*remote sensing*

MDPI

*Article*

# An Efficient Approach Based on Privacy-Preserving Deep Learning for Satellite Image Classification

**Munirah Alkhelaiwi [1], Wadii Boulila [1,2,*], Jawad Ahmad [3], Anis Koubaa [4] and Maha Driss [1,2]**

[1] College of Computer Science and Engineering, Taibah University, Medina 42353, Saudi Arabia; Munirah.Alkhelaiwi@gmail.com (M.A.); maha.idriss@riadi.rnu.tn (M.D.)
[2] RIADI Laboratory, University of Manouba, Manouba 2010, Tunisia
[3] School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, UK; J.Ahmad@napier.ac.uk
[4] Robotics and Internet-of-Things Laboratory, Prince Sultan University, Riyadh 12435, Saudi Arabia; akoubaa@psu.edu.sa
* Correspondence: wadii.boulila@riadi.rnu.tn

**Abstract:** Satellite images have drawn increasing interest from a wide variety of users, including business and government, ever since their increased usage in important fields ranging from weather, forestry and agriculture to surface changes and biodiversity monitoring. Recent updates in the field have also introduced various deep learning (DL) architectures to satellite imagery as a means of extracting useful information. However, this new approach comes with its own issues, including the fact that many users utilize ready-made cloud services (both public and private) in order to take advantage of built-in DL algorithms and thus avoid the complexity of developing their own DL architectures. However, this presents new challenges to protecting data against unauthorized access, mining and usage of sensitive information extracted from that data. Therefore, new privacy concerns regarding sensitive data in satellite images have arisen. This research proposes an efficient approach that takes advantage of privacy-preserving deep learning (PPDL)-based techniques to address privacy concerns regarding data from satellite images when applying public DL models. In this paper, we proposed a partially homomorphic encryption scheme (a Paillier scheme), which enables processing of confidential information without exposure of the underlying data. Our method achieves robust results when applied to a custom convolutional neural network (CNN) as well as to existing transfer learning methods. The proposed encryption scheme also allows for training CNN models on encrypted data directly, which requires lower computational overhead. Our experiments have been performed on a real-world dataset covering several regions across Saudi Arabia. The results demonstrate that our CNN-based models were able to retain data utility while maintaining data privacy. Security parameters such as correlation coefficient (−0.004), entropy (7.95), energy (0.01), contrast (10.57), number of pixel change rate (4.86), unified average change intensity (33.66), and more are in favor of our proposed encryption scheme. To the best of our knowledge, this research is also one of the first studies that applies PPDL-based techniques to satellite image data in any capacity.

**Keywords:** privacy-preserving deep learning; deep learning; remote sensing; privacy-preservation; convolutional neural network; homomorphic encryption; paillier scheme



**Citation:** Alkhelaiwi, M.; Boulila, W.; Ahmad, J.; Koubaa, A.; Driss, M. An Efficient Approach Based on Privacy-Preserving Deep Learning for Satellite Image Classification. *Remote Sens.* **2021**, *13*, 2221. https://doi.org/10.3390/rs13112221

Academic Editors: Do-Hyung Kim, Anupam Anand, Joseph Bullock and Miguel Luengo-Oroz

Received: 14 April 2021
Accepted: 3 June 2021
Published: 6 June 2021



## 1. Introduction

Satellite images of earth are generated by imaging satellites, which may be operated by governments or enterprises. These images are captured through remote sensing (RS) technologies and, generally, RS can be described as the process of collecting and analyzing information about an entity, region or event without being in physical contact with it [1]. RS data is considered a very useful source of information for many applications, such as land use classification, especially when integrated with artificial intelligence technology [2–4]. The size of satellite images is increasing because of the growing demand for

better resolution of images, and along related lines, the growing amount of RS data has enabled the study of various complex research topics [5–7]. However, producing adequate RS images typically requires applying emerging DL-based techniques with complex architecture and computational workload. To provide this, many researchers use cloud computing platforms to apply DL techniques that enable them to extract insights and useful information. However, in such cases, data workflows can be subject to privacy concerns because of the public nature of that data processing and the tools used to manage it. Here, data privacy cannot be ensured, and data leakage may occur. However, there are still several benefits of using cloud computing, such hiding architecture complexity, cost savings, flexibility, and scalability. Thus, the optimal solution is one that can balance the downsides in order to reap the benefits. However, there are also several privacy challenges that require addressing, particularly in cases where satellite images are transmitted or stored using public DL techniques [8]. Al-Rubaie and Chang provide an overview of the privacy-preservation deep learning (PPDL) techniques that can be adopted to safeguard the privacy of either individual or business users [9]. From such work it becomes apparent that PPDL techniques can be used to benefit from public data analytics while also preventing data leakage and keeping sensitive information private from unauthorized access and illegal usage.

DL is commonly used to build predictive models for image processing and both speech and text recognition applications. These models are more precise, especially when trained on large data sets. Prediction is a method of studying available data and then using that expertise to produce new information that was not available before then. In many cases, though, these data also contain sensitive information that likewise requires preservation. Therefore, an important challenge here is to preserve the privacy of such data when they are sent to the public cloud for processing and analysis. In most cases, personnel computers lack the performance capabilities needed to process massive satellite images. Therefore, in order to extract useful knowledge and insights from such RS data, there is a greater need to perform big data analysis using public cloud servers. With this growing reliance on cloud services, the privacy of data collected and processed by cloud service providers during DL training is also becoming a more challenging concern [10]. Satellite images could contain sensitive information, such as oilfield, airport, and military locations, that can be stolen and misused. Likewise, if such images are processed without protection, then this makes it easy for sensitive information to be extracted and used for illegal purposes. This indicates the need to find a reliable privacy method that will ensure that big satellite images are encrypted over cloud servers in ways that cannot be compromised. Thus, exploring PPDL techniques applied to satellite images becomes both a challenging topic and a potentially rewarding one.

The motivation driving this specific research project is to link two cutting-edge research topics, which are DL and privacy. Indeed, the progress of machine learning (ML) and its subfield of deep learning (DL) need not come at the expense of privacy or data security. Therefore, this research work proposes a powerful approach based on PPDL utilized on big satellite images in order to maintain anonymity and safeguard privacy related to data. Our main contribution is to apply PPDL for satellite images' data, which, to the best of our knowledge, is an approach that has not been proposed or attempted anywhere in the literature. In particular, the contributions of this study are:

- Proposing a PPDL technique, namely partially homomorphic encryption (PHE), for privacy-preserving satellite image classification.
- Applying the PHE technique on a proposed DL-based CNN model and existing transfer learning models. (As our results later will demonstrate, this results in promising performance for two test cases.)
- Testing the performance of the proposed PPDL technique on a real-world satellite image dataset.

The remainder of this paper is organized as follows: Section 2 describes the theoretical background, specifically by introducing DL and PPDL as techniques. A review of related

work is presented in Section 3, while Section 4 details the proposed method of using PPML for satellite image classification. Experimental results are reported and discussed in Section 5. Concluding remarks and directions for future work are presented in Section 6.

## 2. Background

This section introduces general background information about DL and PPDL techniques, including their advantages and disadvantages.

### *2.1. Convolutional Neural Network*

There are several different DL approaches commonly used for data processing [11], but the CNN is one of the most common specifically for image-based data processing. CNN is a category of deep neural networks (DNN), which itself is a substantial model of machine learning (ML). CNN is often applied to visual image processing and computer vision [12]. The most substantial presumption regarding issues addressed by CNN is regarding spatially-based characteristics. For instance, in a n object recognition system, we do not need to give thought to the objects' location in the images. The main concern is to detect them throughout the provided images, regardless of their actual location therein. Another essential characteristic of CNN is found in how it acquires conceptual characteristics as data spread into the deeper layers. For instance, the edge could be identified throughout the first layer of the image classification, and then simple features could be identified in the second layer before the top-level features such as objects are identified in the next layers [13]. In this way, CNN addresses the over-fitting issue wherein the neuron within a layer will be connected to the previous layer with a small region rather than all neurons, as usually happens in fully-connected neural networks.

The architecture of CNN models is composed of a set of layers. This set begins with an input layer, continues with a stack of hidden layers, and concludes with an output layer, and in this arrangement, the output of one layer becomes an input of the next layer. Therefore, the basic architecture of CNN consists of three layers, namely convolutional (CONV) layers, pooling (POOL) layers and fully connected (FC) layers, as described below [12,14]. Any middle layers are considered hidden since the activation function and final convolution cover their inputs and outputs [15]. Figure 1 illustrates the basic architecture of a CNN model.

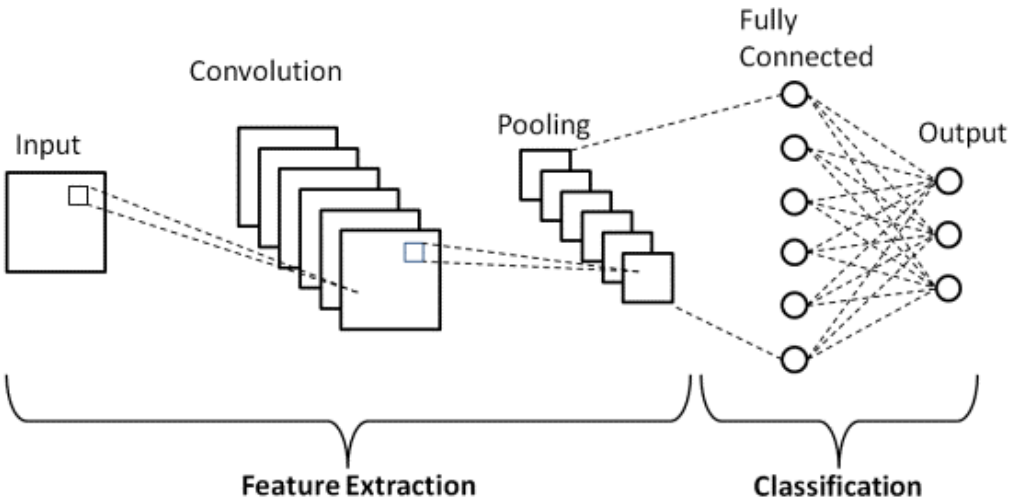


**Figure 1.** CNN Basic Architecture [16].

- Convolutional (CONV) layer: This is the first layer and key component of the CNN [17]. Most of the intensive computational loading is done in such layers. In CONV layers, the image is converted using filters, which are small units applied across the data via a sliding window. CONV layers elicit various features from the input image and, in this way, one image becomes a stack of filtered images.

- Pooling (POOL) layer: CONV layers are followed by POOL layers, which involve sub-sampling of features. Pooling layers progressively reduce the spatial size of the representation, which further decreases the computational load on the network.
- Fully connected (FC) layer: Fully connected layers are used in the last few layers and serve as a classifier. FC layers involve converting the complete pooled feature map matrix into one column, which is then loaded for processing to the neural network.

