can be very tricky—should the AI generators be held accountable, the service provider or the end user? Given these, expose of XAI methods would provide transparency on how models generate or classify facial data. Also, frameworks have to set out clear lines of accountability to ensure liability can be managed properly in cases of misuse or in breaches of an agreement.

- **Environmental Sustainability:** Generating the face or detecting it also has a big price for the environment when developing and deploying AI models. There is also the energy and computation that deeply learning models require to train — especially if this tool have to train a really big model on a really big dataset. The use of AI in face sale contracts therefore carries an ethical dimension and especially so as companies look for environmentally friendly technology practices. To reduce the computational burden, they should be working with energy efficient architectures like EfficientNet.

## 5. PROPOSED FRAMEWORK

Face sale contracts in the context of AI-generated images detection this paper presents further a comprehensive framework. The framework deals with a number of main elements for precise face identification and classification, consisting of data collection, preprocessing, DL model building and evaluation. This structured approach tries to make face sale contracts secure and compliant in the end (see Figure 1).

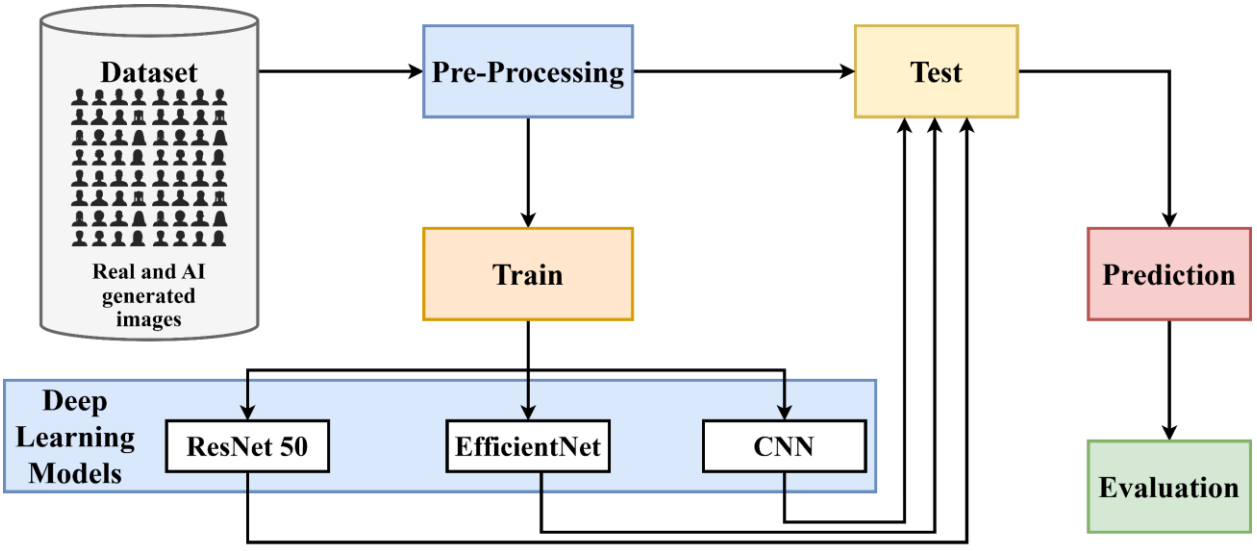


Fig. 1. The Methodology

This work uses a dataset from [46], which contains about 9,600 images of human faces that are specially curated for identifying and classifying real versus synthesized human images as shown in Figure 2. The dataset is organized into two categories optimized for DL applications: real images and AI generated images.

Real Images: This category has 5,000 diverse faces of authentic human faces ranging from diverse demographic backgrounds, expressions and backgrounds.

AI-Generated Images: The 4,630 images in this category were created using advanced AI algorithms, which try to mimic real human facial features as closely as possible.

