Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

An Efficient Privacy-preserving Deep Learning Scheme for Medical Image Analysis

J. Andrew Onesimu*

*Corresponding Author, Ph.D. Candidate, School of Computer Science and Engineering, Vellore Institute of Technology, India. E-mail: onesimu@gmail.com

J. Karthikeyan

Assistant Professor, School of Information Technology and Engineering, Vellore Institute of Technology, India. E-mail: karthikeyan.jk@vit.ac.in

Abstract

In recent privacy has emerged as one of the major concerns of deep learning, since it requires huge amount of personal data. Medical Image Analysis is one of the prominent areas where sensitive data are shared to a third party service provider. In this paper, a secure deep learning scheme called Metamorphosed Learning (MpLe) is proposed to protect the privacy of images in medical image analysis. An augmented convolutional layer and image morphing are two main components of MpLe scheme. Data providers morph the images without privacy information using image morphing component. The human unrecognizable image is then delivered to the service providers who then apply deep learning algorithms on morphed data using augmented convolution layer without any performance penalty. MpLe provides sturdy security and privacy with optimal computational overhead. The proposed scheme is experimented using VGG-16 network on CIFAR dataset. The performance of MpLe is compared with similar works such as GAZELLE and MiniONN and found that the MpLe attracts very less computational and data transmission overhead. MpLe is also analyzed for various adversarial attack and realized that the success rate is as low as 7.9×10^{-90}. The efficiency of the proposed scheme is proved through experimental and performance analysis.

Keywords: Deep learning, Data privacy, Image privacy, Medical image analysis, Data morphing.

DOI: 10.22059/jitm.2020.79191

Introduction

Recently, deep learning has become popular for its improved performance in wide range of application such as communication (Dorner et al., 2018), industrial automation (Liang et al., 2018), e-commerce (Cui et al., n.d.), and healthcare (Miotto et al., 2018). Healthcare services have seen a rise in the performance since the advancements of machine learning and deep learning. Healthcare service systems provide numerous services to the patients such as disease diagnosis, medical prescription, drug discovery, etc. These services are largely depends on patient data. However, patients are reluctant to submit their data to the third party service provider because of privacy breach risks. This impact the performance of deep learning application since it requires huge amount of data to yield a better performance.

Medical image analysis (Shen et al., 2017) is one the prominent areas where large amount of patient medical images are required for disease diagnosis, image segmentation, image processing, and classification. Generally, patients are not willing to disclose their diagnosis or disease to the public. Unauthorized disclosure diagnosis details affect their personal and social life. For example, if the employer of comes to know that one of his employee is diagnosed with some severe illness it affect their social life in the organization. So, it important to keep the diagnosis of a person confidential. However, without such information it is impossible to have an efficient deep learning application. Hence, a secure privacy preserving deep learning scheme is demanding.

Privacy preserving deep learning has been a research focus at present and a good number of literatures are published (Al et al., 2018; Andrew et al., 2019; Andrew J & Karthikeyan J, 2019; Hesamifard et al., 2017; J et al., 2019; X. Ma et al., 2018; Phan et al., 2019; Phong & Phuong, 2019; Shokri & Shmatikov, 2015; Y. Wang et al., 2018; D. Zhang et al., 2018; Q. Zhang et al., 2016). The idea is to separate the privacy information from the original data with additional communication and computational overhead. The privacy preserving deep learning can be achieved through some of the secure approaches such as differential privacy (Biryukov et al., 2011), homomorphic encryption (Gentry & Boneh, 2009), and secure multi-party computation (Goldreich, 1998). However, these techniques suffer from cryptographic limitations, computational overheads, and communication overheads that highly correlated to the deep learning network (Andrew & Kathrine, 2018). The limitations and overheads are directly proportional to the depth of the deep learning network so the increase in the depth makes the privacy preserving approaches impractical in many cases. In some scenarios the privacy preserving techniques transmits limited features than the actual features that degrades the performance of deep learning networks.

The aforementioned drawbacks of the existing privacy preserving techniques are addressed and to overcome the shortcomings a Metamorphose Learning (MpLe) is proposed in this paper. MpLe is a secure and efficient privacy preserving scheme for deep learning

based medical image analysis. The proposed scheme is aimed to deliver the complete data for deep learning training while protecting the privacy of the images. The main contributions of this paper are as follows:

- Image Morphing is one of the components of MpLe that perpetrates shifting of original data to a linear multidimensional space. Image morphing rearranges the matrix values of the original image and replaces the matrix with other values in order to hide the original image feature and makes it unrecognizable for human observers.
- Augmented Convolution layer is the next major component of MpLe that inversely transforms the morphed data and replaces the first layer of the deep learning convolutional neural network (CNN). This operation is focused to restore the performance degradation caused by image morphing in the network.
- The proposed scheme is evaluated for its efficiency using publicly available dataset CIFAR. We also experimented the augmented convolution layer with VGG-16 network and found that the performance of the network is well within the margin of error.
- We perform security analysis of MpLe on security and privacy attacks such as brute-force attack, reversing attack, and pair attack.
- Performance evaluation of the proposed scheme is evaluated based on similar state-of-the-art works such as GAZELLE, and MiniONN.

The remainder of the paper is organized as follows. In section 2, the literature review is presented. In section 3, we introduce a threat model and the notions used in the paper. In section 4, the proposed methodology is presented. In section 5, the details of experimental requirements and experimental analysis are presented. In section 6, the performance evaluation of the proposed system is discussed. Finally, section 7 concludes the paper.

