Deep Learning Evaluation of a Steganographic Algorithm

Peter Eze* and Udaya Parampalli†
* The University of Melbourne, Australia
E-mail: peter.eze@unimelb.edu.au
† The University of Melbourne, Australia
E-mail: udaya@unimelb.edu.au

***Abstract*—The embedding and transmission of secret information through multimedia such as images can be used to achieve privacy in medical and military applications. This form of covert communication is called steganography. As images are also used for training machine learning algorithms for medical diagnosis, it is important to evaluate if embedded data within images have adverse effects on the performance of medical machine learning models. This paper evaluates the effect of a previously developed steganographic algorithm called the C_4S algorithm on deep machine learning models for malaria diagnosis. The steganographic algorithm is used to hide patient data in malaria blood smear images. Both original and watermarked images are used to train different models using deep learning algorithms. The goal is to determine the effect of steganography on deep machine learning algorithms. The deep learning models are used to predict if a blood smear image has a malaria parasite or not. The results show that the basic Convolutional Neural Network (CNN) model trained with combined watermarked and non-watermarked images has the best prediction accuracy of 97.85%. Basic CNN trained with only non-watermarked medical images performed better on the watermarked test set with an accuracy of 97.65% as opposed to lower performance on a non-watermarked test set at 94.50%accuracy. We conclude that it is a false generalisation to assume that medical image steganography reduces the classification accuracy of machine learning algorithms. Instead, medical image steganography algorithms should be subjected to clinical trials and approved on a case-by-case basis for use in digital health applications**

I. Introduction

Steganography is an information hiding technique which embeds information in a cover or host while maximising both the capacity of the hidden information and the fidelity of the cover or host [1], [2]. Some of the examples of the cover or host are network traffic statistics, email message, images, audio and video. Steganography is a form of covert communication in which the existence of a communication channel is concealed and only authorised receivers can view the message.

Management of medical records is one of the recent applications of information hiding where Covert or steganographic channels can be used to secure patients' health information using medical images as cover [3]. To aid medical care, artificial intelligence (AI) and deep learning algorithms are used to aid diagnosis using the medical images. When these images contain medical records as hidden data, it is important to investigate the effect of this embedded information on the accuracy of the deep learning algorithm for medical diagnosis. As different, machine and deep learning algorithms have different levels of robustness of their input features, it is also important to investigate which of the available deep learning algorithms will be the most resistant to any effect of steganography.

The study in [4] tested watermarked Fundus eye scans against some models [5], [6] used to classify Healthily, Macular Edema and Central Serous Chorioretinopathy (CSCR) eyes diseases. The accuracy of their models trained with original data ranged from 95% to 100%. Their results show that there was no difference in classification accuracy for the original and watermarked test set. However, few test data were used (15 to 45). Also, it was not clear if the original model was trained with watermarked training set. Again, this study failed to recognise that if watermarking is adopted for integrity checks, future training sets would contain watermarked data and not just the test data.

Garcia-Hernandez *et al* in [8] performed an objective evaluation of the impact of watermarking on computer-aided diagnosis in medical imaging. Two watermarking algorithms were used on half of the samples. They tried to establish the effect of spread Spectrum DCT (SS-DCT) and High Capacity Data Hiding (HCDH) watermarking on the segmentation and classification accuracy of the lesions in the image. They found that with an appropriate choice of parameters, both watermarking systems can perform well without any adverse effect on segmentation and classification accuracy. However, SS-DCT could alter the accuracy if high embedding strength is used. Their approach is most related to our work but not in the area of pneumonia disease classification using X-ray scans. Besides, they used only about 500 Breast Ultrasound as their dataset.

In our previous work in [9], [13] we developed the C_4S steganographic algorithm [2], [14] and used a traditional machine learning algorithm called support vector machine (SVM) to evaluate the impact of steganography on the classification accuracy of Pneumonia disease. Four feature sets: contrast, homogeneity, energy and entropy were used to train the SVM model. These features were extracted before watermarking and after different amounts of watermark bits were inserted into an 8x8 block of the X-ray image. Apart from recall, other machine learning parameters such as accuracy and

precision linearly degraded with increased watermark. This notion helped us to improve the C_4S algorithm.