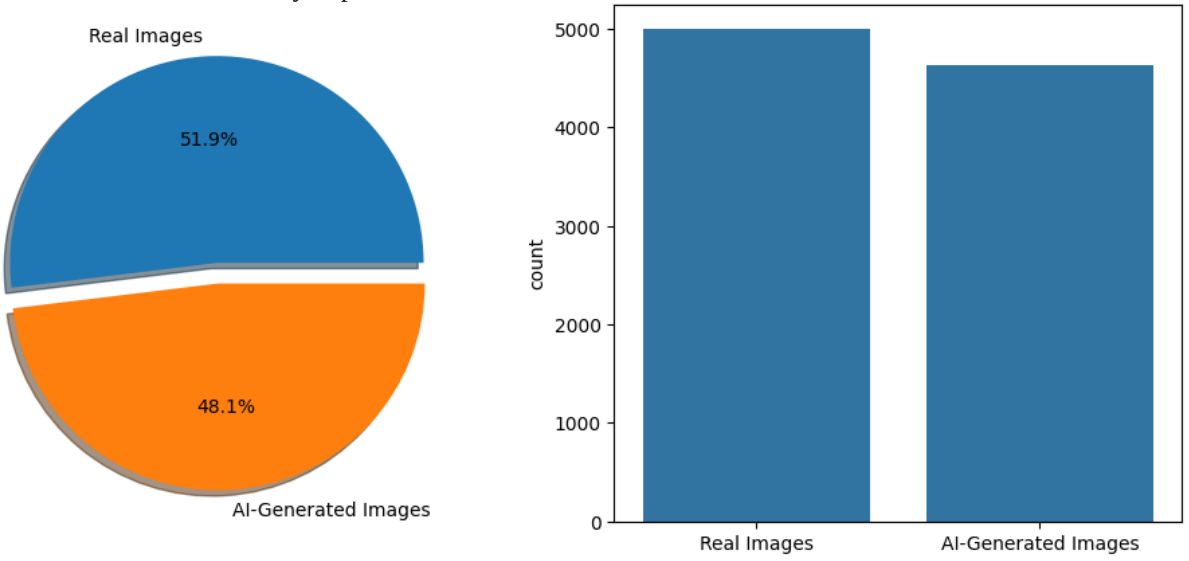


Fig. 2. classifying real versus AI-generated images

Supervised learning model training requires the distribution of the dataset to have the proper labels in order to assist the algorithms in pattern recognition of real and generated by AI images. This structured approach guarantees a solid grounding for proper classification inside the framework proposed here.

### 5.1. Preprocessing

In the preprocessing phase, the dataset is divided into three subsets such that we train, validate and test the models as shown in Figure 3. The data split was performed as follows:

- **Training Set:** I used 70% of the dataset for training the DL models. To optimize ability which the model can learn the features that can distinguish the real and AI generated face, this subset was used.
- **Validation Set:** During model training, 15% of dataset was reserved for validation. In order to avoid overfitting and additionally fine tune the model's hyperparsmeters, the model was validated against this validation set.
- **Test Set:** The remaining 15% of the dataset was used for testing. The performance of the final model on data it had not seen was evaluated on this test set, comparing how well the model performed, giving us an unbiased measure of the model's accuracy.

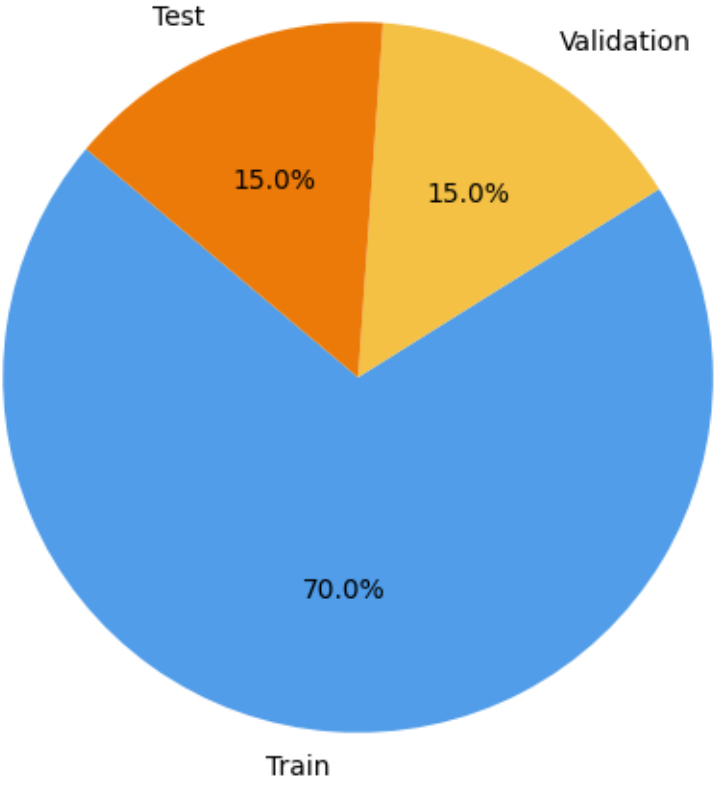


Fig. 3. Dataset Splitting

In addition to splitting the data, standard preprocessing techniques were applied, including image resizing [47], normalization [48], and augmentation [49]. These techniques help improve model performance by ensuring consistency in input image size, scaling pixel values, and introducing variability through augmentations such as rotations and flips. This preprocessing pipeline was essential for enhancing the robustness and generalization of the models during training.

## 5.2. Build DL Models

In this study, three DL models were implemented to classify real and AI-generated images: A custom CNN and ResNet50 and EfficientNet. Each model was chosen for its high performance in image classification tasks, while also having features advantages in feature extraction and efficiency.