Literature Review

In this section, we present state-of-the-art literatures published on various privacy preserving techniques to protect the high-dimensional data. We identified three major privacy preserving techniques they are: (1) Differential Privacy, (2) Homomorphic Encryption, and (3) Secure Multi-Party Computation.

Differential Privacy

Differential privacy is one of the privacy preserving techniques used in deep learning to secure the training data leakage from the model. It adds differential noise to the dataset to protect the original data from leaking. Differential privacy mechanism to protect the deep

learning model is proposed in (Abadi et al., 2016). The sensitive information required to train the model should not be exposed through the model. To achieve this a differential privacy based stochastic gradient descent algorithm is proposed. It adds the differential noise to the SGD to protect the privacy of the model. This is achieved with some considerable information loss without affecting the model accuracy. Ping Li et al. (P. Li et al., 2018) proposed a privacy preserving machine learning approach through double decryption public key cryptography technique and differential privacy. The data from multiple sources are outsourced to semi-honest cloud after encryption using the key generated locally. The cloud server then adds the differential noise to the encrypted data also the dataset is randomized then machine learning task carried out on the randomized data thus it protects the privacy of the outsourced data also decreases the computational complexity. In (Alguliyev et al., 2019) privacy preserving deep learning architecture is proposed. This architecture consists of a sparse denoising autoencoder and CNN. The sparse denoising autoencoder trained on time series data to transform the data. CNN is to classify the transformed data. Privacy preserving big data analysis is achieved by transforming the raw data into black, gray, and white classes. These outputs are given as the input to the CNN layer which performs privacy preserving data classification. The effectiveness of the model is evaluated using MSE (Mean Squared Error).

Differential privacy mechanism is adopted in (Y. Li et al., 2019) to protect the privacy of the model from training data. The model uses depth-wise separable convolution network to speed the image recognition process and privacy of the model is ensured by adding differential noise to the training data along with the fake data generated by GAN. Privacy disclosure issue during similarity joins of dataset is addressed in (Ding et al., 2019). The confidentiality of the data is ensured through differential privacy along with MapReduce. In (Sagirlar et al., 2018) the privacy of IoT smart devices are addressed. The decentralized approach of the smart objects allows the user to set the privacy preference. The privacy of IoT meta data are protected during joins, aggregation and other database actions through the proposed framework. Yang et al. (Yang et al., 2020) addressed the problem of image privacy and possible privacy breach through identification of the images. The authors have proposed a graph based neural network for image privacy to protect the images from leaking the privacy information during image classifications. The technique is based on graph, at first it extracts image features and regions then based on that the graph nodes are initialized. The representations of the image is protected through differential privacy mechanism.

Homomorphic Encryption

Homomorphic Encryption (HE) is also called as Paillier cryptosystem which allows computations to be carried out on an encrypted form of data without needing to decrypt it. Ciphertext computation results can then be decrypted for actual results. Many machine learning algorithms are depending on cloud server so the data holders are required to submit

their personal data to the third party cloud server, there arises privacy risk since the personal data is involved. In order to address the privacy issues related with third party cloud server an homomorphic framework on non-abelian rings is proposed by (J. Li et al., 2020). Displacement matrix concept is introduced to reduce the computational complexity of cipher text. Xiong et al. (Xiong & Dong, 2019) proposed a modified homomorphic encryption scheme called somewhat homomorphic encryption (SHE) to protect the privacy of images. In general the images are encrypted and the computations are carried out over the encrypted images with the help of HE. But SHE is aimed to protect the content of the image by ensuring encryption during the data embedding and extraction process. Block level prediction error expansion method is also used to protect the secret content of the encrypted image.

Securing the deep learning model from privacy leakage through intermediate results are one of the challenging issues. In (P. Li et al., 2017) a homomorphic encryption based scheme is proposed to protect the privacy of the deep learning model and the intermediate results. Multi-Key FHE scheme is utilized to secure the multiple parties who are involved in training the deep learning model. A hybrid structure of double decryption and FHE is utilized to minimize the computation and communication cost. To overcome the issue of Paillier cryptosystem which only supports integers a public key cryptosystem based on homomorphic encryption and Bresson cryptosystem is proposed in (Z. Ma et al., 2019). In this paper, multiple data owners jointly trained a random forest (RF) classifier without leaking their privacy by encrypting their data with their own secret keys. Homomorphic encryption system allows the RF to be trained on encrypted data thus it preserves the privacy of the data owners. Additive Homomorphic Encryption (AHE) based privacy preserving classification scheme is proposed in (T. Li et al., 2019). In this scheme the privacy issues of data and classifier are addressed. The classifier training is outsourced to a trusted server hence the data owner doesn’t involve in training. The computational cost of AHE is high, this makes this scheme less suitable for cost efficient devices.

Secure Multi-party Computation

Secure multi-party computation (SMC) is a privacy preserving technique where multiple parties jointly perform a computation without necessary to share their private data. Although differential privacy provides reasonable privacy, the accuracy of privacy preservation and data utility is less compared to privacy violating models. In (X. Ma et al., 2018) an alternative for differential privacy is proposed through cryptographic tools in the field of deep learning by Ma et al. The proposed model comprises of data verifiability, ElGamal encryption, and Diffie-Hellman key exchange. Multi-party deep learning allows distributed data owners to collaboratively train a model. However, privacy issues such as parameter leakage and adversary threats are viable. Thus data verifiability of the proposed model allows the user to verify the correctness of the data received from the server. ElGamal encryption method is

utilized to encrypt the gradients and parameters from adversaries. This ensures the data privacy and parameter privacy of the model.