In this paper, we leverage the C_4S steganographic algorithm [2], [9], [14] to adding watermark to malaria blood smear image samples. these samples were originally curated for training machine learning algorithms for malaria case finding. We have chosen to apply deep learning instead of traditional machine learning because of the ability of deep learning algorithm to extract detailed features that are robust against various attacks. This feature could make it more resilient and robust to steganographic modifications. Conversely, we need to establish the if the C_4S algorithm and associated images can be safely used for deep machine learning training without compromising the accuracy of the models. Hence, we make the following contributions:

1) ascertain the robustness of deep learning algorithm to C_4S algorithm.
2) compare the accuracy of the machine learning algorithms when they are trained with images containing secret watermark and clean original images.
3) evaluate the impact of the C_4S on deep learning using a large medical image dataset unlike the smaller datasets used in existing literature and our previous works.

While we describe our methods in the next section, we present our results in Section III. We then discuss the implications of these results in Section IV and conclude the work in Section V.

II. Methods

We divide this section into the steganographic algorithm and the Deep learning algorithms employed in this paper. This is because we first apply the steganographic algorithm images to the blood smear images before deep learning training.

A. Steganographic Algorithm

The C_4S spread spectrum algorithm [2], [14], which has been previously developed was applied in this research. A summary of the C_4S insertion strategy is shown in Figure 1.

C_4S is an additive spread spectrum steganographic technique used for either fragile, semi-fragile, or robust watermarking. These variation in the application of C_4S are achieved by adjusting the parameters ρ and ϵ (epsilon). The parameter, ρ is a real-valued number agreed between the sender and the receiver. It is embedded in such a way that the correlation value at the receiver between an image sub-block, X_i and a secret-key-generated sequence, $\boldsymbol{W}$ equals $\pm\rho \pm \epsilon$. Hence, ϵ is a control parameter for determining the level of fragility (or robustness) of the watermarking algorithm. In general, $\epsilon < p$. ϵ is a tolerance parameter. **G** is a gap required between insertion zones to detect tampering.

1) Embedding Function: The C_4S embedding function follows the classical spread spectrum additive embedding method:

Table I: Steganographic Algorithm Hyperparameters

	Parameter	Value
1.	Bits per block	0,1 and 4
2.	ρ	0.6
3.	ϵ	0.5
4.	Channel gap , **G**	2

$$Y_{ij} = \begin{cases} X_{ij} + \alpha W_{ij}, & \text{if } s_k = 1 \\ X_{ij} - \alpha W_{ij}, & \text{if } s_k = 0 \end{cases}, \tag{1}$$

where **Y** is watermarked image block, **X** is the cover image block, α is the embedding strength which depends on ρ and **G**, and **W** is a Pseudo random sequence code. **Y**, **X** and **W** are almost of the same dimension, $\mathbf{m} \times \mathbf{n}$. *i* and *j* refer to individual pixels in the k_{th} sub-block from the global image **X** and **Y**.

2) Extraction Function: The desired constant correlation, ρ, at the receiver is any real number agreed in advance between the sender and receiver. This means that decoding function needs to output values in the range $\pm\rho\pm\epsilon$ for correct watermark detection within an image sub-block. This gives us the watermark decoding function:

$$\overline{s} = \begin{cases} 0, & \text{if } r = \langle Y, W \rangle = \rho \pm \epsilon \\ 1, & \text{if } r = \langle Y, W \rangle = -\rho \pm \epsilon \end{cases}, \tag{2}$$

where ϵ is the tolerated error deviation from unintentional attacks, noise and quantisation errors. The expression $\langle Y, W \rangle$ is the linear correlation between Y and W. In general: $0 < \epsilon \leq \rho$.

Further details about C_4S design and algorithms can be found in [7].

3) Performance measures: In watermarking and steganography, the perceptual and structural distortion of an image are measured by Peak Signal-to-noise ratio (PSNR) and Structural Similarity Index Measure (SSIM).

PSNR is the ratio of the original cover over the noise (standard error) introduced by watermarking.

$$PNSR = 20 * log_2 \frac{B}{\sqrt{MSE}} \tag{3}$$

B is the largest value of signal or the dynamic range for the pixel values (2^n, where n is pixel depth) and **MSE** is the Mean Square Error per pixel. PSNR is a statistical degradation measure.