- **ResNet50**: A DL architecture that is known to solve the problem of vanishing gradients in deep networks. ResNet50 employs the residual learning, which helps the network train deeper layer at least as well as the other networks [50]. We trained the model with Adam optimizer, batch size of 64 and 75 epochs and learning rate of 0.00001. So, using 50 layer architecture was critical for capturing fine detail in facial images, for telling real from AI generated faces.
- **EfficientNet**: EfficientNet was used to achieve balance between computational complexity and accuracy, which is its main characteristic of scalability and efficiency [51]. It was trained with Adam optimizer, a batch size of 32 over 15 epochs, a learning rate of 0.001 and the model was optimized. The precise classification under computational efficiency was perfectly suited for this large dataset using the EfficientNet architecture.
- **CNN**: Using of the Convolutional Neural Network (CNN) is designed for specific classification task. For feature extraction part, it consisted of several convolutional and pooling layers, and fully connected layers for

classification [30]. Adam optimizer, batch size of 32, 10 epochs and learning rate of 0.001 were used for training the CNN.

The study was able to compare the performance of these different architectures by using these different architectures to see how they do in identifying real and AI generated faces. To maximize a relevant measurement of classification accuracy, each model was trained with tuned hyperparameters.

### 5.3. Evaluation

The performance of the three models (ResNet50, EfficientNet, and CNN) was evaluated using several key metrics (Confusion Matrix, Accuracy, Precision, Recall, F1-Score, and Fitting). This comparison allowed to have an overview of how good each model was at classifying real and AI generated images.

- Confusion Matrix: Each model was visualized by a generated confusion matrix to view classification results [29]. The distribution of true positives, true negatives, false positives and false negatives was shown in this matrix to identify the number of correct and incorrect predictions. It gave us some insight about places where the model was not able to classify image correctly, including the successes as well as the misclassifications.
  - **Accuracy**: To know how often real and AI generated images are correctly classified, the accuracy is computed on the entire set. This is a metric that represents how many correctly predicted instances out of all predicted instances [52]. One general benchmark was accuracy and performance of the models was compared.
  - **Precision**: To evaluate the models ability to correctly identify AI generated images, precision was computed. Specifically, as a measurement of how often in cases of an image predicted as AI generated the model is correct, it calculates the proportion of true positive predictions out of all of the positive predictions [53].
  - **Recall**: To assess how well the models found all of the instances of AI generated images, recall was used. More specifically, it computes what proportion of true positives exist among all true positives [54] and tells us how many of the AI generated faces were classified correctly by each model
  - **F1-Score**: When the classes don't balance out, F1 Score is computed to balance precision and recall. It stands for the harmonic mean of precision and recall and have a single indicator for false positives and false negatives and overall models' behaviour [53].
  - **Fitting**: This evaluated was used to know how the fitting process for each model and if it was underfitting nor overfitting [31]. The training and validation loss curves were checked to see if it have converged and if there are any big discrepancies between training and validation performance, if so it could point to over fitting. This was another tool to help ensure the models generalise correctly when the unseen data is presented.

The models were evaluated using these evaluation metrics to determine overall classification accuracy, false positive and false negative minimization and tradeoff between precision and recall. Through these evaluations, these ResNet50, EfficientNet, and CNN models were able to test their effectiveness in detecting real and AI generated faces, and what strengths they exhibited and what their weaknesses were.

## 6. RESULTS AND DISCUSSIONS

To assess the performance of the models for classifying AI-generated and real images, we analyzed the confusion matrices for each model (EfficientNet, ResNet50 and CNN). These confusion matrices tell us something about how much the models can really recognise the difference between images generated by AI and true ones.

In particular, the ResNet50 model demonstrates great capability of distinguishing AI generated images and real ones. From the confusion matrix we know that ResNet50 correctly classified the 700 AI generated images without any false positives or misclassifications. The model also succeeds in finding all 745 real images without any false negative. We can see this means the ResNet50 model was 100% accurate in both categories, confirming that it can extract and interpret features that separate out real content from AI generated content..

For the EfficientNet model, the confusion matrix also shows high accuracy. On 695 of 700 AI generated images it correctly classified all 695 images with none of which was incorrectly labeled as AI generated. Moreover, it also was able to correctly classify all 750 real images with the model, maintaining the robust performance. In fact, for that detection of AI generated images ResNet50 is slightly more accurate but overall EffectiveNet is highly effective for image classification too.

However, the results of the CNN model are more varied. From the confusion matrix, while CNN correctly categorized 516 of the AI generated images as such, it incorrectly identified 484 as real, thus resulting in a relatively high amount of false negatives. Also, 443 real images were correctly identified, 483 of them were wrongly identified as being AI generated.

While CNN is a significantly worse performer compared to ResNet, EfficientNet, the performance is still insightful — that is, for identifying which parts of the model need work, especially regarding false positives and negatives. The confusion matrices for these three models are shown in Figure 4.