Preserving privacy of images in computation less devices is proposed in (Rahim et al., 2018) by Rahim et al. Feature extraction is a high computational task which cannot be carried out in computation less devices such as mobile phone. Hence, the feature extraction part will be carried out separately and the mobile device will encrypt the input images to preserve the privacy. The images are then decrypted and fed to the pretrained CNN model for feature extraction. The mobile device uses cost effective SHA-256 hash function to hash the secret and public keys. Wang et al. (W. Wang et al., 2020), proposed a privacy preserving mixed set operations based on SMC. The protocol considers a semi-honest model where multiple parties joins together for a computation but tries to ascertain others private data. The encoding scheme of SMC and the threshold ElGamal cryptosystems are adopted to protect the privacy of the parties in the network.

According to the state-of-the-art literature review, we understood that there is huge risk involved in deep learning networks. Various cryptography based implementations have been proposed to mitigate the privacy issues in the deep learning networks. However, they suffer from cryptographic limitations, computational overheads, and communication overheads. Also, it is observed that privacy preserving techniques have impacted the performance of the deep learning approaches. Hence, an efficient and secure privacy preserving scheme for deep learning approach is crucial.

Preliminaries

Threat Model

In our work, we assume a scenario where DH_i is a data holder who holds personal data ($i \leq n$) where n is the number of data holders. There is a team of deep learning experts who requests data from the data holders to develop deep learning application. We use the term *developers* to denote the deep learning service provider. The data holders transfers the data to the third party *developers* in order to get appropriate service. DH_i comprises of personal data which needs to be protected. So, the data holder is expected to morph the data before submitting it to the *developers*. Then the *developers* apply deep learning techniques to perform some desired task to provide appropriate services without performance penalty.

In our work, we consider two threat models they are Honest-but-Curious (HBC) and Semi-Honest-Curious (SMC). In both threat models, we consider *developer* as the adversary who tries to acquire more information. In HBC, the adversarial *developer* follows the privacy rules but out of curiosity he tries to acquire more information about the data holder. In SMC, the adversarial *developer* is accessible to data holder's personal data who is does not follow

the rule appropriately and tries to ascertain more knowledge about the data holder. We analyze our proposed scheme using these threat models.

Notations

In this section we describe the various notations used in this paper. (Table 1) shows the notations and the descriptions used in this paper.

Table 1. Notations

Notations	Description
DH_i	Data Holder
D	Image data
D^r	Unrolled row vector
A	Matrices
$A_{x,y}$	Matrix element with (x, y) coordinate
$K_{i,j}$	Convolution layer kernel
α, β	Number of channels in convolution kernels
T^r	Morphed data
M	Morphing matrix

Proposed Methodology

In this section, we present the proposed MpLe system. There are two main components of MpLe they are image morphing and augmented convolution layer. In a medical image analysis system, the data holder first morph their medical images using image morphing which is a key component of MpLe. Augmented convolution layer component is then used to perform deep learning network on the morphed data without performance penalty. The augmented convolution layer is designed in such a way to be compatible with any types of deep learning network. It replaces the first layer of the deep learning network to perform the task on morphed data without performance degradation. (Figure 1) shows the architecture of MpLe. The architecture shows the original neural network architecture with MpLe components. The original data of the data holders are morphed using image morphing and the morphed matrix is then submitted to the deep learning network. The original deep learning network's first layer is replaced with augmented convolutional layer in order to perform the desired task without performance penalty.

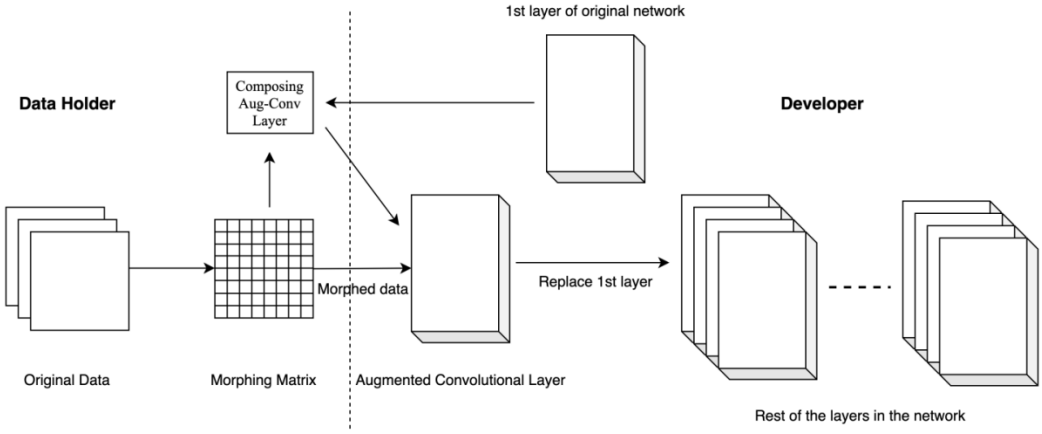


Figure 1. Architecture of Metamorphose Learning

Vectorize Data

The basic operation of MpLe is to replace the first convolution layer in the deep learning network with an equivalent matrix operation. It is the essential step for both image morphing and augmented convolution layer. Unrolling the image to single dimensional is the key for data parallelism in deep learning frameworks. So, at first the input image matrix is unrolled to column vector. The column vector from each sliding window of the input data is then performed matrix to matrix multiplication. The matrix columns are later converted to row vector to form the image matrix. The detailed steps of the vectorizing image data is shown in algorithm 1.

algorithm1 (Vectorize Data)