Transfer Learning

Transfer learning (TL) is a DL approach for transferring knowledge from one pre-trained model to another [18]. TL is commonly used when training a small dataset where the CNN's weights are initialized before being fine-tuned with the new dataset [19]. TL aids in adapting current models trained on large datasets to work in a specific context [20]. There are several pre-trained models approaches based on this research, including VGG16, ResNet-50, Xception and DenseNet121. Each of these common TL approaches is described below:

- VGG16: Simonyan and Zisserman (2014) proposed the architecture of the VGG16 model. VGG16 is a CNN model that consists of 16 hidden layers, including a total with convolutional, max pooling and fully connected layers. VGG16 was trained on the ImageNet dataset, which consists of 1,000,000 images. VGG16 is constructed of five blocks of convolutional layers with a $3 \times 3$ filter and stride of 1. After each convolution, an activation function (ReLU) is executed, followed by a max-pooling process with a $2 \times 2$ max filter and stride of 2. At the end of the five blocks, three FC layers are added: the first two layers with 4096 neurons and an ReLU activation function each, and the third layer with 1000 neurons and a SoftMax activation function [21]. The default input size is $224 \times 224 \times 3$ [22].
- ResNet-50: The ResNet model's architecture was proposed in 2015 by He et al. ResNet-50 is a 50 convolutional neural network layers pre-trained on the ImageNet dataset [23]. The fundamental concept behind the ResNet model is to use shortcut links to bypass blocks of convolutional layers (bottleneck). The CONV layers each have a $3 \times 3$ filter and are designed according to two rules: (1) the layers have the same number of filters with the same output feature map size and (2) the number of filters is multiplied if the feature map size is halved. The convolutional layers conduct the downsampling with a stride of 2. The network ends with an average POOL layer and 1000 FC layers with a SoftMax activation function. The default input size is $224 \times 224 \times 3$ [24].
- Xception: The Xception model's architecture was proposed by Chollet (2017). This model is a CNN-based architecture also trained on the ImageNet dataset. The Xception architecture comprises 36 CONV layers with a $3 \times 3$ filter and stride of 2. These CONV layers are structured into 14 modules, all of which have the ReLu activation function except for the first and last modules. The FC layer is replaced with a global average POOL layer and the default input size is $299 \times 299 \times 3$ [25].
- DenseNet121: Huang et al. (2017) proposed the architecture of the DenseNet121 model, another CNN-based architecture trained on the ImageNet dataset. DenseNet121 is composed of 5 dense blocks. The first block consists of a convolution layer with a $7 \times 7$ filter and stride of 2 and a MaxPooling layer with a $3 \times 3$ max filter and stride of 2. The remaining blocks consist of BatchNormalization, the ReLU activation function, and two CONV layers with $1 \times 1$ and $3 \times 3$ filters. A transition layer follows each block except for the last, which instead is followed by a classification layer. In DenseNet121, all previous feature-maps are used as input in each layer. The default input size is $224 \times 224 \times 3$ [26].

### *2.2. Privacy-Preservation Deep Learning*

Several privacy-preservation techniques focus on allowing different entities to train DL models without revealing secure data. The existing privacy-preserving techniques

already developed in this field, including encryption and differential privacy, are reviewed in this section.

#### 2.2.1. Privacy-Preservation through Encryption

Cryptographic methods could be used to conduct DL training and testing on encrypted data [9]. Such methods allow for privacy protection, but specialized techniques are needed to do useful statistical analysis on encrypted data [27]. To achieve PPDL, the most commonly used cryptographic methods are homomorphic encryption, secret sharing, and secure multi-party computation.

##### Homomorphic Encryption

Homomorphic encryption (HE) is a cryptographic technique that preserves the ability to process and produce data in encrypted forms as if it were unencrypted [28]. Due to its flexibility and its highly desirable outcomes, HE is a unique technique for encryption that can address both privacy and security concerns more easily than some other techniques. Rivest et al. [29] introduced this technique and its four crucial security procedures: key generation, encryption, decryption, and evaluation algorithm, as illustrated in Figure 2. With HE, if the user wants to request information from the cloud server, then that data is first encrypted and saved in the cloud before the user then sends request information to the cloud server. Without identifying the data's characteristics, the cloud server performs a predictive model on encrypted data and sends back that encrypted output to the requesting user. Only by utilizing a unique secret key can the user decrypt the obtained data, which is still encrypted; thus, the data's privacy and security are maintained [30].

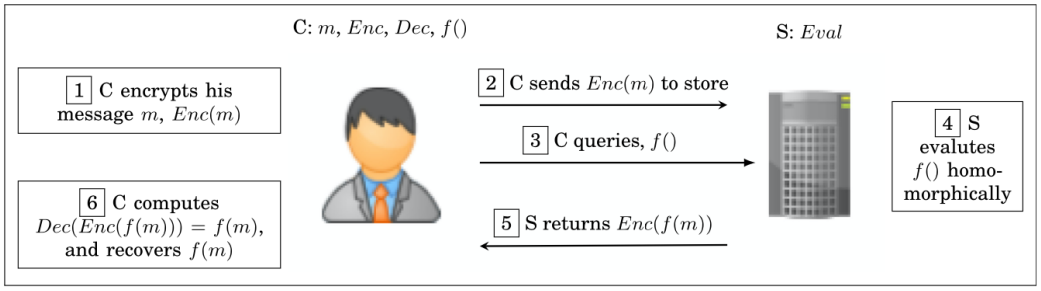


**Figure 2.** Homomorphic Encryption Technique [30].

Based on the number of mathematical processes performed on the encrypted data, HE techniques can be classified into three types: (1) Fully Homomorphic Encryption (FHE), (2) Somewhat Homomorphic Encryption (SHE), and (3) Partially Homomorphic Encryption (PHE). The three types are described in more detail below [30].

- Fully Homomorphic Encryption (FHE): enables the performance of various types of assessment operations on the encrypted data with unbounded range.
- Somewhat Homomorphic Encryption (SHE): all addition and multiplication operations are permissible in SHE, but with only a limited range.
- Partially Homomorphic Encryption (PHE): Only one form of mathematical operation on the encrypted data is permitted in the PHE scheme, such as a multiplication or addition procedure, with an unbounded range.

As previously mentioned, HE is a suite of four functions: key generation, encryption, decryption and evaluation. These are described in further detail below [31].

- Key generation: the client will generate a pair of public keys (PK) and secret keys (SK) to encrypt the plaintext (PT).
- Encryption: the client will encrypt the PT using the PK, and the ciphertext (CT) will be submitted to the server along with the PK.
- Decryption: the client will decrypt the generated evaluation using its SK, and the result will be obtained.

- Evaluation: the server has a ciphertext evaluation function (F) and executes it as in-demand using the PK.

The encryption mechanism in HE is homomorphic; that is, if it is possible to compute Enc(f (PT1, PT2)) from Enc(PT1) and Enc(PT2), where f can be: +, × and without utilizing the secret key (SK) [32]. Therefore, HE has two main properties: multiplicative HE and additive HE.

Multiplicative HE: an HE is multiplicative if it satisfies the following equation [16,32]:

$$Enc(PT1 \times PT2) = Enc(PT1) \times Enc(PT2). \quad (1)$$

Additive Homomorphic Encryption: an HE is additive if it satisfies the following equation [16,32]:

$$Enc(PT1 + PT2) = Enc(PT1) + Enc(PT2). \quad (2)$$

Secret Sharing

The secret sharing (SS) technique was proposed by Shamir and Blakley in 1979 [33]. This technique is the process of distributing a secret between several entities, each holding a share of the whole. Single shares are not useful independently, but the secret can be reconstructed if the shares are merged [9]. Figure A1 illustrates a secret share (S) held by several people (P). The secret can be retrieved if all participants work together [34].

Secure Multi-Party Computation

Secure Multi-Party computation (SMPC) is a set of techniques that allow two or more parties to separate data among themselves in order to perform collaborative computations. As illustrated in Figure A2, SMPC allows each party to acquire its corresponding output without obtaining any other information [35]. This technique can be used to conduct dataset analyses in an encrypted domain without perturbing or compromising the data [36].

#### 2.2.2. Differential Privacy

Differential privacy (DP) is a technique that enables the perturbation of a dataset to hide individual data while preserving the ability to do statistical analysis on that same dataset. This is a method of preserving the dataset's global statistical distribution while at the same time minimizing personally identifiable information [36]. In the context of our research, privacy is the property of both the output and the computation producing that output. DP can be used to resolve most user privacy concerns, since the user can guarantee that the analysis results will not reveal anything specific to him. In addition, if the user information is excluded from the analysis, user privacy is protected because the result of the analysis does not rely on the user-specific information, as illustrated in Figure A3 [37].

#### 2.2.3. Hybrid Techniques

This is a combination of two different PPDL techniques, and thus can provide greater privacy protection than either one used individually. Truex et al. proposed a hybrid approach as a means of overcoming existing PPDL disadvantages by combining both DP and SMPC to minimize noise injection without losing privacy as the number of parties rises [38].

#### 2.2.4. Comparison between PPDL Techniques

PPDL techniques such as HE, SS, SMPC, and DP can guarantee data processing without revealing any information about the input data. However, these techniques also come with limitations. Table 1 depicts advantages and limitations of the PPDL techniques already discussed in this study. Overall, HE is extremely useful for processing sensitive data on a cloud server. Although it is costly in terms of computational overhead, it is

also the most secure approach. SMPC does have certain advantage over HE, such as the possibility of obtaining input from different parties and higher practicality due to higher velocity and less overhead, but it also requires communication during computation where HE does not and thus retains higher degrees of protection. Meanwhile, DP has considerable benefits in terms of practicality. It is faster than both HE and SMPC since no computationally overhead encryption is required. In certain situations, though, DP accuracy can be lower because of the additional noise and the method has some limitations when it comes to security. All PPDL techniques have strengths and limitations, often in terms of greater privacy protection coming at the cost of lower practicality in terms of speed and ease of deployment. In our work, though, privacy is considered the most important aspect. To that end, this research utilizes HE, which offers higher data security and protection.

**Table 1.** Comparison of PPDL Techniques.

| Techniques | Advantages | Disadvantages |
| --- | --- | --- |
| HE | • Performs inference analysis on encrypted data<br>• Does not require interactivity between the data and model owners<br>• Perceives data privacy and security | Extensive computation overhead |
| SS | • Hides data from participants<br>• Offers protection from attacks<br>• Does not require all the shares to reconstruct the secret, but only the threshold | • Does not protect the private key from being stolen if an adversary is involved during the setup<br>• The private key is no longer secure once reconstructed |
| SMPC | • Eliminates the trade-off between data accessibility and data protection<br>• No authorized third parties can see the data | Requires interactivity between data and model owners |
| DP | Ensures the privacy of input data and the privacy of learning models | Requires broad noise to obtain significant privacy |

## 3. Related Works

This section reviews relevant research concerning PPDL. It also provides a comparison between previous research works and the current research.