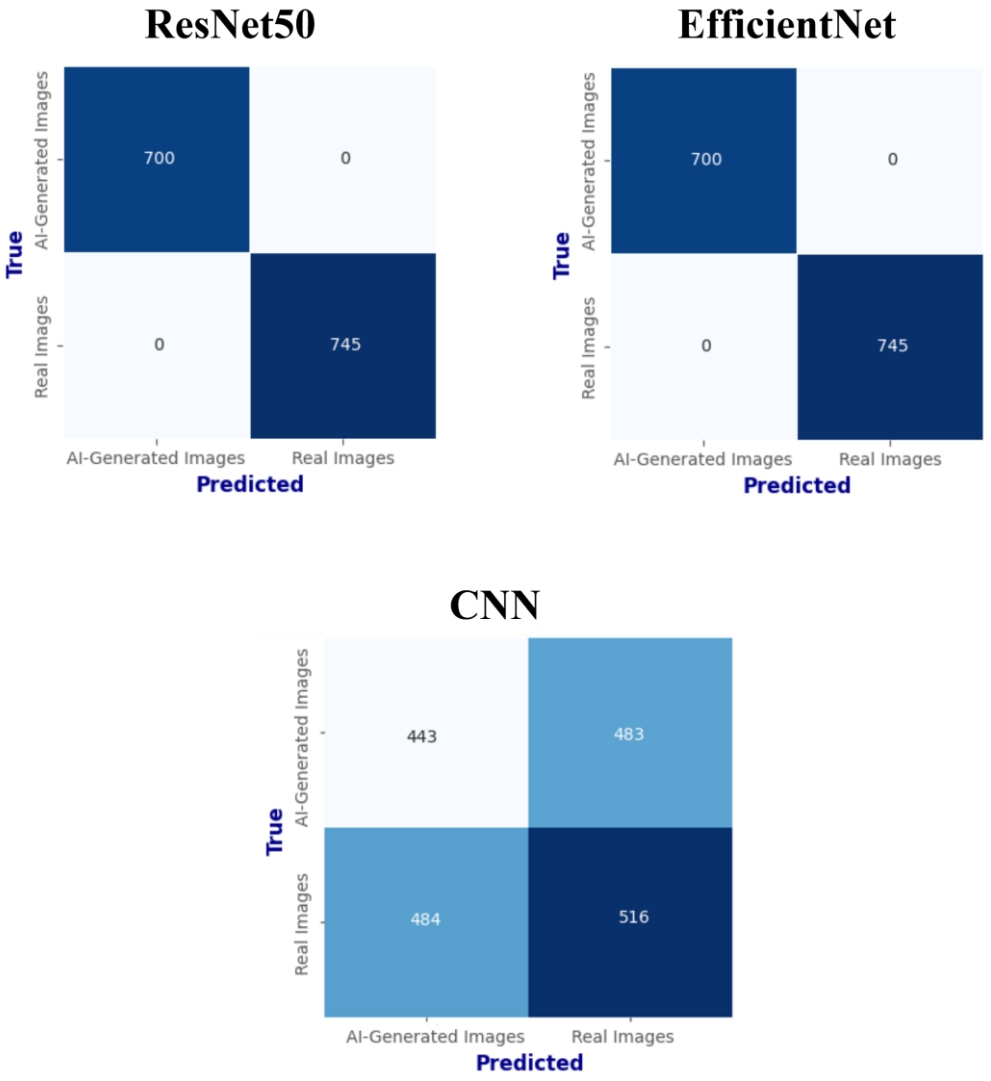


Fig. 4. The confusion matrices of DL models

Table 1 compares the performance of three DL models ResNet50, EfficientNet, and CNN across four key metrics. In the field of machine learning, we have Accuracy, Precision, Recall, and F1–score. All metrics were perfect (score of 1.00) for ResNet50 and EfficientNet, indicating perfect classification result and therefore, perfect precision, recall and balance (precision ratio vs. recall ratio, F1 score). That means that these models were all to accurately label all the instances and had no false positives and false negatives. ResNet50, having deep residual architecture, is effective at learning complex features, since the vanishing gradient problem is addressed, and EfficientNet is designed to be computationally efficient. The generic CNN model performed worse, with corresponding Accuracy of 0.50, Precision of 0.52, Recall of 0.51 and F1-score of 0.52.

TABLE I. COMPARES THE PERFORMANCE OF DL MODELS

| DL Method | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| **Resnet50** | *1.00* | *1.00* | *1.00* | *1.00* |
| **EfficientNet** | *1.00* | *1.00* | *1.00* | *1.00* |
| **CNN** | *0.50* | *0.52* | *0.51* | *0.52* |

The two higher performances inferred (ResNet50 and EfficientNet models) are studied, and the results are assessed based on the accuracy and loss metrics provided for training and validation. Figure 5 shows how the loss decrease and the accuracy of ResNet50 increase with epochs at the same time, showing good learning and no overfitting. The aligned training and validation curves show that. On the other hand, in comparison, the EfficientNet has loss and accuracy fluctuations and accordingly training process appears to be less stable and overfitting risk rises significantly even when validation accuracy doesn't keep improving properly.