1: Input: Image data D
2: Output: Unrolled Row Vector D^rUnroll the input data D
3: assign smaller row index values for all channel vector
4: concatenate channel vectors $D^r[1 \times \alpha m^2]$
5: replace the convolutional layer using Matrix C
6: C dimension $\alpha m^2 \times \beta n^2$
7: Initialize C $\Rightarrow$ 0
8: **foreach** weight $(k_{(i,j),(a,b)})$ **do**
9: **if** (x, y) satisfies **then**
10: $x = im^2 + am + cm + b + d$ where $i \in [0, \alpha)$
11: $y = jn^2 + cn + d$ where $j \in [0, \beta)$
12: $C_{x,y} = k_{(i,j),(a,b)}$
13: $c, d \in [0, m - p + 1)$; $a, b \in [0, p)$
14: $F^r[1 \times \beta n^2] = D^r . C$
15: **end if**
16: **end for**
17: return F^r

Image Morphing

Image morphing is one of the essential component of MpLe scheme to protect the privacy of the images. Image morphing component gets the original image as input and replaces the matrix values with morphing data to make the original image human unrecognizable. D^r is the input of image morphing component and the output is the morphed image T^r. In order to achieve this, it has to follow certain requirements they are, similar input and output image size, optimal computational cost, and unrecognizable image morphing. The detailed steps of the requirement and image morphing process is shown in algorithm 2.

algorithm 2 (Image Morphing)

1: Input: D^r unrolled data

2: Output: T^r morphed data

3: construct matrix *M'[q,q]=M.random.randint(1,)*

4: $q \Rightarrow k = \frac{\alpha m^2}{q} \epsilon Z$ where *k* is morphing factor

5: construct diagonal *M'[q,q] to M'[αm² × αm²]*

6: $M_{x,y} = \begin{cases} M'_{x-Nq,y-Nq} \\ 0, \ otherwise \end{cases}$

7: $T^r = D^r . M$

8: return T^r

Augmented Convolutional Layer

Augmented convolutional layer is another main component of MpLe. It comprises of feature extraction, conv-layer formation, original data restoration, and channel order randomization steps. In the feature extraction step, the main focus is to extract the features from the image similar to the original images. It is done by performing dot matrix product of the unrolled data vector and convolution layer matrix. The calculated output are then transformed to perform matrix multiplication with the morphed data to get the original features of the image. In convolution layer formation, a convolution layer is trained on similar dataset to extract the low level features of the image. In original data restoration, the original features are restored to reduce the performance penalty. Channel order randomization is performed to ensure the deep learning network from inverse attack. The unrolled image vector are randomly distributed using *rand* function in order to protect the data from inverse attack. The detailed steps of augmented convolution layer is shown in algorithm 3.

algorithm 3 (Augmented Convolution Layer)

Feature extraction

1: $F^r[1\times\beta n^2] = D^r.C$

2: inverse matrix $M \Rightarrow M^{-1}$

3: $T^r.M^{-1} = D^r$

Conv-layer formation

4: $C^{ac} = M^{-1}.C$

original data restoration

5: $T^r.C^{ac} = T^r.M^{-1}.C = D^r.C = F^r$

Channel order randomization

6: output feature shuffling $F^{'} = rand(F)$

7: $D^r \neq rand(F^r).C^{-1}$

8: reduce the vulnerability of C^{ac}

9: Divide C^{ac} to β groups

10: n^2continuous columns

11: shuffle the order of column randomly

12: construct Aug-Conv layer matrix

13: randomization of features

Experiments

In this section, we discuss the experimental setup, dataset details, experimental analysis, and security analysis of the proposed scheme.

Experimental Setup

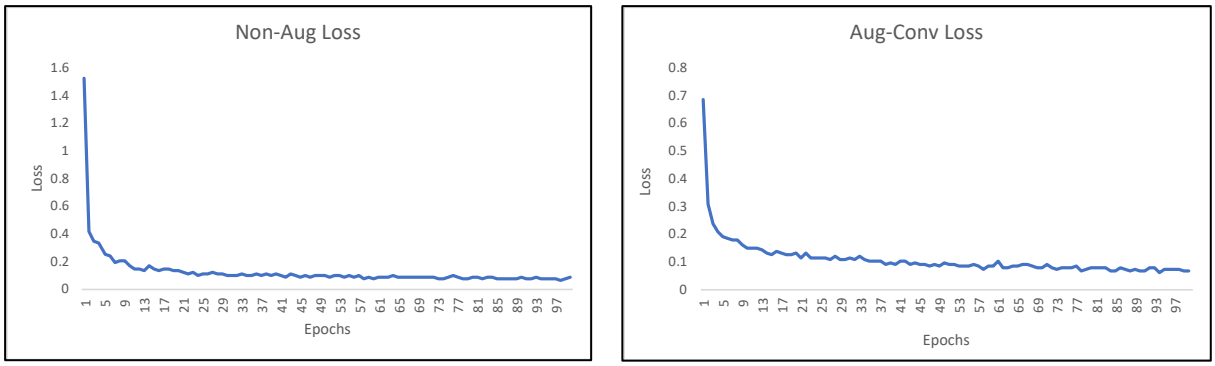
Our experiments are conducted on GPU machine with 16 GB of primary memory and 256 GB of secondary storage. The programs are written in Python language and executed in Mac OS environment.

Dataset details

We experimented our proposed scheme on CIFAR (Krizhevsky et al., 2009) dataset. CIFAR dataset is image dataset used for classification tasks. There are two variants based on the number of classes they are CIFAR-10 and CIFAR-100. The integer represents the number of classes in the dataset. Each class contains 600 images of size 32 X 32.