SSIM is a better measure of perceptual fidelity between two images, as proven in [10], [11].

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{4}$$

Where μ_x, μ_y and σ_x^2, σ_y^2 are the corresponding mean and variance of the images x and y respectively. The parameter σ_{xy} is covariance of x and y. The value of SSIM ranges from 0 to 1, where 1 indicates perfect similarity between two images.

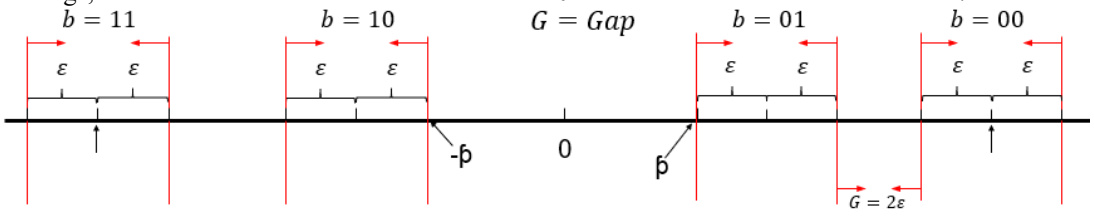


Figure 1: C_4S Insertion Strategy: Predefined correlation channels within the image represents groups of bits. Each channel is separated by a channel gap, ***G***. The width of each channel is 2ϵ

From steganographic perspective, performance analysis will be examine using the above parameters. However, the major focus of this work is on the deep learning performance parameters as discussed in Section II-B.

B. Deep learning algorithms

Deep learning is a class of machine learning that is based on the hierarchical feature learning of the human brain. It is a structured arrangement of artificial neural networks (ANN) that applies automatic feature extraction to learn good representation from raw data [16]. This class of algorithms is unlike the support vector machine (SVM) that we used in [9]. In classical machine learning algorithms, features are manually extracted from data and fed into the machine learning algorithms. This is not the case with deep learning algorithms. In this work, we have considered two deep learning algorithms:

1) Basic Convolutional Neural Network (CNN) - Convolutional neural networks (CNN) are forms of deep learning models that are suited for feature extraction from structured array of data such as images. These features are patterns such as lines, gradients, contours and shapes. The advantage of deep neural networks such as CNN is that they can work raw input image datasets to extract the features they need [15]. Unlike traditional machine learning algorithms where specific features will be extracted and fed into the training algorithm, CNN extracts the required features themselves. Fig. 2 shows the general architecture of a CNN.

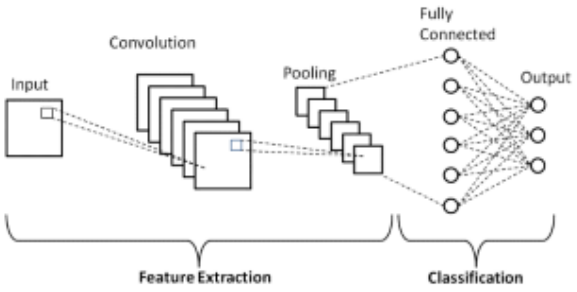


Figure 2: **Basic CNN Architecture**: We employed three convolution layers alternated with 3 pooling layers. The output is either malaria or healthy

A CNN is made of two major parts: a convolution tool and a fully connected layer. Feature extraction occurs at the convolutional layers of the CNN. Based on the features extracted from the convolution process, the fully connected layer predicts the class of the image. Apart from the convolution and fully connected layers, the input layer receives data input as images, the output layer provides the predictions in form of labels or probabilities while the pooling layer reduces the size of the convolved features. There are often more than one pooling layer as each is often placed after each convolution layer. The stacking of layers shows why CNN is called deep learning.

2) MobileNet Version 2 - This is a lightweight CNN model architecture that targets machine learning deployment into mobile devices [17]. The goal of the MobileNet model is to optimise compute power and memory. some accuracy trade-offs resulting from these optimal computational and memory size improvements. The improvement of the Mobilenet architecture from V1 to V2 has some accuracy improvement while still maintaining low memory size and computational cost.