Phong et al. [39] have proposed an HE-based approach. In their work, the authors considered the main issues of the Shokri and Shmatikov system, which tends to leaking users' local data to the cloud server as many users perform neural network-based DL. Phong et al. address this by building an enhanced DL system using additive HE. This system can prevent information leakage while still preserving accuracy. The results of Phong et al.'s work demonstrate that the system does not leak any user information to an honest-but-curious cloud service. Meanwhile, the use of HE adds a reasonable amount of overhead to the DL system.

Another solution based on data encryption using FHE was proposed by Vizitiu et al. [40]. Here, the authors proposed to encrypt the input data and send it to the server to predict their results. Their approach took advantage of the MORE framework, which does not reveal patient records. The model performance was evaluated using medical imaging and the MINST dataset. Compared to the plain form, the experiment results indicated that Vizitiu et al.'s proposed solution achieves similar accuracy over the clinical dataset and approximately identical precision over the MNIST dataset. As for performance, the encrypted approach improved latency relative to unencrypted approaches, while as for

privacy, the authors noted that while the MORE framework gives a certain degree of privacy, it is still susceptible to chosen plaintext attacks.

Wang and Chang [41] also developed an approach based on DP techniques. The authors considered a two-party image classification issue where the data owners retain the images, and unreliable data users train the ML model with any of these images as input. Wang and Chang aimed to preserve data usability on image classification while at the same time maintaining data privacy. The authors proposed to use a randomized reply to perturb the image locally, which satisfies local DP. They also introduced DCA-Conv, a supervised image feature extractor, to manage the trade-off between usability and privacy. The results they achieved demonstrated that DCA-Conv can achieve a high degree of data usability while still maintaining privacy.

Abadi et al. [42] have proposed another work based on DP. Here, the authors considered the training models problem, which may expose private and sensitive information in image datasets. Within the framework of DP, Abadi et al. introduced techniques to enhance the computational performance of DP, including computing gradients algorithms, dividing works into smaller batches, and applying differentially private principal projection at the input nodes. This experiment was conducted on a ML TensorFlow framework, and the results indicated that the training model can achieve relatively high efficiency, privacy protection, and model consistency.

Huang et al. [33] have proposed a framework-based on the SS technique. These authors considered mobile sensing data protection and response time on cloud computing. Using an encryption-based secret exchange strategy, they developed a privacy-preserving CNN for feature extraction. Instead of cloud servers, the massive computational process was moved to edge servers in order to reduce delays between the cloud server and the mobile device. The results they achieved demonstrated the safety, efficacy, and reliability of their scheme based on theoretical analysis and empirical studies.

Another work based on the SS technique has been proposed by Ma et al., who considered the privacy of facial image data on cloud servers [43]. They proposed an AdaBoost-based system for face recognition (POR), which was designed to protect users' facial characteristics and the service providers' privacy based on additive secret sharing techniques. This system consisted of two edge servers assigned to the complex POR computing operation. The authors enhanced the additive secret sharing-based technique features by increasing the efficient input domain. Through theoretical analysis, they demonstrated the consistency and security of the technique. The results of their experiment indicated a decrease in computational error as compared to the current differential privacy-based framework.

Xia et al. [44] have proposed a scheme inspired by additive secret sharing techniques. They considered the encrypted image problem, which restricted the effectiveness of the image usage. They also proposed a set of additive protected computation protocols on numbers and equations with higher efficiency. With the assistance of their protocols, including the total operation of image classification in the unencrypted domain, they extracted CNN characteristics, reduced the dimension of characteristics, and generated the index safely. They also evaluated the execution of the suggested scheme in terms encryption reliability, recovery precision, and recovery efficiency using the Corel image dataset. The experiment's results showed the higher reliability and efficiency of this new scheme.

Erkin et al. [45] have proposed a framework based on the SMPC technique. The authors considered a case in which one party provides a facial image while the other party has access to a facial database, then introduced an extreme privacy-enhanced facial recognition framework that effectively protects both the input data and the server output that is running the matching function. The experiment's results proved that the privacy-preserving framework is accurate, and also that it is possible to perform the protocol on modern hardware platforms.

According to the literature summarized here, we can conclude that many studies have used different PPDL techniques as well as diverse image datasets. However, as shown in Table 2, none of these works discuss PPDL in satellite images data.

**Table 2.** Comparison of Relevant Studies about PPML Techniques.

| Ref. | Domain of Application | DL Models | PPDL Techniques | Dataset |
|---|---|---|---|---|
| Phong et al. [39] | • Grayscale images<br>• Labeled street view images | Deep neural networks (DNNs) | Additively homomorphic encryption | • MNIST<br>• SVHN |
| Vizitiu et al. [40] | • Grayscale<br>• Medical imaging | Deep neural networks (DNNs) | Additively homomorphic encryption | • MNIST<br>•X-ray coronary angiographies |
| Wang & Chang. [41] | • Frontal facial images<br>• Grayscale images<br>• Grayscale fashion images | • Naive Bayes<br>• K-nearest neighbors (KNN) | Local differential privacy | • Yale Face B<br>• MNIST<br>• Fashion-MNIST |
| Abadi et al. [42] | • Grayscale images<br>• Color images | Neural networks | Differential privacy | • MNIST<br>• CIFAR-10 |
| Huang et al. [33] | • Color images<br>• Grayscale images | Convolutional neural network (CNN) | Additive secret-sharing technique | • CIFAR-10<br>• MNIST |
| Ma et al. [43] | Face images | Neural networks | Additive secret-sharing technique | FERET Database |
| Xia et al. [44] | Concept images such as castle | Convolutional neural network (CNN) | Additive secret-sharing technique | • Corel-1k<br>• Corel-10k |
| Guajardo et al. [45] | Face images | GNU GMP library | Secure multi-party computation | ORL Database |
| This research | Vegetation, road, bare soil and urban images | Convolutional neural network (CNN) | Partially homomorphic encryption | Satellite Dataset |

All previous studies do demonstrate the ability to maintain data privacy while preserving data usability when processing DL models, though. Thus, based on those works, our research will be evaluated on a proposed DL model, namely CNN, with a new dataset of real-world satellite images. Additionally, several experiments will be conducted using state-of-the-art DL transfer techniques to evaluate the performance of the method we propose.

## 4. Proposed Method

This section will describe the proposed method for PPDL in the specific case of satellite image classification. The proposed workflow, which is based on PHE and CNN, is illustrated in Figure 3. Pre-processing, the data is encrypted on the client side with a public key and cannot be decrypted without knowing the private key. Therefore, only the encrypted data is accessible to the CNN-based model (cipher-images). Direct training on cipher-images data is achieved through the partially homographic feature of the Paillier encryption scheme; more details about PHE and Paillier scheme are presented in Section 4.1. After the training process is completed, the model will undergo the testing phase with new cipher-images data, which are encrypted using the same public key as the training process: the proposed CNN-based model architecture is presented in Section 4.2. Finally, since the cloud server operates directly on cipher-images data, data privacy is maintained during both training and testing. As a result, satellite image data processing is carried out securely and unauthorized parties cannot decipher the data.

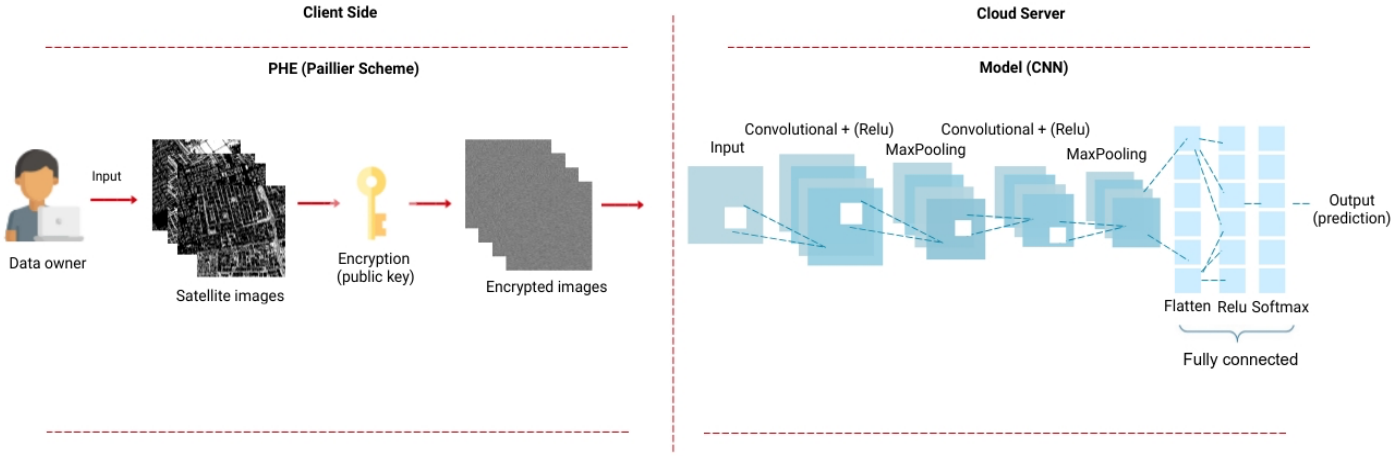


**Figure 3.** The Proposed Workflow.

### *4.1. The Proposed Encryption Method for Satellite Images*

As previously mentioned in Section 2.2.1, the HE technique has three sub-types that can be used for preserving data privacy. In this paper, the technique used for encrypting the data is PHE, which we selected because it does not require as much overhead required for executing computations. However, FHE requires a lattice-based cryptosystem, and it is not a realistic scheme either conceptually or technically. Especially in terms of computation, the bootstrapping section (which is the intermediate refreshing method of a processed ciphertext), is expensive and considerably more complicated than most other options. On the other hand, SHE allows for the performance of a limited number of sequential ciphertext multiplication and addition operations while PHE allows for an unbounded number of times [16,46].

The phases of the PHE technique will be described in Section 4.1.1.

#### 4.1.1. Partially Homomorphic Encryption Schemes

There are many PHE schemes, each of which enhances a particular aspect of PHE. The first achievement of the public key (PK) cryptosystem is Rivest-Shamir-Adleman (RSA). Public-key encryption is asymmetric key encryption, which is a type of algorithm that demands two different keys, one being private and unique to certain users while the other is public [47]. The RSA scheme was established by Rivest, Shamir and Adleman (1977) as the first public-key cryptosystem for asymmetric-key encryption with the homomorphic property. It also defined the properties of multiplicative HE. However, strong security principles are not necessarily fulfilled here because, in order to accomplish semantic security, RSA must pad a message with random bits before encryption, which results in losing the homomorphic property [48]. Since RSA does not fulfill strong security requirements, then, this research implementation is based on the Paillier scheme. The cryptosystem of Paillier is created by Pascal Paillier (1999) as a probabilistic asymmetric algorithm for public-key cryptography [47]. The Paillier scheme has a homomorphic property, unlike RSA, and it is limited to addition. So, the property of additive HE is realized by the Paillier cryptosystem. Electronic voting is an implementation of additive HE. Each vote will be encrypted, and only the total will be decrypted [32].

##### Paillier Scheme

The encryption scheme of Paillier is composed of three phases: key generation, encryption, and decryption, as depicted in Algorithms 1, 2, and 3 respectively [32,47].

