JKAU: Comp. IT. Sci., **Vol**. **13** No. 2, pp: 86 – 117 (2024 A.D.)
DOI: 10.4197/Comp.13-2.7

Optimizing Federated Learning for Medical Image classification: A Comparative Study of Pre-Trained Models on Compressed X-ray Imagery

Sawsan ALwadaie[1], Amani Jamal[2], Samar Alkhuraiji[3], and Lamiaa Elrefaei [4]

[1 2 3]*Department of Computer Science, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia*
[4]*Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University,Cairo , Egypt*
Salwadaie@stu.kau.edu.sa, Atjamal@kau.edu.sa, Salkhuraiji@kau.edu.sa,
lamia.alrefaai@feng.bu.edu.eg

Abstract— Machine learning, particularly deep learning, has revolutionized a number of fields, including medical diagnostics. In this study, federated learning FL is employed to address privacy concerns and data access limitations inherent in medical imaging. A simulated FL environment was used to investigate the performance of five pre-trained neural network models: DENSENET121, RESNET18, VGG-NET11, GOOGLENET and INCEPTION-V3. It emphasizes the optimization of training duration as well as the application of lossy image compression techniques such as JPEG in order to improve communication efficiency. We conducted a comparative analysis of the models' performance before and after image compression by evaluating the Area Under the Receiver Operating Characteristic Curve and the training time. According to the results, image compression can maintain or improve model performance while affecting training time, underscoring the trade-offs between model accuracy and computational efficiency.

Keywords—federated learning, CheXpert, Simulated Environment, JPEG Algorithm.

I. INTRODUCTION

Advancements in machine learning and deep learning DL over the past decade have substantially impacted various sectors, particularly medical applications. These technologies have revolutionized the medical field by improving disease detection and diagnosis through medical imaging, assisting radiologists in the early and accurate identification of conditions such as pneumonia and other lung diseases, and ultimately expediting patient recovery [1] [2] [3].

A key evolution in DL is the advent of pre-trained convolutional neural networks CNNs, known as transfer learning TL. This approach utilizes networks, such as AlexNet, ResNet18, and DenseNet, which are trained on extensive datasets, such as ImageNet. In TL, a model trained for a specific task is repurposed for a different DL task using the same weights [4][5] [6]. However, the application of DL in critical areas, such as healthcare, where accuracy is paramount, presents challenges. Deep neural networks in the medical domain require vast image datasets, which are difficult to acquire due to patient privacy and data confidentiality concerns [7].

Federated learning FL offers a solution to these challenges by enabling access to data from various devices while preserving data ownership. FL enables collaborative model training without sharing individual device data, ensuring privacy [8]. In FL, numerous communication rounds

occur, where each round involves data processing on multiple FL clients, such as smartphones and computers. These clients participate in training and then transmit their results to a central server, forming a global model. In a simulated FL environment, a single dataset is divided into multiple subsets, each representing a client. Training occurs on these subsets in multiple rounds, and the results are transmitted to a central server to update the global model [9].

Studies, such as [10], have explored FL in medical image classification, emphasizing privacy and addressing data constraints. Similarly, research by [11] explored the use of FL with different numbers of clients and training epochs, using models such as ResNet50 and DenseNet121 on chest X-ray datasets, achieving accuracy rates up to 92%.

Our research contributes to this field by investigating compression techniques within FL on X-ray images. Other studies have focused on neural network compression (NNC) by MPEG (ISO)[12] and the Parameter Compression Method [13]. These techniques aim to compress neural network models and enhance data privacy, respectively. In contrast, our study introduces lossy JPEG image compression to optimize medical X-ray image classification. This method has proven effective in CNN model training [14] and plays a key role in minimizing communication overhead in FL.

This study adopts a novel approach by implementing JPEG lossy image compression within a simulated FL environment. Our objective is to reduce communication costs during data transmission by applying the compression algorithm to X-ray dataset images. We have implemented and analyzed various pre-trained CNN models with JPEG compression in this simulated FL setting, training five different models (DenseNet121, ResNet18, VGG-Net11, GoogLeNet, and Inception-V3) on X-ray image datasets. Our research evaluates their performance before and after applying JPEG compression, marking a significant advancement in FL, especially in medical imaging, by optimizing model performance while addressing efficiency and privacy challenges. In our research paper, we systematically explore the application of federated learning to the field of medical imaging, particularly X-ray image classification. This introductory section has been used to set the stage for the study by highlighting the advancements and significance of machine learning in medical diagnostics. The remainder of this paper is structured into additional sections as follows. The Literature Review section is subdivided into two subsections. The Background sub-section explores the concept of federated learning, its challenges, and types, and is followed by the Related Work sub-section, which reviews current methodologies in the field. The Materials section describes the CheXpert dataset used for model training, while the section, Experimental Setup, outlines the procedural framework of our study, including the training of pre-trained models on both original and compressed images. The Results section presents the findings of our model evaluations, and the following section, Discussion, interprets these results in the context of federated learning frameworks. The paper culminates in a final section, Conclusion, that encapsulates the study's findings and its implications for future medical imaging practices. All sources are cited in the References section, and detailed results are relegated to the Appendix.