Experimental Analysis

We first train the network on the CIFAR-10 and CIFAR-100 dataset. After training the VGG16 network we generate augmented convolution layer matrix. We test the accuracy of the VGG16 network on using morphed data as the training data. Testing of the network along with the augmented layer using morphed data as the training data is performed. Vulnerability of the network is assessed under inverse training scenario to generate the inverse training dataset. We calculate the Mean Squared Error Loss between the original image and the retrieved image.



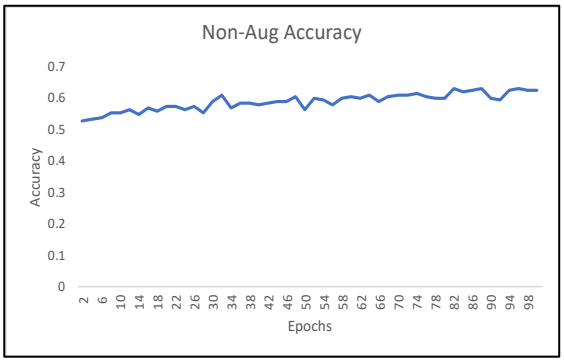
Loss of non-augmented network

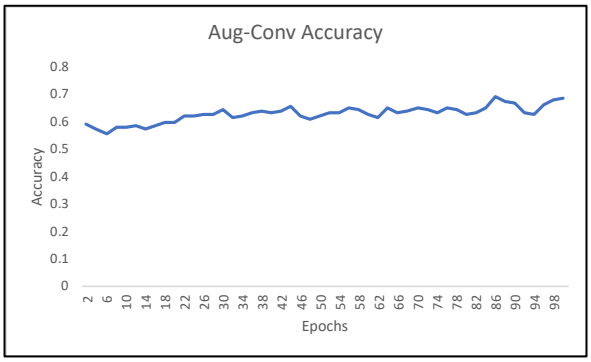
Loss of augmented convolution network

Figure 2. Model Training Loss

(Figure 2) shows the model training loss of original network i.e., without augmented convolution layer and the network that replaces the first layer as augmented convolution layer. (Figure 2. (a)) shows the non-augmented loss and (Figure 2. (b)) shows the augmented convolution network loss. It is observed from that augmented convolution layer loss is slimmer compared to non-augmented.

(Figure 3) shows the accuracy of the original network without the augmented convolution layer and the modified network with augmented convolutional layer. The graph values of compared with respect to the number of epochs versus the improvement in accuracy. It is noticed that the accuracy of the model increases with the number of epochs. The model is iterated up to 100 epochs. (Figure 3. (a)) shows the accuracy of non-augmented layer and (Figure 3. (b)) shows the network with augmented layer. It is observed that the network with augmented convolutional layer has comparatively better performance than the original network. This proves that the proposed augmented convolution layer is efficient in extracting the actual features from the morphed images.





Accuracy of non-augmented network

Accuracy of augmented convolution network

Figure 3. Model Accuracy

We further experimented our scheme on different types of medical images such as MRI and X-ray images. The image morphing can be performed based on the privacy requirement. (Figure 4) shows the various morphing of various images. The first column of the images are the original images from Brain MRI and Chest X-ray datasets. The second column shows the morphed medical images based on various privacy requirements. In some cases the images are slightly recognizable and in some cases the images are unrecognizable. This is achieved through image morphing component of MpLe to protect the security and privacy of the medical images.