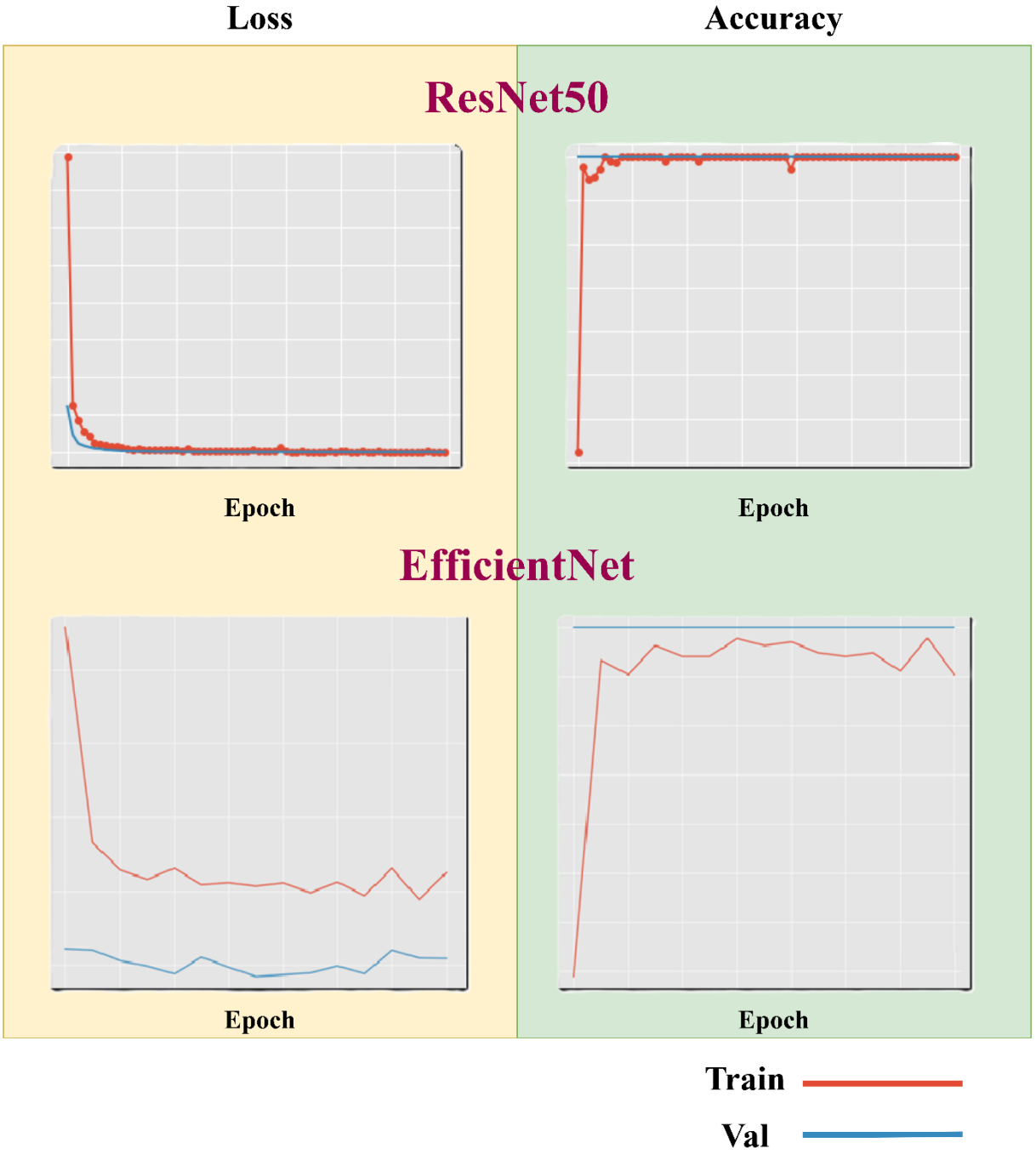


Fig. 5. ResNet50 and EfficientNet models

Observations on these correlated data suggest that ResNet50 is the more stable and more robust model for this classification task as it is able to fit the data better than the other 9 models. Therefore, ResNet50 is the best model to separate real and AI generated images.

Our study in the testing phase of the ResNet model tested its ability to classify images as either AI generated or real. The label in the model classification column shows what the ResNet model predicted for labels, 0 for AI generated images and 1 for real images. As can be seen in Table 2 the ResNet showed high accuracy in classifying the majority of the images as predicted. Suppose, for instance, the model correctly classified an AI-generated image (label encoded as 0) of something in row 0 as 0. Just like in row 1, a real image (labeled as 1) was accurately predicted as 1. This pattern is consistent through the dataset. There are many instances of correct classifications out of the 1,445 images (rows 2, 3, 4, 1442, 1443 and others).