The original images, obtained from the Kaggle Malaria Parasite Detection competition [12], are first used to train each of the deep learning classification algorithms. The images are then watermarked with one and then with four bits of data per 8x8 block. The stego-images generated by the watermarking system are then used to train another model. The same test set (containing some watermarked and non-watermarked images) are used to perform classification using the separate models - from zero watermark, 1-bit watermark and 4-bit watermark. There were 13076 malaria-infected blood smear images and 13054 healthy blood smear images. We then created a fourth dataset by combining the original dataset with half of the images from 1-bit watermarking and from 4-bit watermarking. This gave us a fourth dataset of 25118 healthy images and 25876 malaria images. To train each model the total dataset for each experiment were divided into 70:10:20 for training, validation and testing, respectively.

There are four possible outcomes of a classification or prediction algorithm: True Positive (TP), False Positive (FP), True negative (TN) and False negative (FN). This form what is called a **Confusion Matrix**, C.

$$C = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix}$$

TP is a correctly predicted positive class; FP is a false positive prediction where a subject in negative class is pre-

dicted as being in a positive class. FN is the opposite of FP, where a subject in a positive class is predicted as being in the negative class, while TN is when a negative class is predicted as negative class.

With the confusion matrix, other performance parameters: accuracy, recall (sensitivity), precision and $F_1 - score$ can be defined. The equations that follow define the performance parameters used in this study.

Accuracy is a measure of the ratio of all correct predictions, whether positive or negative, to the entire test set. It is not a good parameter if the number of subjects in each class is not the same.

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{5}$$

In this study, accuracy is the ratio of the sum images correctly predicted as *healthy* and those correctly predicted as having *malaria* to the number of images.

Recall or Sensitivity or True Positive Rate (TPR) is the ratio of correct positive predictions to the total number of the positive class (P).

$$recall = \frac{TP}{TP + FN} \tag{6}$$

This is the ratio of the images correctly predicted as having *malaria* to the total number of people who really have *malaria*.

Precision or Positive Prediction Value (PPV) is the ratio of correct positive predictions to the total positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

In our experiment, this is the ratio of the number of images with *malaria* to the number of those predicted as having *malaria*.

The F-score is a measure of a test's accuracy and it is computed from precision and recall. F1-score is the harmonic mean of precision and recall given as:

$$F_1 - score = 2.\frac{precision \times recall}{precision + recall} \tag{8}$$

$F_1 - score$ do not take true negatives into consideration. Hence, it is a strong measure for the positive class. In this work therefore, it is a strong measure of the ability of an algorithm to truly detect a case of malaria from blood smear images.

With equal class ratio, all of accuracy, precision, recall and F1-score tend to be the same. In such situation, only accuracy can be used to represent the performance of the algorithm on the dataset.

III. Results

A. Steganographic performance

We use PSNR and SSIM to measure the performance of a steganographic algorithm. Figs. 3 and 4 show the PSNR distribution of the healthy and malaria dataset after watermarking.

The PSNR values follows an approximate normal distribution. The lowest PSNR value out among the 26,000 images