The decryption of a Paillier scheme requires a cipher-text that is generated by the encryption process. The public key for encryption is (n, g) and the private key for decryption is ($\lambda$, $\mu$ ).

**Algorithm 1:** Key generation (p, q)

**Input:** Generate two unique prime numbers, p, and q and confirm that: $gcd(p \times q, (p-1)(q-1)) = 1$, where $gcd$ represents the greatest common divisor.

**Output:** (pk, sk)

**if** *length(p)==length(q).* **then**

  Compute $n = p \times q, \lambda = lcm(p-1, q-1)$, where $lcm$ represents the least common multiple.

Choose a random integer $g \in Z^*_{n^2}$( between 1 and $n^2$ )

Define the function: $L(x) = ((x-1)/n)$.

Verify the existence of the following modular multiplicative inverse to ensure that $n$ divides g's order: $\mu = L(g^{\lambda} \mod (n^2))^{-1} \mod (n)$

**Algorithm 2:** Encryption (m, pk)

**Input:** Message to encrypt where $m \in Z_n$

**Output:** $c \in Z_{n2}$

Choose a random integer $r \in Z^*_{n^2}$( between 1 and $n^2$)

Compute the ciphertext as: $c = (g^m \times r^n) \mod (n^2)$

**Algorithm 3:** Decryption

**Input:** c $\in Z_{n^2}$

**Output:** $m \in Z_{n^2}$

Calculate the plaintext message as: $m = L(c^{\lambda} \mod (n^2)) \times \mu \mod (n)$

The Paillier cryptosystem is characterized by the following homomorphic properties [47]:

- Addition of plaintexts: the result of multiplying two ciphertexts would decrypt the sum of their respective plaintexts, as described in the following formula:

$$D_{priv}(E_{pub}(m1)E_{pub}(m2) \mod (n^2) = m1 + m2 \mod (n). \quad (3)$$

- The ciphertext results through raising g to the plaintext would decrypt to the sum of their respective plaintexts, as described in the following formula:

$$D_{priv}(E_{pub}(m1)g^{m2} \mod (n^2) = m1 + m2 \mod (n). \quad (4)$$

### *4.2. Proposed Convolutional Neural Network (CNN)*

As previously discussed, the main goal of this study is to ensure the privacy of satellite images when using public DL methods. Thus we propose to develop a custom CNN model and test its performance on satellite images encrypted using the proposed technique. The proposed CNN model is composed of the following layers: three convolution layers, three polling layers, a dropout layer, a flattening layer, two fully connected layers, and an activation function (ReLu, Softmax). Table 3 presents the proposed CNN architecture.

Here the CNN was trained on encrypted data and passed through a stack of three convolution layers with 32 filters for the first and the second convolution and then 64 layers for the third convolution. The max-pooling layers that subsample the image by filters of 2 × 2 were placed after each convolution layer with 32, 32 and 64 filters, sequentially. The last two fully connected layers were loaded with 64 and 4 nodes sequentially, and the ReLu activation functions were utilized throughout the network, except for the last layer. For regularization, the dropout layer was utilized after the last pooling layer to prevent overfitting. A flattening layer was placed after the dropout layer and before the first fully connected layers in order to adjust the whole pooled feature map into one column. A

Softmax activation function was placed in the last fully connected layer (output layer) to provide a class prediction.

**Table 3.** CNN Architecture.

| No. | Layers | Output Shape | Parameters | Dropout Rate |
|---|---|---|---|---|
| 1 | Input | 128 × 128 × 3 | — | — |
| 2 | Convolutional | 128 × 128 × 32 | 896 | — |
| 3 | Activation (ReLu) | — | — | — |
| 4 | MaxPooling | 64 × 64 × 32 | — | — |
| 5 | Convolutional | 64 × 64 × 32 | 9248 | — |
| 6 | Activation (ReLu) | — | — | — |
| 7 | MaxPooling | 32 × 32 × 32 | — | — |
| 8 | Convolutional | 32 × 32 × 64 | 18,496 | — |
| 9 | Activation (ReLu) | — | — | — |
| 10 | MaxPooling | 16 × 16 × 64 | — | — |
| 11 | Dropout | 16 × 16 × 64 | — | 0.4 |
| 12 | Flatten | 16,384 | — | — |
| 13 | Fully Connected | 64 | 1,048,640 | — |
| 14 | Activation (ReLu) | — | — | — |
| 15 | Fully Connected | 4 | — | — |
| 16 | Activation (softmax) | — | 260 | — |

#### 4.2.1. Data Augmentation

Data augmentation is a technique for increasing the amount of data available for training the proposed CNN model without actually acquiring new data [49]. This technique is used to expand the dataset into a larger one more appropriate for DL model training. There are various strategies used for data augmentation, including rotation, zoom, horizontal, and vertical shift. These techniques assist in enhancing the efficiency of CNNs [50].

Different data augmentation strategies have been utilized in this research. For instance, we applied a 90-degree rotation range, a zoom and shear range of 20%, a brightness scale between 0.2 to 1.0 and a shift range of 20% in height and width. Finally, a horizontal flip and vertical flip have also been applied. Table 4 shows the detailed parameters of our data augmentation processes

**Table 4.** Augmentation Parameters.

| Augmentation | Parameter |
|---|---|
| Rotation | 90° |
| Zoom | 20% |
| Shear | 20% |
| Horizontal shift | 20% |
| Vertical shift | 20% |
| Brightness | [0.2, 1.0] |
| Horizontal flip | Yes |
| Vertical flip | Yes |

## 5. Experiments

This section describes the dataset we utilized as well as the environment in which our experiment took place. It will also present image encryption results, clarify the encryption schema's efficiency, and examine the CNN model's performance with both encrypted and plain data in order to evaluate the efficiency of the proposed encryption method. Experiments with pre-trained models will also be presented as another means of evaluating the efficiency of the proposed encryption method.

### 5.1. Dataset Description

In this study, experiments are conducted using satellite images produced by the French Satellite pour l'Observation de la Terre (SPOT) satellite. These satellite images were acquired using Spot 6 and Spot 7 with a spatial resolution of 2.5 m. They have been corrected both radiometrically and geometrically using ortho-rectification and spatial registration with sub-pixel accuracy and through close comparisons against a global reference system. Four land-cover types are identified in these regions, namely: urban, bare soil, vegetation, and road.

The dataset used in this paper is comprised of 37,774 images, as further illustrated in Table 5.

**Table 5.** Land cover types with number of samples.

| Land Cover Type | No. of Samples |
|---|---|
| Urban | 9730 |
| Vegetation | 8440 |
| Bare soil | 9124 |
| Road | 10,480 |

The considered dataset is further divided into three datasets, resulting in 22,666 images for training the model (training set), 7554 images for validating the trained model (validation set), and 7554 images for assessing the model performance (testing set).

To obtain this dataset, a semantic segmentation is conducted using our previous works [51–53]. The four classes—namely, urban, bare soil, vegetation and road—are extracted from the satellite images, meaning that the resulting images each contain both the real value of pixels of the extracted class and zero for the values of the other classes. Then, each image containing a given class is divided into non-overlapping blocks of $256 \times 256$ pixels and saved into folders, each with the name of the corresponding class. A sample of this dataset is depicted in Figure 4, wherein the white represents a given land cover class and the black represents the values of other classes.

### 5.2. Experimental Set-Up

The hardware configuration and software used for the encryption process are:

- Graphics processing unit: Intel Core i5-3210M (2.50 GHZ).
- Memory: 4 GB.
- Operating system: Windows 10 Professional.
- Visual Studio Code with Python 3.9 extension.

These CNN experiments are conducted using the Google Colab repository, which allowed us to execute Python code through the browser and also provided access to NVIDIA graphics processing unit (GPUs). The libraries used to conduct these experiments are the Keras DL library and TensorFlow backend, a DL platform. The proposed model was trained using Stochastic Gradient Descent (SGD) optimizer with a learning rate of 0.001, 32 batch size and 100 epochs.

### 5.3. Experimental Results

#### 5.3.1. Images Encryption

This section presents the results of image encryption we obtained using the Paillier scheme, or PHE. In addition, it also evaluates the efficiency of the encryption in terms of its reliability.


**Figure 4.** Sample of the dataset.

The Paillier encryption scheme enables researchers to train and test the CNN model without visual information. From this, Figure 5 shows the results of image encryption for four samples representing the four-land cover classes: namely urban, vegetation, bare soil, and road. The original images were encrypted using the public key and decrypted using the private key. The measurement of encryption efficiency and security is a significant feature of image encryption scheme. Visual observation is appropriate in certain situations, but it does not indicate the amount of information hidden. Therefore, the correlation coefficient measurement has been utilized in order to evaluate the Paillier encryption scheme's efficiency.