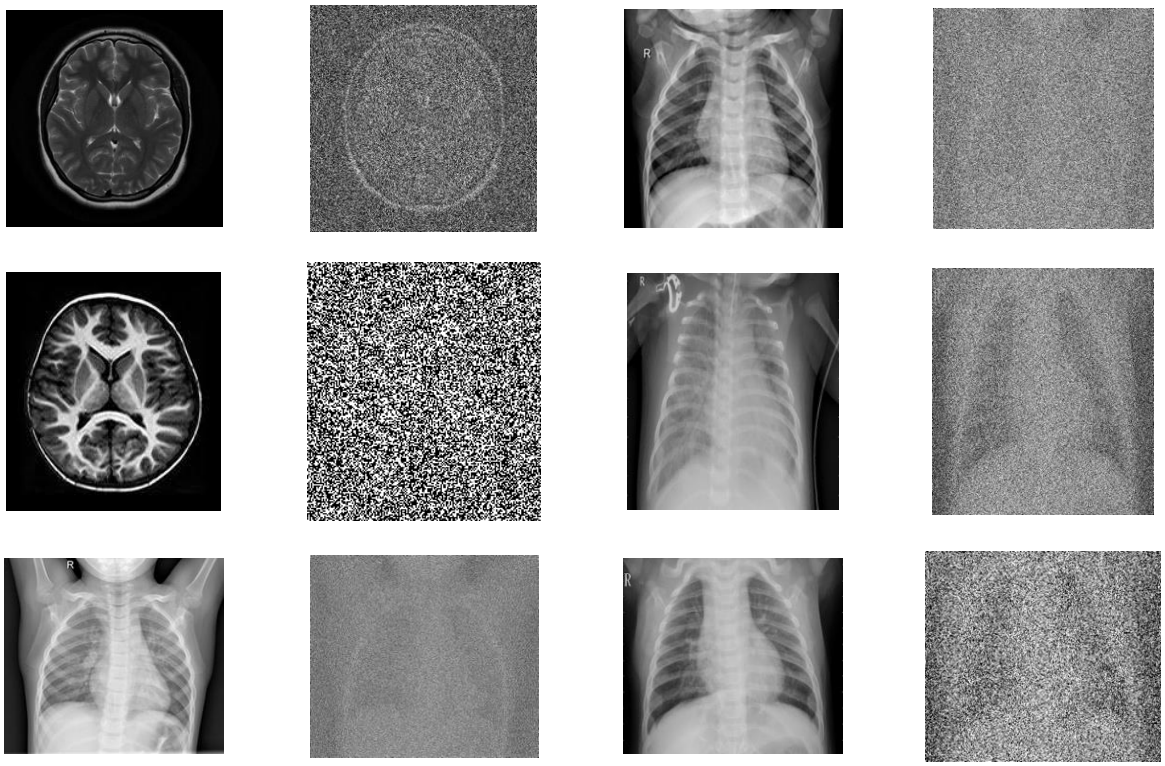


Figure 4. Image Morphing of Medical Images

Security Analysis

We perform security analysis of our proposed scheme against security and privacy attacks to ascertain the personal or sensitive information. We identified the possible attacks such as brute-force attack, reversing attack, and pair attack and provided mathematical proof for the resistance of the proposed system.

Brute Force Attack: It is one of the straightforward attack on metamorphose learning scheme on the HBC threat model. In general brute force attack means trying different possible combination to reveal the actual data. In our scheme, the matrix **M** is targeted to reveal the original feature of the images. Unlike other formats of data, any close approximation of an image data can reveal the actual information. In our proposed scheme, the image matrix are replaced with a calculated value of morphed data that makes our scheme more resilient to brute force attack. It is nearly impossible for the attacker to realize the original data features with the random combinations. Based on our network calculations the success rate of an attack is as low as 7.9×10^{-90}.

***Reversing Attack*:** This is another possible attack on HBC threat model which targets the augmented convolution layer. The augmented convolution layer is formulated in such a way to extract the original features from the morphed images for the possible classification problem. Here the HBC adversary tries to ascertain the original data features from the augmented convolutional layer through reversing attack. In our proposed scheme, this attack is neutralized through randomization techniques. One of the component of augmented convolution layer is randomization, that randomizes the network features using *rand* function with incurring a performance penalty to the deep learning network.

Pair Attack: This is a possible attack on SBC threat model. Pair attack is performed by the adversarial *developer* who has the access to the original image features from the augmented convolutional layer. The adversary tries to extracts the image feature from the augmented convolution layer and to perform the pair wise comparison to measure the performance penalty. At this time the SBC may try to acquire the actual data and tries to recognize the original image. In our proposed scheme, this attack is nullified through unrolling the original image matrix and performing matrix multiplication with unrolled vector and morphed data vector.

Performance Evaluation

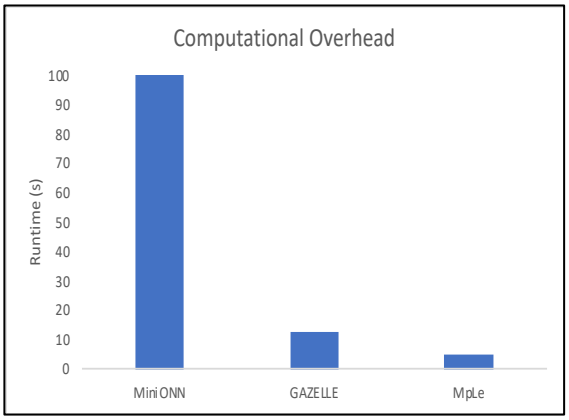
In this section, the performance of the proposed scheme is evaluated with the similar works that used CIFAR-10 datasets based on the computational overhead, communication overhead, performance penalty, and attack success rate. The proposed scheme is compared with GAZELLE (Juvekar et al., n.d.), MiniONN (Liu et al., n.d.) and ARDEN (J. Wang et al.,

2018). GAZELLE scheme is based on secure multiparty computation technique, MiniONN is based on homomorphic encryption, and ARDEN based on feature transmission technique. (Table 2) shows the comparison of performance penalty and attack success rates. It describes that the proposed system did not incur any performance penalty similar to the GAZELLE and outperforms ARDEN. In the comparison with attack success rate it is evident that our MpLe scheme is more resilient to privacy and security attacks.

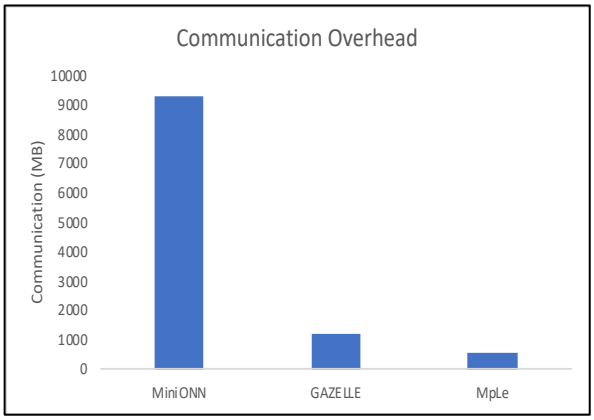
Table 2. Performance evaluation of MpLe with similar framework

Framework	Performance Penalty	Attack Success Rate
MpLe	0	7.9×10^{-90}
GAZELLE	0	2.9×10^{-39}
ARDEN	0.628	Highest probability

Computational and communication overheads are other performance metrics that are compared with the existing frameworks. We compared the performance of our proposed scheme with MiniONN and GAZELLE frameworks. (Figure 5) shows the graphical comparison of the performance of the proposed scheme with other similar frameworks. (Figure 5. (a)) shows the computational overhead and it is noticed that the performance of the proposed MpLe is has the lowest overhead compared to GAZELLE and MiniONN frameworks. Similarly our proposed scheme outperforms other existing system in terms of communication overhead also. (Figure 5.b) shows the comparative analysis of communication overhead with existing frameworks.