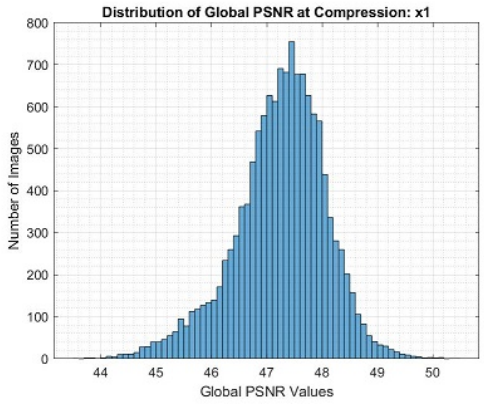


Figure 3: Healthy: Global PSNR Distribution for 1-bit per block

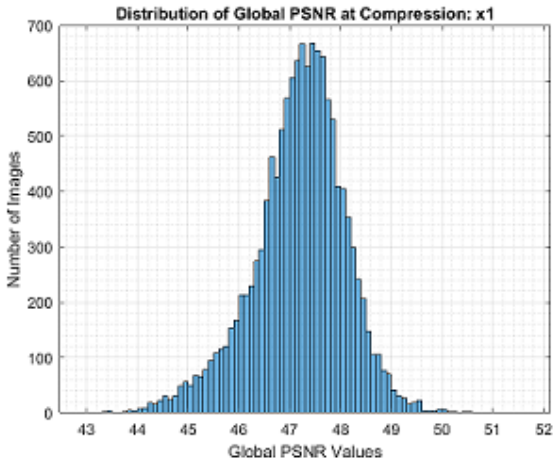


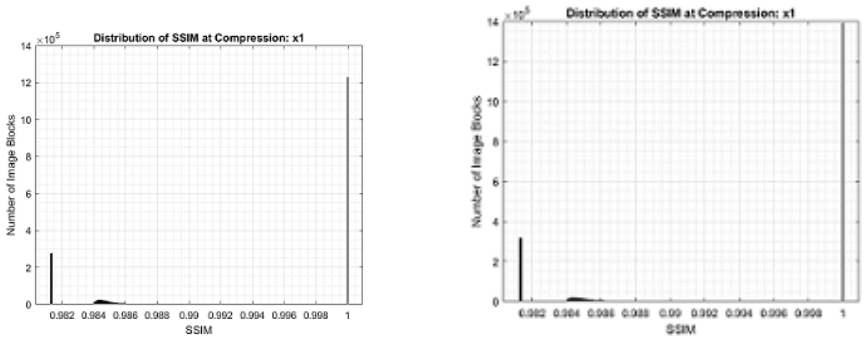
Figure 4: Malaria: Global PSNR Distribution for 1-bit per block

is above 42 dB for both healthy and malaria datasets. The is well above the 40dB benchmark for 8-bit medical images.

With 4-bits of medical data added per 8x8 block, we noticed that the distribution of PSNR reduced to a mean of 32.5 dB as opposed to 47.5dB for the 1-bit embedding capacity per 8x8 block. This is considered a significant degradation for medical images but not for other natural images as the value is still above 30dB.

Figure 5 shows that the SSIM value for individual 8x8 blocks are mostly unity. This means that the structural form of the images were not degraded after embedding. However, some image blocks whose SSIM score are less than unity suffered some degradation. As the SSIM values are all above 0.98, the structural degradation is very minimal.

Not withstanding the interpretation we may give to the distribution of PSNR and SSIM for the watermarked images, the purpose of this work is to find out how they after deep learning performance on original and watermarked medical



(a) Healthy (b) Malaria

Figure 5: Global SSIM Distribution for the Blood Smear Dataset

images. Hence, we focus on the results provided in the next sub-section.

B. Machine learning performance

We trained a total of five models: **Basic CNN Original**, which is a B-CNN architecture trained with original non-watermarked images; **Basic CNN watermarked**, which was trained with 1-bit per 8x8 block watermarked images; **MobileNetV2 Original**, which is a Mobilenetv2 architecture trained with original non-watermarked images; **MobileNetV2 watermarked**, which was trained with 1-bit per 8x8 block watermarked images and **Basic CNN Combined**, which is a B-CNN architecture trained with 50% non-watermarked image, 25% 4-bit watermarked image and 25% 1-bit watermarked image.

The evaluation parameters include accuracy, precision, recall and F1-score. However, the healthy and malaria data sets are equal. This equality means that only the accuracy parameter can be used to represent the other parameters. Fig. 6 uses the accuracy parameter to compare the five models across the four datasets.

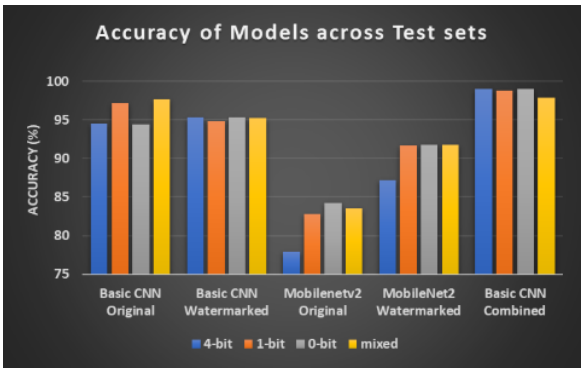


Figure 6: Model Accuracy across test sets: **Basic CNN Combined** trained with 50% non-watermarked image, 25% 4-bit watermarked image and 25% 1-bit watermarked image has the best performance across all test sets.