TABLE II. TEST THE RESNET50 MODEL

| ID | image_path | label | Label encoded | Model classification |
|---|---|---|---|---|
| 0 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |
| 1 | /content/drive/MyDrive/Colab Notebooks/CR/cr/R... | Real Images | 1 | 1 |
| 2 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |
| 3 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |
| 4 | /content/drive/MyDrive/Colab Notebooks/CR/cr/R... | Real Images | 1 | 1 |
| ... | ... | ... | ... | ... |
| 1440 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |
| 1441 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |
| 1442 | /content/drive/MyDrive/Colab Notebooks/CR/cr/R... | Real Images | 1 | 1 |
| 1443 | /content/drive/MyDrive/Colab Notebooks/CR/cr/R... | Real Images | 1 | 1 |
| 1444 | /content/drive/MyDrive/Colab Notebooks/CR/cr/A... | AI-Generated Images | 0 | 0 |

This table shows the model's performance which verifies the conclusion made earlier that ResNet50 is the best model for this task.

## 7. CONCLUSION AND FUTURE WORKS

In this study we evaluated how well ResNet50, EfficientNet and CNN models distinguish real and AI generated images. Detailed training and validation metrics analysis for demonstrated ResNet50's greater stability in learning ability and robustness without any signatures of overfitting and its training and validation accuracy and loss stayed aligned with the typical ResNet50 behaviour. However, EfficientNet model performed very differently from the others on loss and accuracy: much more varied on loss and accuracy, implying a more unstable training process with a higher possibility of overfitting in the training process. With these results, we find ResNet50 to be the best model for the classification of AI-generated vs real images and demonstrated high accuracy of identifying between images that have been generated by AI and those created by humans. Finally, the ResNet50 achieved the final testing results, where it was able to classify most of test images with a precision. Future work will delve into other deep learning models and architectures beyond ResNet50 and EfficientNet to build a better classification accuracy. Investigated will be techniques such as ensemble learning in which multiple models are combined to increase predictions performance. Furthermore, by including a wider range of AI generated images, specifically the ones generated by state of the art generative models will help us better understand how robust the classification models are. In addition, methods to explainability like Grad-CAM can be implemented to interpret the model's decision making process and explain how the model differentiate one from the other between real and AI image. They will form part of the development of more resilient and more transparent image classification systems.

### Conflicts of interest

The author's disclosure statement confirms the absence of any conflicts of interest.

### Funding

The author's paper clearly indicates that the research was conducted without any funding from external sources.

### Acknowledgement

The author extends appreciation to the institution for their unwavering support and encouragement during the course of this research.

## References


[1] B. Khoo, R. C.-W. Phan, and C.-H. Lim, "Deepfake attribution: On the source identification of artificially generated images," *WIREs Data Min. Knowl. Discov.*, vol. 12, no. 3, p. e1438, 2022, doi: https://doi.org/10.1002/widm.1438.

[2] L. L. Dhirani, N. Mukhtiar, B. S. Chowdhry, and T. Newe, "Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review," *Sensors*, vol. 23, no. 3, 2023, doi: 10.3390/s23031151.

[3] W. K. Hon, "Chapter 23: Artificial intelligence: challenges and risks," in *Technology and Security for Lawyers and Other Professionals*, Cheltenham, UK: Edward Elgar Publishing, 2024, pp. 512–540. doi: 10.4337/9781803923918.00035.

[4] B. van der Sloot, *Regulating the Synthetic Society: Generative AI, Legal Questions, and Societal Challenges*. Bloomsbury Academic, 2024.

[5] J. Afzal, "Implementation of Digital Law as a Legal Tool in the Current Digital Era." Springer.

[6] M. Brundage *et al.*, "Toward trustworthy AI development: mechanisms for supporting verifiable claims," *arXiv Prepr. arXiv2004.07213*, 2020.

[7] A. S. Albahri *et al.*, "A systematic review of trustworthy artificial intelligence applications in natural disasters," *Comput. Electr. Eng.*, vol. 118, p. 109409, 2024, doi: 10.1016/j.compeleceng.2024.109409.

[8] Y. L. Khaleel, M. A. Habeeb, A. S. Albahri, T. Al-Quraishi, O. S. Albahri, and A. H. Alamoodi, "Network and cybersecurity applications of defense in adversarial attacks: A state-of-the-art using machine learning and deep learning methods," *J. Intell. Syst.*, vol. 33, no. 1, 2024, doi: 10.1515/jisys-2024-0153.

[9] J. J. Bird and A. Lotfi, "CIFAKE: Image Classification and Explainable Identification of AI-Generated Synthetic Images," *IEEE Access*, vol. 12, pp. 15642–15650, 2024, doi: 10.1109/ACCESS.2024.3356122.

[10] Y. S. Taspinar and I. Cinar, "Distinguishing Between AI Images and Real Images with Hybrid Image Classification Methods," in *2024 13th Mediterranean Conference on Embedded Computing (MECO)*, 2024, pp. 1–4. doi: 10.1109/MECO62516.2024.10577770.

[11] F. M. Salman and S. S. Abu-Naser, "Classification of real and fake human faces using deep learning," 2022.

[12] D. J. Stephens, "Introduction: The contract for the Sale of Goods," in *Sale of Goods*, Informa Law from Routledge, 2020, pp. 1–55.