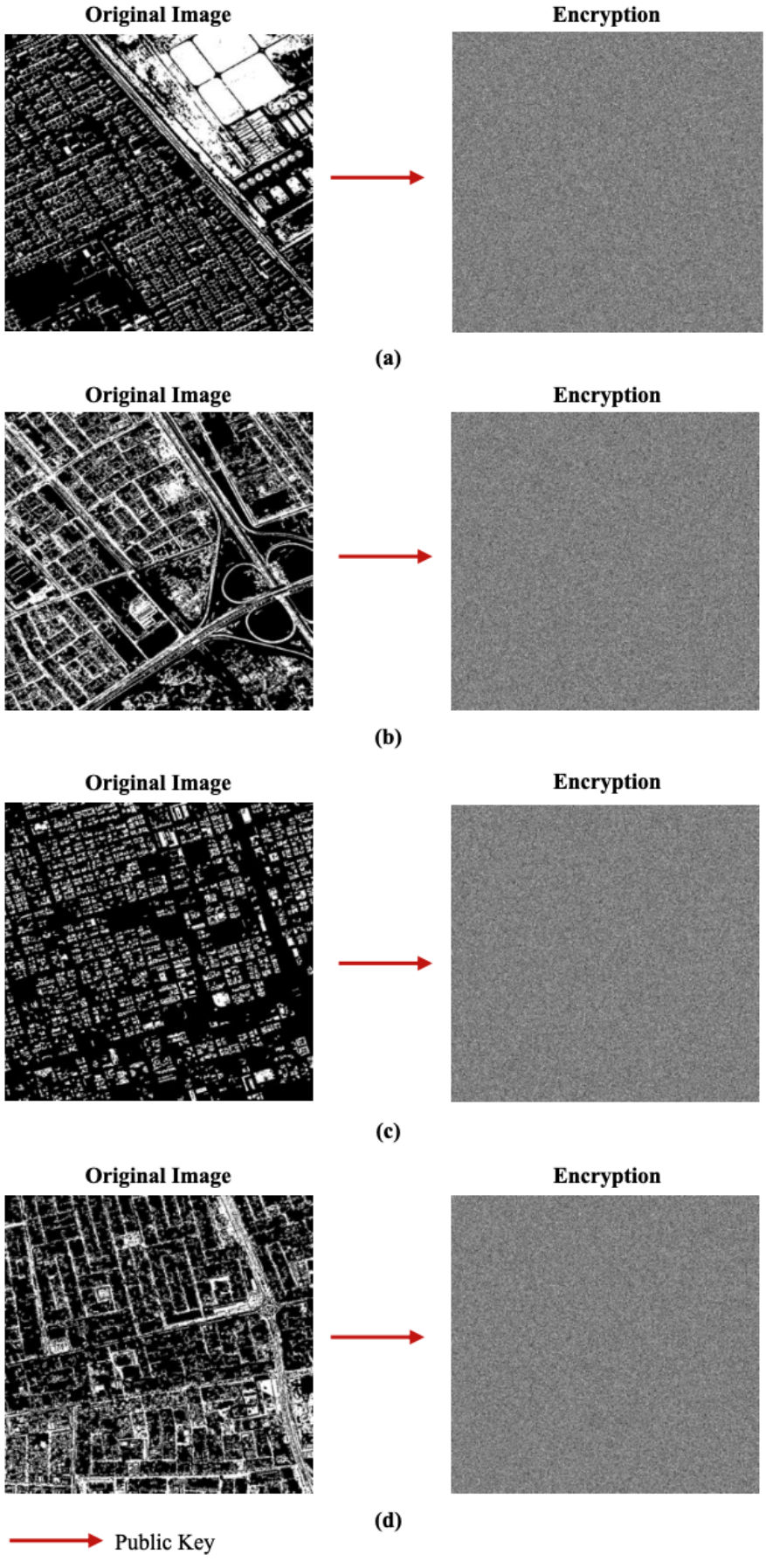


**Figure 5.** (**a**) Bare soil class; (**b**) Road class; (**c**) Urban class; (**d**) Vegetation class.

Security Evaluation

In this section, different security measurement has been utilized to evaluate the Paillier encryption scheme's efficiency.

The correlation coefficient (CC) is used to measure the degree to which two variables are related. The CC value range is between −1.0 and 1.0. A negative correlation is represented by a correlation of −1.0, while a positive correlation is represented by a correlation of 1.0. The correlation coefficient is calculated using the following equation [54–57]:

$$CC = Z_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) / \sqrt{Z_{i=1}^{N}(x_i - \bar{x})^2 (y_i - \bar{y}^2)}, \quad (5)$$

where $x$ represents a plain-image, and y represents an encrypted image, and $\bar{x}$, $\bar{y}$ are the mean of the plain-image and encrypted image, respectively.

An image cryptosystem is considered efficient if the encryption scheme covers all features of a plaintext image, while the encrypted image is also completely unpredictable and strongly uncorrelated. Therefore, the encryption scheme's efficiency can be determined if the correlation coefficient between the encrypted image and plain-image close to zero or −1.

The CC result of the plain image and its corresponding encrypted image from the satellite image dataset is −0.0041, which is a negative relationship. This result indicates the efficiency of this encryption scheme where there are no visual features identified in the encrypted image.

Moreover, a number of other security parameters are evaluated, and the results are shown in Table 6. From Table 6, it is clear that the encrypted images are secure and an intruder cannot get any idea from the encrypted information. Ideally, entropy should be close to 8, and we can see from Table 6 that the value of entropy for ciphertext is 7.9596, much higher than 3.12 for plaintext image. The higher value of contrast (10.57) indicates a secure image. Low energy and homogeneity values show that an encryption scheme is robust and highly secure. In our case, energy and homogeneity values are 0.0156 and 0.3884, respectively. These values are lower than for plain images. For MSE key sensitivity and unified average change intensity, higher values are required. From Table 6, higher values are evident. Additionally, the lower value of peak signal to noise ratio and structural similarity index highlighted the security of the encrypted image.

**Table 6.** Security evaluation.

| Parameter | Plain-Images | Cipher-Images |
|---|---|---|
| Entropy [55] | 3.1206 | 7.9596 |
| Contrast [58] | 5.3191 | 10.5767 |
| Energy [55] | 0.5363 | 0.0156 |
| Homogeneity [55] | 0.9044 | 0.3884 |
| Mean Square Error (MSE) [54,59] | - | $2.1236 \times 10^4$ |
| Peak signal to noise ratio [55] | - | 4.8601 |
| Key sensitivity [60] | - | 99.5725 |
| Unified average change intensity [58] | - | 33.66 |
| Structural similarity index [61] | - | 0.0018 |

5.3.2. CNN Performance

This section will present the performance of the custom CNN proposed in this study over both encrypted and plain data.

Data Augmentation

The first step in presenting the performance of this CNN model is to describe the results of our data augmentation processes. This latter operation is ensured by using the ImageDataGenerator function from the Keras deep learning library. The results obtained

from our selected data augmentation techniques, including rotation, zoom, shear, height and width shift, brightness, and horizontal and vertical flip, are illustrated in Figure 6. A sample image from the satellite dataset was used to demonstrate these results. As shown in Figure 6a, the rotation results in pixels out of the images frame, leaving blank areas with no pixel details, while zoom augmentation in Figure 6b results in making the images' objects larger. As shown in Figure 6c, shearing has been used to shift one part of the images, resulting in a parallelogram shape. Figure 6d,e shows that all images' pixels have been moved in one direction, either horizontally or vertically, while maintaining the original images' dimensions. Furthermore, as shown in Figure 6f, images have been randomly darkened or brightened to further augment the dataset. The images' brightness enables the CNN models to generalize through trained images under varying lighting conditions. Data augmentation retains the features that are essential for predictions. As shown in Figure 6g,h, the pixels are completely rearranged when flipping the images horizontally and vertically, but the features are retained.

Model Performance

The goal of this section is to evaluate the performance of the proposed PPDL process in the context of encrypting satellite images using two criteria: namely, accuracy and applicability. Therefore, the data-driven model's efficiency and outcome were analyzed for by implementing the model on unencrypted (plain images) and encrypted (cipher images) data. Furthermore, quantitative analysis is conducted using accuracy metrics to evaluate the performance of the proposed CNN model for satellite image classification. Accuracy is defined as the total number of correct classifications, either truly positive (TP) or truly negative (TN), as compared to the total number of images in the dataset, as given in Equation (6). Training accuracy refers to a model's accuracy on the training data it was built on, while test accuracy refers to a model's accuracy on test data it is encountering after training.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)100, \quad (6)$$

where TP is the number of instances where the prediction was correct and FN is the number of instances where the prediction was incorrect. The number of correctly-predicted negative instances is depicted as TN, while the number of wrongly predicted negative instances is known as false positives or FP [50].

The proposed CNN model has been trained using both unencrypted (plain-images) and encrypted (cipher-images) data. Consequently, to demonstrate the network's capacity to learn from encrypted data, we also evaluated the CNN model's capacity to preserve performance by conducting and comparing results for two cases.

- Case 1: Training and testing the proposed CNN model with unencrypted data. The results from training the model with plain data achieved 96.91% and 96.84% accuracy for training and validation, respectively, on the training dataset, as well as 93.84% accuracy for the testing dataset. Figure 7 demonstrates the accuracy results for training and validating the CNN model over plain satellite images.
- Case 2: Training and testing the proposed CNN model over encrypted data. The results of training the model on encrypted data achieved 94.04% and 94.30% accuracy for training and validation, respectively, on the training dataset, as well as 90.92% accuracy for the testing dataset. Figure 8 demonstrates the accuracy results for training and validating the CNN model on encrypted satellite images.

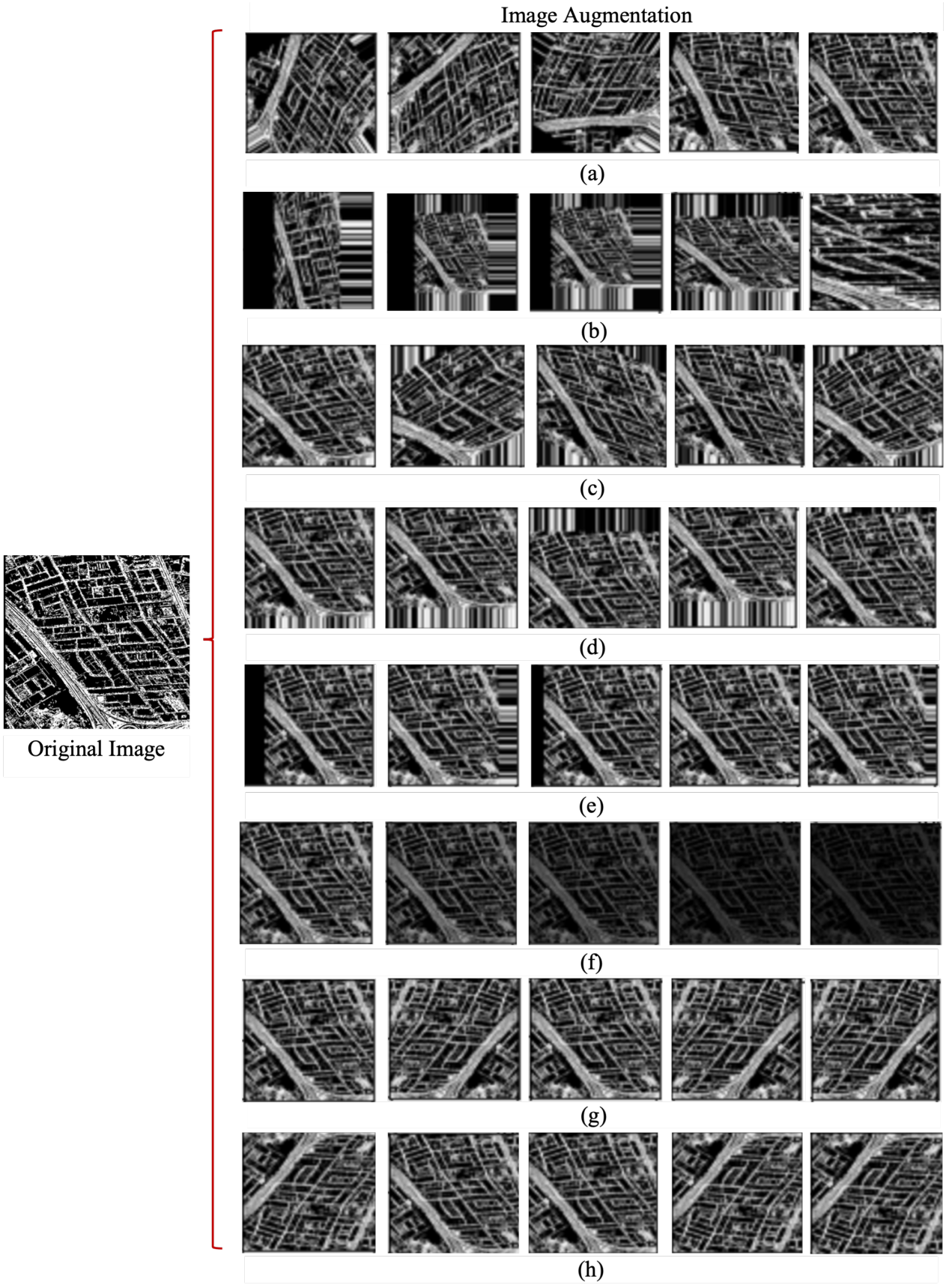


**Figure 6.** Sample image data augmentation. (**a**) Rotation results; (**b**) Zoom results; (**c**) Shearing results; (**d**) Horizontal shift results; (**e**) Vertical shift results; (**f**) brightness results; (**g**) Horizontal flip results; (**h**) Vertical flip results.