In Fig. 6, we applied different test sets to the five models (having different architectures and/or train sets). The test sets are: 4-bit per block watermarked images, 1-bit per block watermarked images, original (0-bit) images, and a proportional combination of all the other three data sets. For all the test sets, the **Basic CNN Combined** has the highest accuracy. This model outperformed any of the models trained with either the original or the watermarked images only. It also outperformed all other models in both original and watermarked image test sets.

Fig. 7(a) is a summary of the best model trained with baseline dataset. The baseline dataset is the dataset without any watermark added and without mixture of watermarked and non-watermarked images. The Basic CNN performed better than MobileNetV2 in all cases.

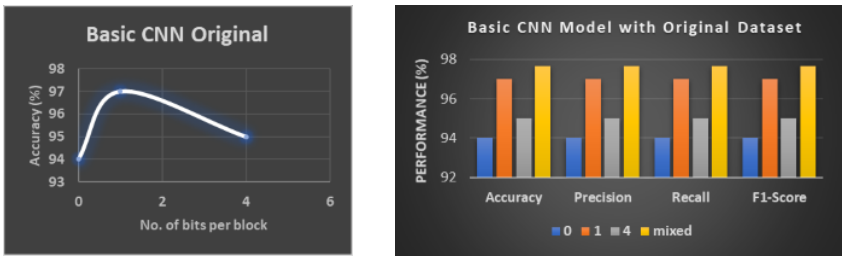
Machine learning model trained with unwatermarked dataset is the currently accepted method of training medical image classification models. Hence, Fig. 7(b) shows the performance of an original Basic CNN model across the four datasets.

Fig. 7(b) shows that *Basic CNN Model trained with original Dataset* performed best on a test set that contains a mixture of data from non-watermarked and differently watermarked images. The second best performance is with 1-bit test data followed by 4-bit test data. The least performance is from non-watermarked test set. This is counter-intuitive with the popular belief that watermarking reduces machine learning accuracy for medical images.

Fig. 8 shows that a model trained with both original and watermarked images will perform slightly better on all test sets than the model trained with original dataset only.

IV. Discussion

In this research paper, we examined the impact of a spread spectrum steganographic algorithm called C_4S algorithm on the performance of deep learning classification models. The aim of this study is to further evaluate the suitability of C_4S algorithm for medical image security in digital Health applications. We have chosen basic CNN and MobileNetV2 deep learning models as they are benchmark models for desktop and mobile machine learning applications respectively.



(a) Embedding capacity vs Accuracy (b) Performance for Individual and Mixed Test sets

Figure 7: Basic CNN trained with original dataset but tested across watermarked, non-watermarked and mixed datasets

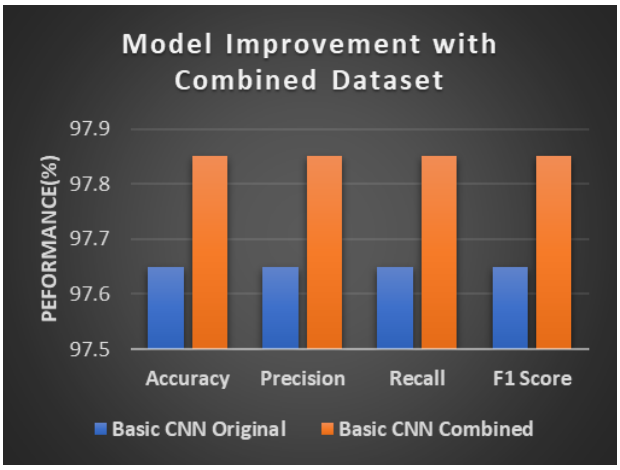


Figure 8: Basic CNN trained with original data set performed better with test images that contains watermarked images

The accuracy and robustness of a machine learning method depends on the method, architecture, training dataset and training hyperparameters. In some cases, pre-processing of the data is introduced to also increase the chance of better performance. In this work, we have applied deep learning method, basic and Mobilenetv2 CNN architectures, malaria blood smear dataset and consistent hyperparameters across all experiments. No image pre-processing or hyperparameter tuning were performed to ensure that variations in performance are only due to steganography (for each deep learning method) and not any other image processing algorithms.