Computational Overhead



Communication Overhead

Figure 5. Performance Evaluation of MpLe with GAZELLE and MiniONN

Conclusion

In this paper, we presented a privacy preserving scheme for deep learning network. At first, we addressed the privacy issues of deep learning applications and the inefficiency of the existing approaches. Then, a secure and efficient privacy preserving scheme called for Metamorphose Learning is proposed to address the privacy issues of medical image analysis. The proposed scheme comprises of image morphing and augmented convolution layer. The image morphing components morphs the input image to human unrecognizable format and the augmented convolution layer is designed to replace the first layer of any deep learning network to effectively extract the original image features from the morphed data. The proposed scheme is analyzed for various privacy attacks and realized the success rate is only 7.9×10^{-90}. Experimental and performance analysis proved that our proposed scheme is efficient in protecting the privacy of the medical image data with minimal computational and communication overhead. In future, we are looking forward to implementing our scheme for various healthcare applications.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep Learning with Differential Privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, 308–318. https://doi.org/10.1145/2976749.2978318

Al, M., Chanyaswad, T., & Kung, S.-Y. (2018). Multi-Kernel, Deep Neural Network and Hybrid Models for Privacy Preserving Machine Learning. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2891–2895. https://doi.org/10.1109/ICASSP.2018.8462336

Alguliyev, R. M., Aliguliyev, R. M., & Abdullayeva, F. J. (2019). Privacy-preserving deep learning algorithm for big personal data analysis. *Journal of Industrial Information Integration*. https://doi.org/10.1016/j.jii.2019.07.002

Andrew, J., & Kathrine, G. J. W. (2018). An intrusion detection system using correlation, prioritization and clustering techniques to mitigate false alerts. In *Advances in Intelligent Systems and Computing* (Vol. 645). https://doi.org/10.1007/978-981-10-7200-0_23

Andrew, J., Mathew, S. S., & Mohit, B. (2019). *A Comprehensive Analysis of Privacy-preserving Techniques in Deep learning based Disease Prediction Systems*. 0–9. https://doi.org/10.1088/1742-6596/1362/1/012070

Andrew J, & Karthikeyan J. (2019). Privacy-Preserving Internet of Things: Techniques and Applications. *International Journal of Engineering and Advanced Technology (IJEAT)*, *8*(6), 3229–3234. https://doi.org/10.35940/ijeat.F8830.088619

Biryukov, A., De Cannière, C., Winkler, W. E., Aggarwal, C. C., Kuhn, M., Bouganim, L., Guo, Y., Preneel, B., Bleumer, G., Helleseth, T., Canetti, R., Varia, M., Peters, C., Kaliski, B., Desmedt, Y., Kesidis, G., De Soete, M., Bleumer, G., Schoenmakers, B., … Smith, S. W. (2011). Differential Privacy. In *Encyclopedia of Cryptography and Security* (pp. 338–340). Springer US. https://doi.org/10.1007/978-1-4419-5906-5_752

Cui, L., Huang, S., Wei, F., Tan, C., Duan, C., & Zhou, M. (n.d.). *SuperAgent: A Customer Service Chatbot for E-commerce Websites*. 97–102. https://doi.org/10.18653/v1/P17-4017

Ding, X., Yang, W., Raymond Choo, K.-K., Wang, X., & Jin, H. (2019). Privacy preserving similarity joins using MapReduce. *Information Sciences*, *493*, 20–33. https://doi.org/10.1016/J.INS.2019.03.035

Dorner, S., Cammerer, S., Hoydis, J., & Brink, S. Ten. (2018). Deep Learning Based Communication over the Air. *IEEE Journal on Selected Topics in Signal Processing*, *12*(1), 132–143. https://doi.org/10.1109/JSTSP.2017.2784180

Gentry, C., & Boneh, D. (2009). *A fully homomorphic encryption scheme*. http://cs.au.dk/~stm/local-cache/gentry-thesis.pdf

Goldreich, O. (1998). *Secure Multi-Party Computation*.

Hesamifard, E., Takabi, H., Ghasemi, M., & Jones, C. (2017). Privacy-preserving Machine Learning in Cloud. *Proceedings of the 2017 on Cloud Computing Security Workshop - CCSW '17*. https://doi.org/10.1145/3140649.3140655

J, A., Karthikeyan, J., & Jebastin, J. (2019). Privacy Preserving Big Data Publication On Cloud Using Mondrian Anonymization Techniques and Deep Neural Networks. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 722–727. https://doi.org/10.1109/ICACCS.2019.8728384

Juvekar, C., Mtl, M., Vaikuntanathan, V., & Chandrakasan, A. (n.d.). *GAZELLE: A Low Latency Framework for Secure Neural Network Inference*. Retrieved August 9, 2020, from https://www.usenix.org/conference/usenixsecurity18/presentation/juvekar

Krizhevsky, A., Hinton, G., & others. (2009). *Learning multiple layers of features from tiny images*.