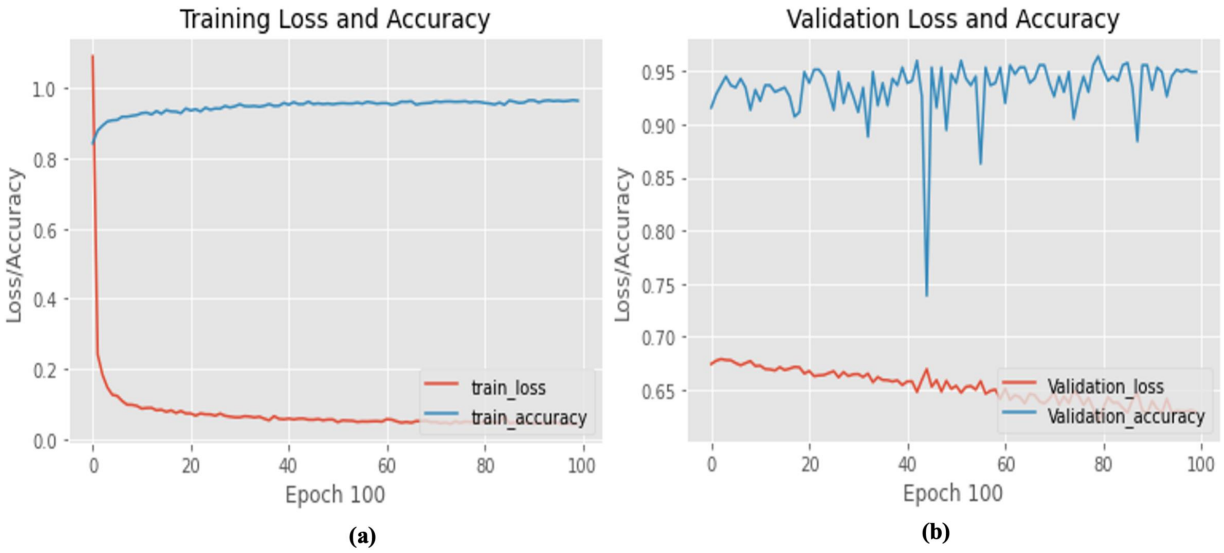


**Figure 7.** Model accuracy over plain data. (**a**) Training accuracy; (**b**) Validation accuracy.

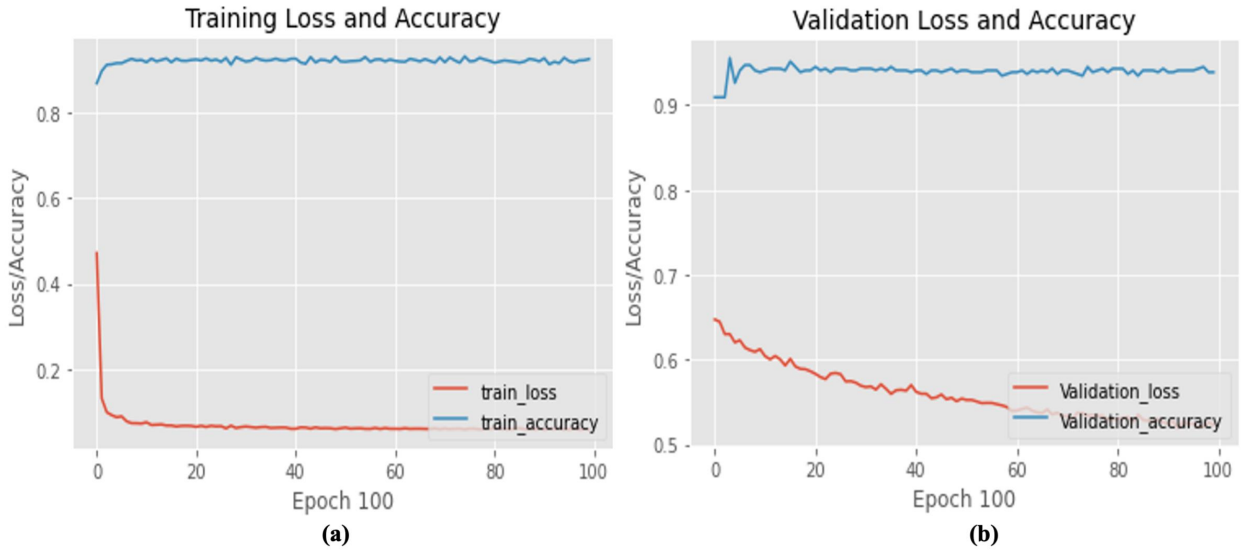


**Figure 8.** Model accuracy over encrypted data. (**a**) Training accuracy; (**b**) Validation accuracy.

The results obtained from the training and validation of the proposed CNN model with plain data reached about 2% higher accuracy over the encrypted data, as presented in Table 7. Moreover, the model's performance with plain data reached about 3% higher accuracy over the encrypted data, which indicates a low tradeoff between data utility and privacy. Therefore, the approach that we propose here provides maximum privacy protection while also maintaining data utility. The loss of data utility due to data encryption is measured by comparing the model's performance with both the encrypted data and the plain data. To allow such a comparison, the same CNN model was used for both encrypted data and plain data.

**Table 7.** The proposed model's performance over plain and encrypted data.

| Type of Dataset | Training Accuracy | Validation Accuracy | Prediction Accuracy |
|---|---|---|---|
| Plain | 96.91% | 96.84% | 93.84% |
| Encrypted | 94.04% | 94.30% | 90.92% |

### *5.4. Evaluation of the Privacy-Preserving*

In order to evaluate the performance of the proposed privacy-preservation scheme, we also conducted additional experiments using pre-trained models. Here, four common pre-trained models (namely VGG16, Xception, ResNet50 and DenseNet121) were considered. Each was modified by fine-tuning the last layers, whereas the models' previous layers were preserved frozen. We used the same batch size, initial learning rate, number of epochs and input image resolution for all models, as depicted in Table 8.

**Table 8.** Characteristics of the four considered pre-trained CNN models considered here.

| CNN Models | Batch Size | Initial Learning Rate | Epochs | Input Image Resolution | Parameter |
|---|---|---|---|---|---|
| VGG16 | 32 | 0.001 | 100 | 128 × 128 | 14.847.044 |
| Xception | 32 | 0.001 | 100 | 128 × 128 | 29.916.844 |
| ResNet50 | 32 | 0.001 | 100 | 128 × 128 | 31.924.484 |
| DenseNet121 | 32 | 0.001 | 100 | 128 × 128 | 8.002.756 |

A comparison of the CNN models' classification accuracy, as obtained when trained on encrypted images, is presented in Table 9. The results show good classification accuracy of all the pre-trained CNN models. As we note, DenseNet121 achieved the best accuracy among the four models, with an accuracy of 93.36% for training, 90.93% for validation and 90.5% for testing.

**Table 9.** Comparison of classification accuracy between pre-trained CNN models using encrypted images.

| CNN Models | Training Accuracy | Validation Accuracy | Prediction Accuracy |
|---|---|---|---|
| VGG16 | 89.52% | 90% | 89% |
| Xception | 92.65% | 90.93% | 89.17% |
| ResNet50 | 93.36% | 90.93% | 89.8% |
| DenseNet121 | 93.36% | 90.93% | 90.5% |
| The proposed model | 94.04% | 94.30% | 90.92% |

Figure 9 illustrates the training accuracy of the five CNN models: both the one we propose and the four pre-trained ones. We note that our proposed model provides better training accuracy compared to all four pre-trained CNN models.

Execution Time

All runtimes reported in this section were measured on the Google Colab repository with a CPU running at 2.30GHz. Table 10 presents a detail of the runtime for each CNN model. The training runtime for VGG16, Xception, and ResNet50 is 22.27, 21.26, and 30.8 min, respectively. Additionally, the prediction runtime for these models is 0.803, 0.827, and 1.835 s, respectively. The training runtime for DenceNet121 and the proposed model is 18.34 and 16.43 min, and the prediction runtime is 4.134 and 0.976 s, respectively. Accordingly, the computation overhead varies from one model to another. However, PHE data are significantly fast during both training and prediction and therefore the computation overhead of the proposed encryption schema is low.

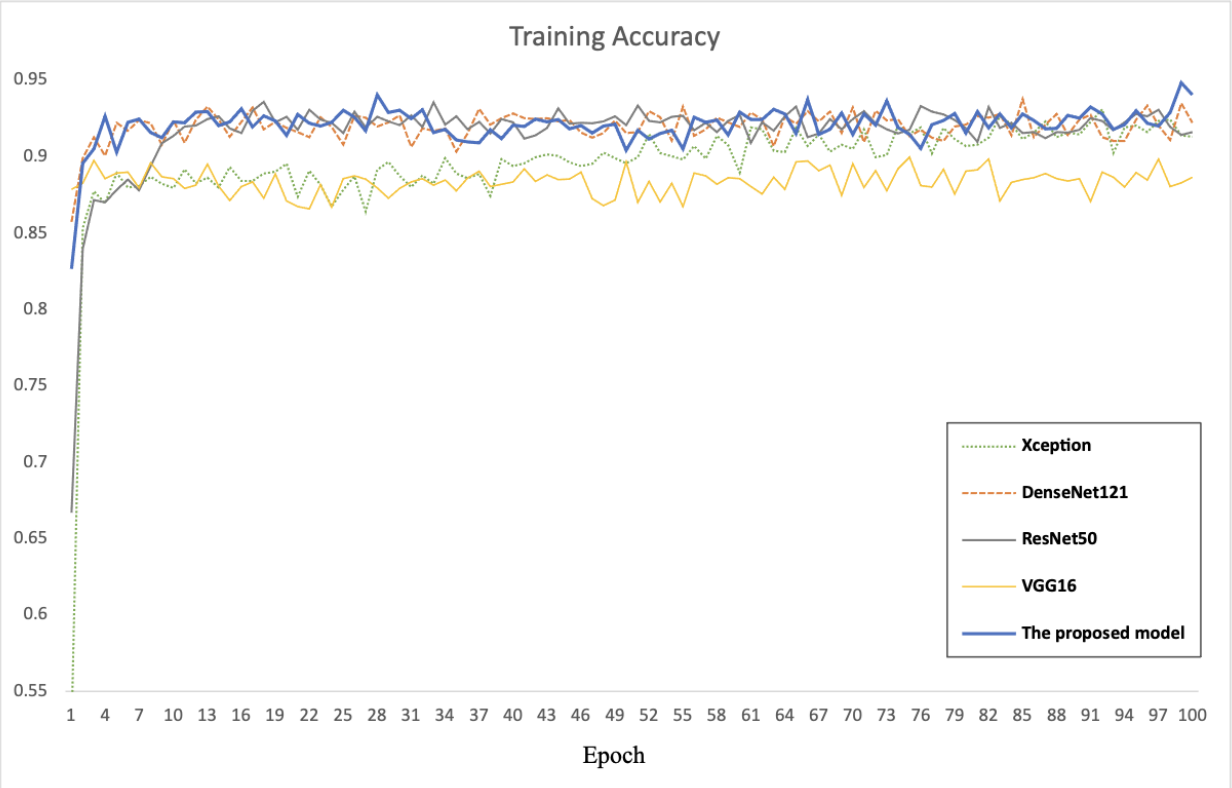


**Figure 9.** Training accuracy of different CNN models.

**Table 10.** CNN models' run-time.

| CNN Models | Training Run-Time | Prediction Run-Time |
|---|---|---|
| VGG16 | 22 min 27 s | 0.803 s |
| Xception | 21 min 26 s | 0.827 s |
| ResNet50 | 30 min 8 s | 1.835 s |
| DenseNet121 | 18 min 34 s | 4.134 s |
| The proposed model | 16 min 43 s | 0.976 s |