In comparison with traditional machine learning methods such as the support vector machine (SVM), deep learning methods are more robust to steganography. In previous works in [9], [13] where SVM was used to evaluate the C_4S algorithm, the classification decreased from 86.14% before the addition of watermark to 82.18% after the addition of watermark. Recall increased from 82.18% to 85.15%. However, in the case of deep learning, both accuracy and recall increased from 94.5% to 97.65% before and after watermarking, respectively. The overall explanation for higher performance with deep learning is that it can extract more differentiating features between classes than a traditional machine learning methods. Secondly, the increase in performance after watermarking implies that the C_4S acts like a pre-processing algorithm that helps to better differentiate a healthy image from a malaria image. Not all steganographic methods will have this positive effect on deep learning algorithms.

Going further, the results in this work demonstrated that models trained with original non-watermarked data can even perform well or better with watermarked image. For the Basic CNN model trained with original dataset, the accuracy on 1-bit watermarked test set was 97.19% as opposed to an accuracy of 94.36% on an original non-watermarked test. Even for the 4-bit watermarked test, the accuracy reduced to 94.49%, which is still higher than 94.36% with an original test set. Hence, the challenge lies in finding the optimal embedding capacity that will just improve deep learning algorithms rather than degrading it. Figure 7(a) shows that at some embedding capacity greater than 5.5bits per block, the accuracy of the deep learning algorithm may become worse than that of non-watermarked test set. The design process for digital health that involves steganographic information security should include a method of finding this optimal embedding capacity. A framework that helps to achieve this goal was introduced in The framework presented in [19].

There has been numerous assertions that watermarking and steganography would alter the fine-grained features in medical images and reduce the ability of medical personnel to detect biomarkers or symptoms that indicates a sickness [18]. This research shows that such a blanket statement is not the case. Instead medical experts should treat each medical image information hiding algorithm just as a drug discovery process. Algorithms should be taken through "clinical" trials to find out their impact on diagnostic accuracy. Apart from using machine learning evaluation like this work, medical doctors should verify that secret watermarks do not change how humans and other medical equipment read and interpret medical images. Generally, medical image alterations are not recommended, however, we recognize that machine learning algorithms are applicable to pre-processing medical images before diagnosis occurs. The results of this study indicate that steganography and watermarking could be part of signal processing with the extra advantage of its security features

wherever they are applicable.

V. Conclusions

It is a false generalisation to assume that watermarking and steganography reduces the accuracy of machine learning algorithms for medical image classification. However, similar evaluation should be performed for different data-sets to determine their applicability. Furthermore, deep learning architectures also should be experimented on how they are robust to steganography and watermarking. The framework presented in [19] is an attempt to quickly evaluated a dataset on several steganographic algorithms and deep learning algorithms in order to select the best combination of algorithms for a given medical image application. It is recommended that one follows similar framework while designing a medical image security application over an open network and where machine learning algorithms will be run on the transmitted medical image that contains a watermark.

In future works, we intend to evaluate more datasets, more steganographic algorithms and more deep learning architectures. Also, an online evaluation platform will be created to enable developers to easily evaluate their dataset, data hiding algorithms against selected deep learning algorithms before integrating them into an application for digital health delivery.

References

[1] O. M. Al-Quershi and B. E. Khoo. "Authentication and Data Hiding using a hybrid ROI-based watermarking schemes for DICOM images," *Journal of Digital Imaging,* vol24, 1, pp. 114–125, 2011.

[2] P. U Eze. *A Robust and Reliable Tele-medical data Security and Authentication System using Spread Spectrum Steganography*. Ph.D. Thesis, School of Computing and Information Systems, Melbourne School of Engineering, The University of Melbourne, Australia, 2020, pp. 133 - 160. url: https://minerva-access.unimelb.edu.au/handle/11343/253877.