Li, J., Kuang, X., Lin, S., Ma, X., & Tang, Y. (2020). Privacy preservation for machine learning training and classification based on homomorphic encryption schemes. *Information Sciences*, *526*, 166–179. https://doi.org/10.1016/j.ins.2020.03.041

Li, P., Li, J., Huang, Z., Li, T., Gao, C.-Z., Yiu, S.-M., & Chen, K. (2017). Multi-key privacy-preserving deep learning in cloud computing. *Future Generation Computer Systems*, *74*, 76–85. https://doi.org/10.1016/J.FUTURE.2017.02.006

Li, P., Li, T., Ye, H., Li, J., Chen, X., & Xiang, Y. (2018). Privacy-preserving machine learning with multiple data providers. *Future Generation Computer Systems*, *87*, 341–350. https://doi.org/10.1016/J.FUTURE.2018.04.076

Li, T., Li, X., Zhong, X., Jiang, N., & Gao, C. (2019). Communication-efficient outsourced privacy-preserving classification service using trusted processor. *Information Sciences*, *505*, 473–486. https://doi.org/10.1016/J.INS.2019.07.047

Li, Y., Wang, Y., & Li, D. (2019). Privacy-preserving lightweight face recognition. *Neurocomputing*, *363*, 212–222. https://doi.org/10.1016/J.NEUCOM.2019.07.039

Liang, J., Mahler, J., Laskey, M., Li, P., & Goldberg, K. (2018). Using dVRK teleoperation to facilitate deep learning of automation tasks for an industrial robot. *IEEE International Conference on Automation Science and Engineering*, *2017-August*, 1–8. https://doi.org/10.1109/COASE.2017.8256067

Liu, J., Juuti, M., Lu, Y., & Asokan, N. (n.d.). *Oblivious Neural Network Predictions via MiniONN transformations*. Retrieved August 9, 2020, from http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.

Ma, X., Zhang, F., Chen, X., & Shen, J. (2018). Privacy preserving multi-party computation delegation for deep learning in cloud computing. *Information Sciences*, *459*, 103–116. https://doi.org/10.1016/J.INS.2018.05.005

Ma, Z., Ma, J., Miao, Y., & Liu, X. (2019). Privacy-preserving and high-accurate outsourced disease predictor on random forest. *Information Sciences*, *496*, 225–241. https://doi.org/10.1016/J.INS.2019.05.025

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, *19*(6), 1236–1246.

Phan, N. H., Vu, M. N., Liu, Y., Jin, R., Dou, D., Wu, X., & Thai, M. T. (2019). Heterogeneous Gaussian mechanism: Preserving differential privacy in deep learning with provable robustness. *IJCAI International Joint Conference on Artificial Intelligence*, *2019-Augus*(June), 4753–4759. https://doi.org/10.24963/ijcai.2019/660

Phong, L. T., & Phuong, T. T. (2019). Privacy-Preserving Deep Learning via Weight Transmission. *IEEE Transactions on Information Forensics and Security*, 1–1. https://doi.org/10.1109/TIFS.2019.2911169

Rahim, N., Ahmad, J., Muhammad, K., Sangaiah, A. K., & Baik, S. W. (2018). Privacy-preserving image retrieval for mobile devices with deep features on the cloud. *Computer Communications*, *127*, 75–85. https://doi.org/10.1016/J.COMCOM.2018.06.001

Sagirlar, G., Carminati, B., & Ferrari, E. (2018). Decentralizing privacy enforcement for Internet of Things smart objects. *Computer Networks*, *143*, 112–125. https://doi.org/10.1016/J.COMNET.2018.07.019

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, *19*(1), 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 909–910. https://doi.org/10.1109/ALLERTON.2015.7447103

Wang, J., Zhu, X., Zhang, J., Cao, B., Bao, W., & Yu, P. S. (2018). Not just privacy: Improving performance of private deep learning in mobile cloud. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2407–2416. https://doi.org/10.1145/3219819.3220106

Wang, W., Li, S., Dou, J., & Du, R. (2020). Privacy-preserving mixed set operations. *Information Sciences*, *525*, 67–81. https://doi.org/10.1016/j.ins.2020.03.049

Wang, Y., Adams, S., Beling, P., Greenspan, S., Rajagopalan, S., Velez-Rojas, M., Mankovski, S., Boker, S., & Brown, D. (2018). Privacy Preserving Distributed Deep Learning and Its Application in Credit Card Fraud Detection. *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, 1070–1078. https://doi.org/10.1109/TrustCom/BigDataSE.2018.00150

Xiong, L., & Dong, D. (2019). Reversible data hiding in encrypted images with somewhat homomorphic encryption based on sorting block-level prediction-error expansion. *Journal of Information Security and Applications*, *47*, 78–85. https://doi.org/10.1016/j.jisa.2019.04.005

Yang, G., Cao, J., Chen, Z., Guo, J., & Li, J. (2020). Graph-based neural networks for explainable image privacy inference. *Pattern Recognition*, *105*, 107360. https://doi.org/10.1016/j.patcog.2020.107360

Zhang, D., Chen, X., Wang, D., & Shi, J. (2018). A Survey on Collaborative Deep Learning and Privacy-Preserving. *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, 652–658. https://doi.org/10.1109/DSC.2018.00104

Zhang, Q., Yang, L. T., & Chen, Z. (2016). Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning. *IEEE Transactions on Computers*, *65*(5), 1351–1362. https://doi.org/10.1109/TC.2015.2470255

Bibliographic information of this paper for citing:

Onesimu, J. Andrew & Karthikeyan, J. (2020). An Efficient Privacy-Preserving Deep Learning Scheme for Medical Image Analysis. *Journal of Information Technology Management*, Special Issue, 50-67.