### *5.5. Discussion*

Recent years have seen increasing concerns about protecting the privacy of confidential information when processing data using models. This leads to the need for cryptographic techniques to solve privacy concerns in data-driven models. Several PPDL techniques have been proposed in the literature to solve these concerns. This research is, to the best of our knowledge, the first work that investigates PPDL for satellite image classification.

In this study, we have proposed a PHE-based Paillier scheme as a means of preserving data privacy. This PHE scheme enables several operations to be performed directly on the encrypted data (cipher-images) without the need to access the unencrypted data (plain-images). Furthermore, the proposed encryption scheme offers high security, as measured with different security parameters. Because the data are encrypted with a PHE scheme, the images contain no identifiable information and thus do not reveal nothing sensitive. The capability of the Paillier encryption scheme in DL models was further demonstrated by tackling satellite image classification.

The performance of CNN models is highly dependent on the presence of a large dataset, which often constitutes a limitation in this research area since only comparatively small datasets are available. For example, the satellite dataset used in this research comprises 37,774 images divided into four classes. Therefore, this research has utilized data augmentation techniques to help increase the satellite dataset variety and thus improve

the CNN model's performance. An important observation here is that the PHE technique used in this study has ensured the privacy of data without compromising the classification accuracy of DL models. The results we have achieved show that the classification accuracy of different DL models is relatively close for both encrypted and plain data. Therefore, the proposed encryption scheme ensures good classification of satellite images while preserving the privacy and security of data included within these images.

## 6. Conclusions and Future Works

DL has become the core technology in many forms of data analysis. Therefore, various security threats and corresponding defensive PPDL techniques have attracted much attention in both the research community and in global interests such as military operations and business. With such an increased interest in processing satellite image data, there also comes a great demand for preserving privacy when using public DL technique for processing satellite images.

In this study, we proposed a PHE-based technique to protect sensitive information in satellite images when applying public DL models. To the best of our knowledge, this study constitutes the first research work that focused on PPDL as applied to satellite images. The encryption scheme developed in this research enables both the security of data and good classification accuracy. To evaluate the encrypted scheme's efficiency, we conducted several experiments on both our custom model and several pre-trained CNN models. The results show a high level of efficiency for both the plain and encrypted data. In general, all models achieved good results in satellite image classification, but ours was the best by a slight margin. However, although our proposed PHE encryption scheme is efficient and provides good classification accuracy while preserving sensitive information within satellite images, several possible extensions can be considered in future work. We plan to apply the proposed approach to other datasets and test their performance. In addition, applying other PPDL techniques such as SS, SMPC, and DP, and evaluating their efficiency with the proposed method will be considered as a future perspective of this work. Moreover, we plan to explore the possibility of developing a hybrid technique that integrates more than one PPDL technique and then evaluate its performance, particularly in terms of privacy and classification accuracy.

**Author Contributions:** Conceptualization, M.A., W.B. and J.A.; methodology, M.A. and W.B.; software, M.A. and W.B.; validation, M.A., W.B. and J.A.; formal analysis, M.A., W.B. and A.K.; investigation, M.A. and M.D.; resources, M.A., W.B. and J.A.; data curation, W.B.; writing—original draft preparation, M.A. and W.B.; writing—review and editing, M.A., W.B., J.A., A.K. and M.D.; visualization, M.A.; supervision, W.B.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data will be available upon request to the corresponding author.

**Acknowledgments:** The authors would like to thank King Abdul-Aziz City for Science and Technology (KACST) in Riyadh, Saudi Arabia for providing satellite data used in this study. Also, the authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

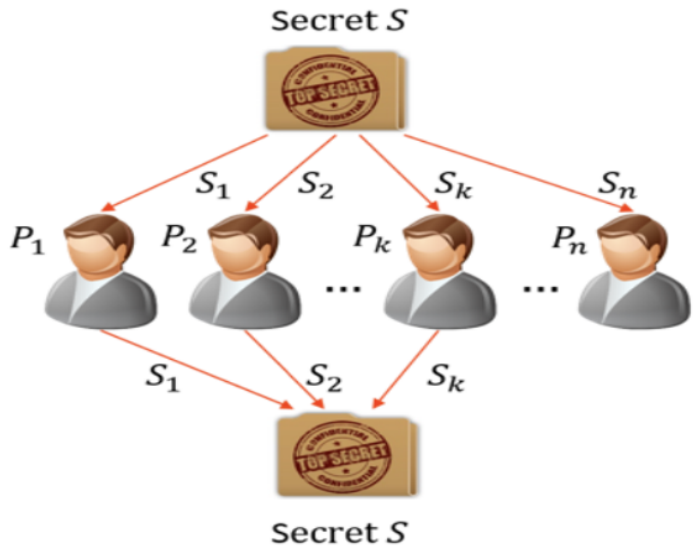


**Figure A1.** Secret Sharing Technique [34].

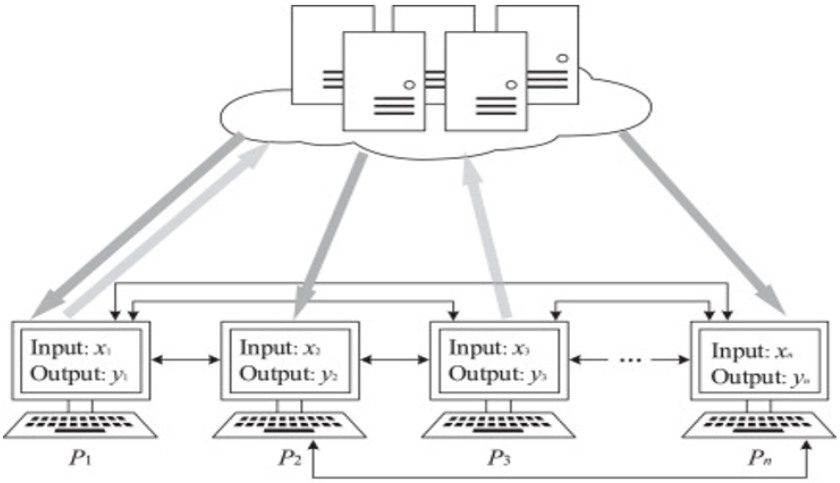


**Figure A2.** Secure Multi-Party Computation Technique [35].

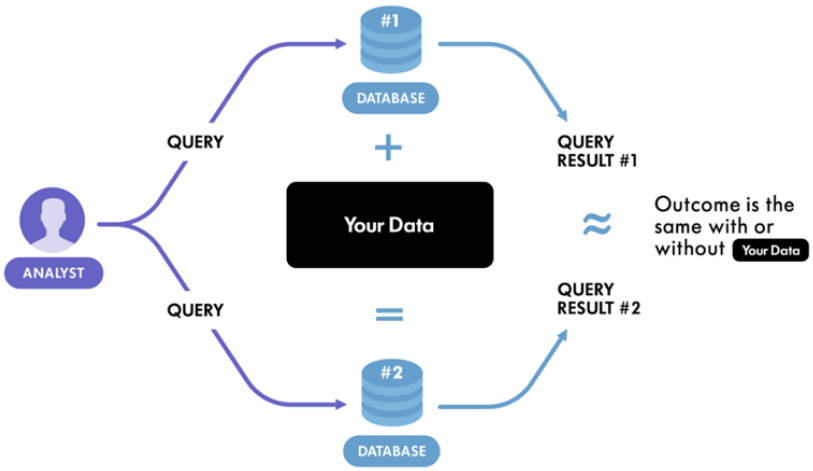


**Figure A3.** Differential Privacy Technique. Available online: https://medium.com/ydata-ai/differential-privacy-a-brief-introduction-fee4756d19e (accessed on 6 June 2021).