[3] H. Hedieh Sajedi. Applications of data hiding techniques in medical and healthcare systems: a survey. Network Modeling Analysis in Health Informatics and Bioinformatics (2018) 7:6 https://doi.org/10.1007/s13721-018-0169-x.

[4] B. Hassan, R. Ahmed, and O. Hassan. 2019. *An Imperceptible Medical Image Watermarking Framework for Automated Diagnosis of Retinal Pathologies in an e-Health Arrangement*. 7 (2019), 69758 – 69775. https://doi.org/10.1109/ACCESS.2019.2919381.

[5] B. Hassan, R. Ahmed, and B. Li. 'Computer aided diagnosis of idiopathic central serous chorioretinopathy'. In *Proceedings of 2nd IEEE Advanced Management, Communication, Automation Control (IMCEC)*. 824 – 828, 2018.

[6] B. Hassan, R. Ahmed, B. Li, O. Hassan, and T. Hassan. *Automated retinal edema detection from fundus andoptical coherence tomography scans*. In Proceedings of 5th ICCAR , Beijing, China. pp. 1 – 6, 2019.

[7] P. U Eze. *A Robust and Reliable Tele-medical data Security and Authentication System using Spread Spectrum Steganography*. Ph.D. Thesis, School of Computing and Information Systems, Melbourne School of Engineering, The University of Melbourne, Australia, 2020, pp. 76 - 88. url: https://minerva-access.unimelb.edu.au/handle/11343/253877.

[8] J. J. Garcia-Hernandez, W. Gomez-Flores, and J. Rubio-Loyola. Analysis of the impact of digital watermarking oncomputer-aided diagnosis in medical imaging. *Computers in Biology and Medicine* vol.68, 2016, pp. 37–48.

[9] P. U. Eze, U. Parampalli, R.J Evans and D. Liu. "Evaluation of the Effect of Steganography on Medical Image Classification Accuracy". *Journal of Bioinformatics and Computational Biology*, August 2020 In-Press.

[10] M. Fakhredanesh, R. Safabakhsh, and M. Rahmati. 2014. A Model-Based ImageSteganography Method Using WatsonâĂŹs Visual Model.ETRI Journal36, 3(2014), 479–489.

[11] Q. Huynh-Thu and M. Ghanbari (Eds.). 2008. Scope of validity of PSNR in im-age/video quality assessment.Electronics letters44, 13 (2008), pp. 800–801.

[12] Kaggle. Malaria Parasite Detection: parasite detection in thin blood smear image , 2020. retrieved from https://www.kaggle.com/c/malaria-parasite-detection/data

[13] P. U Eze. *A Robust and Reliable Tele-medical data Security and Authentication System using Spread Spectrum Steganography*. Ph.D. Thesis, School of Computing and Information Systems, Melbourne School of Engineering, The University of Melbourne, Australia, 2020, pp. 161 - 197. url: https://minerva-access.unimelb.edu.au/handle/11343/253877.

[14] P. Eze, U. Parampalli, R.J Evans and D. Liu, "Spread Spectrum Steganographic Capacity Improvement for Medical Image Security in Teleradiology." in Proceedings of *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Honolulu, Hawaii, USA*, 17th – 21st July 2018, pp. 766-769. DOI: 10.1109/EMBC.2018.8512344.

[15] S. Karen and A. Zisserman. (2015),*Very Deep Convolutional Networks for Large-Scale Image Recognition*. url: https://arxiv.org/pdf/1409.1556v6.pd

[16] Y. Bengio Deep Learning of Representations for Unsupervised and Transfer Learning. *Workshop on Unsupervised and Transfer Learning. JMLR: Workshop and Conference Proceedings* 27:17–37, 2012.

[17] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.

[18] Chen, Y.-P.; Fan, T.-Y.;Chao, H.-C. WMNet: A Lossless Watermarking Technique Using Deep Learning for Medical Image Authentication. Electronics 2021, 10,932. https://doi.org/10.3390/electronics10080932.

[19] P.U. Eze, U. Parampalli, R. Evans and D. Liu, "A New Evaluation Method for Medical Image Information Hiding Techniques," *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 6119-6122, doi: 10.1109/EMBC44109.2020.9176066.