[13] S. K. Singh, A. Tiwary, and N. Gupta, *Business Law*. RAJEEV BANSAL, 2021.

[14] E. Apaydin, "The principle of good faith in contracts," *Under Int. Unif. laws cisg, unidroit Princ. Princ. Eur. Contract law.–2019*, 2019.

[15] A. Kull, "Restitution as a Remedy for Breach of Contract," in *Restitution*, Routledge, 2020, pp. 293–346.

[16] M. G. Bridge, *The international sale of goods*. Oxford University Press, 2017.

[17] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Pract. Guid. 1st Ed., Cham Springer Int. Publ.*, vol. 10, no. 3152676, pp. 10–5555, 2017.

[18] R. Li, *Artificial intelligence revolution: How AI will change our society, economy, and culture*. Simon and Schuster, 2020.

[19] S. Sheikh, *Understanding the role of artificial intelligence and its future social impact*. IGI Global, 2020.

[20] I. H. Sarker, "AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems," *SN Comput. Sci.*, vol. 3, no. 2, p. 158, 2022, doi: 10.1007/s42979-022-01043-x.

[21] S. F. Ahmed *et al.*, "Deep learning modelling techniques: current progress, applications, advantages, and challenges," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13521–13617, 2023, doi: 10.1007/s10462-023-10466-8.

[22] C. Sarkar *et al.*, "Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development," *Int. J. Mol. Sci.*, vol. 24, no. 3, 2023, doi: 10.3390/ijms24032026.

[23] L. Alzubaidi *et al.*, "Reliable deep learning framework for the ground penetrating radar data to locate the horizontal variation in levee soil compaction," *Eng. Appl. Artif. Intell.*, vol. 129, p. 107627, 2024, doi: 10.1016/j.engappai.2023.107627.

[24] A. Bin Rashid and M. D. A. K. Kausik, "AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications," *Hybrid Adv.*, vol. 7, p. 100277, 2024, doi: https://doi.org/10.1016/j.hybadv.2024.100277.

[25] A. S. Albahri, Y. L. Khaleel, and M. A. Habeeb, "The Considerations of Trustworthy AI Components in Generative AI; A Letter to Editor," *Appl. Data Sci. Anal.*, vol. 2023, pp. 108–109, 2023, doi: 10.58496/adsa/2023/009.

[26] A. S. Albahri *et al.*, "A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion," *Inf. Fusion*, vol. 96, pp. 156–191, 2023, doi: 10.1016/j.inffus.2023.03.008.

[27] L. Alzubaidi *et al.*, "Towards Risk-Free Trustworthy Artificial Intelligence: Significance and Requirements," *Int. J. Intell. Syst.*, vol. 2023, p. 4459198, 2023, doi: 10.1155/2023/4459198.

[28] S. S. Rajest, B. Singh, A. J Obaid, R. Regin, and K. Chinnusamy, "Recent developments in machine and human