## References


1. Sanderson, R. Introduction to Remote Sensing. Available online: http://faculty.kfupm.edu.sa/crp/bramadan/crp514/readings/7%20-%20Introd_Remote_Sensing_Dr_Sanderson_New_Mexico_State_Univ_38Pages.pdf (accessed on 6 June 2021).
2. Carranza-García, M.; García-Gutiérrez, J.; Riquelme, J.C. A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sens.* **2019**, *11*, 274. [CrossRef]
3. Costache, R.; Bao Pham, Q.; Corodescu-Roșca, E.; Cîmpianu, C.; Hong, H.; Thi Thuy Linh, N.; Ming Fai, C.; Najah Ahmed, A.; Vojtek, M.; Muhammed Pandhiani, S.; et al. Using GIS, remote sensing, and machine learning to highlight the correlation between the land-use/land-cover changes and flash-flood potential. *Remote Sens.* **2020**, *12*, 1422. [CrossRef]
4. Nowakowski, A.; Mrziglod, J.; Spiller, D.; Bonifacio, R.; Ferrari, I.; Mathieu, P.P.; Garcia-Herranz, M.; Kim, D.H. Crop type mapping by using transfer learning. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *98*, 102313. [CrossRef]
5. Pan, X.; Yang, F.; Gao, L.; Chen, Z.; Zhang, B.; Fan, H.; Ren, J. Building Extraction from High-Resolution Aerial Imagery Using a Generative Adversarial Network with Spatial and Channel Attention Mechanisms. *Remote Sens.* **2019**, *11*, 917. [CrossRef]
6. Hajjaji, Y.; Boulila, W.; Farah, I.R.; Romdhani, I.; Hussain, A. Big data and IoT-based applications in smart environments: A systematic review. *Comput. Sci. Rev.* **2021**, *39*, 100318. [CrossRef]
7. Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, 1–13. [CrossRef]
8. Fan, Y.; Bai, J.; Lei, X.; Zhang, Y.; Zhang, B.; Li, K.C.; Tan, G. Privacy preserving based logistic regression on big data. *J. Netw. Comput. Appl.* **2020**, *171*, 102769. [CrossRef]
9. Al-Rubaie, M.; Chang, J.M. Privacy-preserving machine learning: Threats and solutions. *IEEE Secur. Priv.* **2019**, *17*, 49–58. [CrossRef]
10. Raynal, M.; Achanta, R.; Humbert, M. Image Obfuscation for Privacy-Preserving Machine Learning. *arXiv* **2020**, arXiv:2010.10139.
11. Boulemtafes, A.; Derhab, A.; Challal, Y. A review of privacy-preserving techniques for deep learning. *Neurocomputing* **2020**, *384*, 21–45. [CrossRef]
12. Tu, F.; Yin, S.; Ouyang, P.; Tang, S.; Liu, L.; Wei, S. Deep convolutional neural network architecture with reconfigurable computation patterns. *IEEE Trans. Very Large Scale Integr. (Vlsi) Syst.* **2017**, *25*, 2220–2233. [CrossRef]
13. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
14. Li, Q.; Cai, W.; Wang, X.; Zhou, Y.; Feng, D.D.; Chen, M. Medical image classification with convolutional neural network. In Proceedings of the 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 10–12 December 2014; pp. 844–848.
15. Tanuwidjaja, H.C.; Choi, R.; Baek, S.; Kim, K. Privacy-Preserving Deep Learning on Machine Learning as a Service—A Comprehensive Survey. *IEEE Access* **2020**, *8*, 167425–167447. [CrossRef]
16. Acar, A.; Aksu, H.; Uluagac, A.S.; Conti, M. A survey on homomorphic encryption schemes: Theory and implementation. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–35. [CrossRef]
17. Shafee, A.; Awaad, T.A. Privacy attacks against deep learning models and their countermeasures. *J. Syst. Archit.* **2020**, *114*, 101940. [CrossRef]
18. Huang, Z.; Pan, Z.; Lei, B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sens.* **2017**, *9*, 907. [CrossRef]
19. Ng, H.W.; Nguyen, V.D.; Vonikakis, V.; Winkler, S. Deep learning for emotion recognition on small datasets using transfer learning. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, Seattle, WA, USA, 9–13 November 2015; pp. 443–449.
20. Serra, E.; Sharma, A.; Joaristi, M.; Korzh, O. Unknown landscape identification with CNN transfer learning. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018; pp. 813–820.
21. Rezende, E.; Ruppert, G.; Carvalho, T.; Theophilo, A.; Ramos, F.; de Geus, P. Malicious software classification using VGG16 deep neural network's bottleneck features. In *Information Technology-New Generations*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 51–59.
22. Liu, B.; Zhang, X.; Gao, Z.; Chen, L. Weld defect images classification with vgg16-based neural network. In Proceedings of the International Forum on Digital TV and Wireless Multimedia Communications, Shanghai, China, 8–9 November 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 215–223.
23. Rezende, E.; Ruppert, G.; Carvalho, T.; Ramos, F.; De Geus, P. Malicious software classification using transfer learning of resnet-50 deep neural network. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; pp. 1011–1014.
24. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
25. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
26. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

27. Briggs, C.; Fan, Z.; Andras, P. A Review of Privacy-preserving Federated Learning for the Internet-of-Things. *arXiv* **2020**, arXiv:2004.11794.
28. Tariq, M.I.; Memon, N.A.; Ahmed, S.; Tayyaba, S.; Mushtaq, M.T.; Mian, N.A.; Imran, M.; Ashraf, M.W. A Review of Deep Learning Security and Privacy Defensive Techniques. *Mob. Inf. Syst.* **2020**, *2020*, 6535834. [CrossRef]
29. Rivest, R.L.; Adleman, L.; Dertouzos, M.L. On data banks and privacy homomorphisms. *Found. Secur. Comput.* **1978**, *4*, 169–180.
30. Shrestha, R.; Kim, S. Integration of IoT with blockchain and homomorphic encryption: Challenging issues and opportunities. In *Advances in Computers*; Elsevier: Amsterdam, The Netherlands, 2019; Volume 115, pp. 293–331.
31. Parmar, P.V.; Padhar, S.B.; Patel, S.N.; Bhatt, N.I.; Jhaveri, R.H. Survey of various homomorphic encryption algorithms and schemes. *Int. J. Comput. Appl.* **2014**, *91*. [CrossRef]
32. Tebaa, M.; El Hajji, S.; El Ghazi, A. Homomorphic encryption applied to the cloud computing security. In Proceedings of the World Congress on Engineering, London, UK, 4–6 July 2012; Volume 1, pp. 4–6.
33. Huang, K.; Liu, X.; Fu, S.; Guo, D.; Xu, M. A lightweight privacy-preserving CNN feature extraction framework for mobile sensing. *IEEE Trans. Dependable Secur. Comput.* **2019**. [CrossRef]
34. Tso, R.; Liu, Z.Y.; Hsiao, J.H. Distributed E-voting and E-bidding systems based on smart contract. *Electronics* **2019**, *8*, 422. [CrossRef]
35. Zhao, C.; Zhao, S.; Zhao, M.; Chen, Z.; Gao, C.Z.; Li, H.; Tan, Y.A. Secure multi-party computation: Theory, practice and applications. *Inf. Sci.* **2019**, *476*, 357–372. [CrossRef]
36. Kaissis, G.A.; Makowski, M.R.; Rückert, D.; Braren, R.F. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat. Mach. Intell.* **2020**, *2*, 305–311. [CrossRef]
37. Wood, A.; Altman, M.; Bembenek, A.; Bun, M.; Gaboardi, M.; Honaker, J.; Nissim, K.; O'Brien, D.R.; Steinke, T.; Vadhan, S. Differential privacy: A primer for a non-technical audience. *Vand. J. Ent. Tech. L.* **2018**, *21*, 209. [CrossRef]
38. Truex, S.; Baracaldo, N.; Anwar, A.; Steinke, T.; Ludwig, H.; Zhang, R.; Zhou, Y. A hybrid approach to privacy-preserving federated learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, London, UK, 15 November 2019; pp. 1–11.
39. Aono, Y.; Hayashi, T.; Wang, L.; Moriai, S. Privacy-preserving deep learning via additively homomorphic encryption. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 1333–1345.
40. Visvikis, D.; Le Rest, C.C.; Jaouen, V.; Hatt, M. Artificial intelligence, machine (deep) learning and radio (geno) mics: Definitions and nuclear medicine imaging applications. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 2630–2637. [CrossRef]
41. Wang, S.; Chang, J.M. Privacy-Preserving Image Classification in the Local Setting. *arXiv* **2020**, arXiv:2002.03261.
42. Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 308–318.
43. Ma, Z.; Liu, Y.; Liu, X.; Ma, J.; Ren, K. Lightweight privacy-preserving ensemble classification for face recognition. *IEEE Internet Things J.* **2019**, *6*, 5778–5790. [CrossRef]
44. Xia, Z.; Gu, Q.; Xiong, L.; Zhou, W.; Weng, J. Privacy-Preserving Image Retrieval Based on Additive Secret Sharing. *arXiv* **2020**, arXiv:2009.06893.
45. Erkin, Z.; Franz, M.; Guajardo, J.; Katzenbeisser, S.; Lagendijk, I.; Toft, T. Privacy-preserving face recognition. In Proceedings of the International symposium on privacy enhancing technologies symposium, Seattle, WA, USA, 5–7 August 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 235–253.
46. Morris, L. Analysis of partially and fully homomorphic encryption. *Rochester Inst. Technol.* **2013**, 1–5. Available online: http://gauss.ececs.uc.edu/Courses/c6056/pdf/homo-outline.pdf (accessed on 3 June 2021).
47. Yi, X.; Paulet, R.; Bertino, E. Homomorphic encryption. In *Homomorphic Encryption and Applications*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 27–46.
48. El Makkaoui, K.; Ezzati, A.; Beni-Hssane, A. Cloud-RSA: An enhanced homomorphic encryption scheme. In *Europe and MENA Cooperation Advances in Information and Communication Technologies*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 471–480.
49. Phung, V.H.; Rhee, E.J. A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. *Appl. Sci.* **2019**, *9*, 4500. [CrossRef]
50. Kassani, S.H.; Kassani, P.H. A comparative study of deep learning architectures on melanoma detection. *Tissue Cell* **2019**, *58*, 76–83. [CrossRef] [PubMed]
51. Boulila, W. A top-down approach for semantic segmentation of big remote sensing images. *Earth Sci. Inform.* **2019**, *12*, 295–306. [CrossRef]
52. Boulila, W.; Sellami, M.; Driss, M.; Al-Sarem, M.; Safaei, M.; Ghaleb, F.A. RS-DCNN: A novel distributed convolutional-neural-networks based-approach for big remote-sensing image classification. *Comput. Electron. Agric.* **2021**, *182*, 106014. [CrossRef]
53. Boulila, W.; Ghandorh, H.; Khan, M.A.; Ahmed, F.; Ahmad, J. A novel CNN-LSTM-based approach to predict urban expansion. *Ecol. Inform.* **2021**, *64*, 101325. [CrossRef]
54. Ahmad, J.; Ahmed, F. Efficiency analysis and security evaluation of image encryption schemes. *Computing* **2010**, *23*, 25.
55. Qayyum, A.; Ahmad, J.; Boulila, W.; Rubaiee, S.; Masood, F.; Khan, F.; Buchanan, W.J. Chaos-based confusion and diffusion of image pixels using dynamic substitution. *IEEE Access* **2020**, *8*, 140876–140895. [CrossRef]

56. Masood, F.; Boulila, W.; Ahmad, J.; Arshad; Sankar, S.; Rubaiee, S.; Buchanan, W.J. A novel privacy approach of digital aerial images based on mersenne twister method with DNA genetic encoding and chaos. *Remote Sens.* **2020**, *12*, 1893. [CrossRef]
57. Khan, J.S.; Boulila, W.; Ahmad, J.; Rubaiee, S.; Rehman, A.U.; Alroobaea, R.; Buchanan, W.J. DNA and plaintext dependent chaotic visual selective image encryption. *IEEE Access* **2020**, *8*, 159732–159744. [CrossRef]
58. Ahmad, J.; Hwang, S.O. A secure image encryption scheme based on chaotic maps and affine transformation. *Multimed. Tools Appl.* **2016**, *75*, 13951–13976. [CrossRef]
59. Ali, N.H.M.; Abead, S.A. Modified Blowfish Algorithm for Image Encryption using Multi Keys based on five Sboxes. *Iraqi J. Sci.* **2016**, *57*, 2968–2978.
60. Rad, R.M.; Attar, A.; Atani, R.E. A new fast and simple image encryption algorithm using scan patterns and XOR. *Int. J. Signal Process. Image Process. Pattern Recognit.* **2013**, *6*, 275–290. [CrossRef]
61. Dosselmann, R.; Yang, X.D. A comprehensive assessment of the structural similarity index. *Signal Image Video Process.* **2011**, *5*, 81–91. [CrossRef]