intelligence," 2023.
[29] Y. L. Khaleel, "Fake News Detection Using Deep Learning," University of Miskolc, 2021. doi: http://dx.doi.org/10.13140/RG.2.2.31151.75689.
[30] M. A. Habeeb, "Hate Speech Detection using Deep Learning Master thesis," University of Miskolc, 2021. [Online]. Available: http://midra.uni-miskolc.hu/document/40792/38399.pdf
[31] S. Dadvandipour and Y. L. Khaleel, "Application of deep learning algorithms detecting fake and correct textual or verbal news," *Prod. Syst. Inf. Eng.*, vol. 10, no. 2, pp. 37–51, 2022, doi: 10.32968/psaie.2022.2.4.
[32] D. Grewal, A. Guha, C. B. Satornino, and E. B. Schweiger, "Artificial intelligence: The light and the darkness," *J. Bus. Res.*, vol. 136, pp. 229–236, 2021, doi: https://doi.org/10.1016/j.jbusres.2021.07.043.
[33] M. S. Bhuiyan, "The role of AI-Enhanced personalization in customer experiences," *J. Comput. Sci. Technol. Stud.*, vol. 6, no. 1, pp. 162–169, 2024.
[34] A. Akyüz and K. Mavnac\io\uglu, "Marketing and Financial Services in the Age of Artificial Intelligence," in *Financial Strategies in Competitive Markets: Multidimensional Approaches to Financial Policies for Local Companies*, H. Dinçer and S. Yüksel, Eds., Cham: Springer International Publishing, 2021, pp. 327–340. doi: 10.1007/978-3-030-68612-3_23.
[35] I. Kulkov, "The role of artificial intelligence in business transformation: A case of pharmaceutical companies," *Technol. Soc.*, vol. 66, p. 101629, 2021, doi: https://doi.org/10.1016/j.techsoc.2021.101629.
[36] J. Paschen, M. Wilson, and J. J. Ferreira, "Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel," *Bus. Horiz.*, vol. 63, no. 3, pp. 403–414, 2020, doi: https://doi.org/10.1016/j.bushor.2020.01.003.
[37] S. Mishra and A. R. Tripathi, "AI business model: an integrative business approach," *J. Innov. Entrep.*, vol. 10, no. 1, p. 18, 2021.
[38] B. Libai *et al.*, "Brave New World? On AI and the Management of Customer Relationships," *J. Interact. Mark.*, vol. 51, no. 1, pp. 44–56, Aug. 2020, doi: 10.1016/j.intmar.2020.04.002.
[39] F. Lagioia, A. Jabłonowska, R. Liepina, and K. Drazewski, "AI in Search of Unfairness in Consumer Contracts: The Terms of Service Landscape," *J. Consum. Policy*, vol. 45, no. 3, pp. 481–536, 2022, doi: 10.1007/s10603-022-09520-9.
[40] M. Stone *et al.*, "Artificial intelligence (AI) in strategic marketing decision-making: a research agenda," *Bottom Line*, vol. 33, no. 2, pp. 183–200, 2020, doi: 10.1108/BL-03-2020-0022.
[41] S. Youn and S. V. Jin, "'In A.I. we trust?' The effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging 'feeling economy,'" *Comput. Human Behav.*, vol. 119, p. 106721, 2021, doi: https://doi.org/10.1016/j.chb.2021.106721.
[42] S. Rusthollkarhu, S. Toukola, L. Aarikka-Stenroos, and T. Mahlamäki, "Managing B2B customer journeys in digital era: Four management activities with artificial intelligence-empowered tools," *Ind. Mark. Manag.*, vol. 104, pp. 241–257, 2022, doi: https://doi.org/10.1016/j.indmarman.2022.04.014.
[43] R. Max, A. Kriebitz, and C. Von Websky, "Ethical Considerations About the Implications of Artificial Intelligence in Finance," in *Handbook on Ethics in Finance*, L. San-Jose, J. L. Retolaza, and L. van Liedekerke, Eds., Cham: Springer International Publishing, 2021, pp. 577–592. doi: 10.1007/978-3-030-29371-0_21.
[44] N. Balasubramaniam, M. Kauppinen, K. Hiekkanen, and S. Kujala, "Transparency and Explainability of AI Systems: Ethical Guidelines in Practice," in *Requirements Engineering: Foundation for Software Quality*, V. Gervasi and A. Vogelsang, Eds., Cham: Springer International Publishing, 2022, pp. 3–18.
[45] B. Memarian and T. Doleck, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review," *Comput. Educ. Artif. Intell.*, vol. 5, p. 100152, 2023, doi: https://doi.org/10.1016/j.caeai.2023.100152.
[46] K. Dhote, "Human Faces Dataset." https://www.kaggle.com/datasets/kaustubhdhote/human-faces-dataset
[47] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance," in *Applications in Electronics Pervading Industry, Environment and Society*, S. Saponara and A. De Gloria, Eds., Cham: Springer International Publishing, 2022, pp. 10–17.
[48] S. Albert *et al.*, "Comparison of Image Normalization Methods for Multi-Site Deep Learning," *Appl. Sci.*, vol. 13, no. 15, 2023, doi: 10.3390/app13158923.
[49] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," *Pattern Recognit.*, vol. 137, p. 109347, 2023, doi: https://doi.org/10.1016/j.patcog.2023.109347.
[50] D. Theckedath and R. R. Sedamkar, "Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks," *SN Comput. Sci.*, vol. 1, no. 2, p. 79, 2020, doi: 10.1007/s42979-020-0114-9.
[51] H. Le Duc, T. T. Minh, K. V. Hong, and H. L. Hoang, "84 Birds Classification Using Transfer Learning and EfficientNetB2," in *Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and*

*Industry 4.0 Applications*, T. K. Dang, J. Küng, and T. M. Chung, Eds., Singapore: Springer Nature Singapore, 2022, pp. 698–705.

[52] F. K. H. Mihna, M. A. Habeeb, Y. L. Khaleel, Y. H. Ali, and L. A. E. Al-saeedi, "Using Information Technology for Comprehensive Analysis and Prediction in Forensic Evidence," *Mesopotamian J. CyberSecurity*, vol. 2024, pp. 4–16, Mar. 2024, doi: 10.58496/MJCS/2024/002.

[53] M. A. Habeeb, Y. L. Khaleel, and A. S. Albahri, "Toward Smart Bicycle Safety: Leveraging Machine Learning Models and Optimal Lighting Solutions," in *Proceedings of the Third International Conference on Innovations in Computing Research (ICR'24)*, K. Daimi and A. Al Sadoon, Eds., Cham: Springer Nature Switzerland, 2024, pp. 120–131.

[54] H. M. Abdulfattah, K. Y. Layth, and A. A. Raheem, "Enhancing Security and Performance in Vehicular Adhoc Networks: A Machine Learning Approach to Combat Adversarial Attacks," *Mesopotamian J. Comput. Sci.*, vol. 2024, pp. 122–133, 2024, doi: 10.58496/MJCSC/2024/010.