II. LITERATURE REVIEW

This literature review delves into two critical areas: the foundational concepts and challenges of FL (addressed in the Background section), and pivotal advancements and applications (detailed in the Related Work section). This review serves to frame the current landscape of FL, setting the stage for our study's focus on medical image analysis.

A. BACKGROUND

This section delves into the intricacies of FL, a collaborative machine-learning paradigm designed to maintain data privacy across diverse applications. It explores FL's evolution, challenges, and its integration with differential privacy, highlighting its significance in the realm of secure, decentralized data processing.

1) Federated Learning:

FL is a collaborative machine learning approach that enables the use of multiple devices for training a shared model while ensuring that the privacy of each data source is maintained. It was introduced by Google in 2017 to enhance voice recognition while preserving user privacy [15][8]. The process involves a central server distributing an initial model to local devices, which then train the model using their data. The locally trained models are aggregated into a new model on one central server. FL has been applied in healthcare, finance, and telecommunications, where data privacy is crucial [16] [17].

2) Challenges In Federated Learning:

FL is a promising approach in machine learning; however, it faces several challenges. One of the most common challenges is non-independent and identically distributed data (non-IID), where different devices have varying data distributions, leading to performance degradation of the global model, and causing a mismatch between the local and global data [18]. Communication efficiency is another significant challenge, particularly when FL is applied in cases with a large number of devices with limited communication resources, resulting in issues related to communication efficiency [15]. Furthermore, the devices within the FL framework may differ in computational capabilities, battery life, and data storage, posing challenges in designing algorithms that can accommodate this heterogeneity [17]. Personalized data privacy and Personalized data privacy and security are of paramount importance in FL, necessitating appropriate encryption techniques and mechanisms to ensure privacy preservation and security [19]. Finally, FL models can be vulnerable to attacks by malicious nodes that deploy poisoned updates or leakage-prone updates, which can lead to the entire system malfunctioning [20]. These challenges must be thoroughly considered and carefully addressed to effectively implement federated learning systems in real-world applications.

3) Differential Privacy in Federated Learning:

In the realm of FL, safeguarding data privacy is a foundational concept, as it ensures collaborative model training without compromising sensitive information. In [7], the authors explore differentially private FL for protecting data privacy in the binary classification of medical image (chest X-ray) data. They compare two neural network architectures and observe that non-private models achieve high accuracy but are susceptible to privacy attacks through image reconstruction. They mitigate this risk by integrating a privacy mechanism based on Renyi differential´ privacy with a Gaussian noise mechanism during local model training. This approach helps enhance privacy protection while maintaining reasonable classification accuracy. In [21], the authors introduce a privacy-preserving solution for collaborative machine learning among data-owning entities, such as hospitals. Their study addresses the challenge of malicious parties trying to disrupt the model. The proposed approach, distance-based outlier suppression (DOS), calculates distances between local model updates from different clients, detects outliers using copula-based outlier detection (COPOD), and uses weighted averages to update the global model. DOS is robust against various poisoning attacks, making it suitable for medical imaging datasets, such as CheXpert and HAM10000. Table 1 illustrates the strategies employed to effectively mitigate privacy concerns within the context of federated learning. Researchers employ a variety of datasets to investigate methods that enhance privacy. The primary objective is to achieve a harmonious equilibrium between safeguarding data privacy and maximizing the efficacy of FL, thereby providing significant contributions to the field of privacy-preserving collaborative machine learning. FL can be applied in simulation environments to verify model performance and functionality [7]. It is also implemented in real-world settings, including devices, cloud, and IoT environments, to improve machine learning models while maintaining privacy [16]. FL uses

Table I: OVERVIEW OF RECENT STUDIES ON PRIVACY MEASURES IN FL: THIS TABLE SUMMARIZES KEY RESEARCH FROM 2021

Year	Ref.	Datasets	Privacy Methods
2022	[21]	CheXpert HAM10000	Method (DOS) designed to enhance the robustness and privacy of federated learning systems by effectively identifying and suppressing malicious behavior
2022	[7]	CheXpert Mendeley	They apply Differential Privacy to various neural network architectures in federated learning for chest X-ray classification, assessing the trade-off between performance and privacy across different privacy budgets
2021	[30]	EMNIST	This work involves adding constraints related to Differential Privacy (DP) to a specific approach used in Federated Learning (FL) called DP-SCAFFOLD. These constraints are designed to enhance the privacy of the data used in the FL process
2021	[31]	MNIST CIFAR	They present FEDMD-NFDP ensures the privacy of the data used in the federated learning process. It incorporates Noise-Free Differential Privacy (NFDP) mechanisms to protect data privacy

various global model algorithms, including federated averaging [8], FedProx [22], federated MDL (model and data level) [16], and federated meta learning [23].

4) Types of federated learning:

FL has attracted extensive research since its inception, leading to various proposed approaches in the literature. These approaches may be generally categorized into five types and are presented here in two groups.

The first group of approaches addresses methods for data sharing between different sources: vertical federated learning, horizontal federated learning, and federated transfer learning. In a vertical FL approach, different entities (such as hospitals) hold different features or attributes for the same set of entities (for example, patients). This method is used when collaborating entities have different types of data about the same individuals. For example, multiple hospitals may contribute their medical images to assist in disease diagnosis. Each entity contributes different features to create a more comprehensive understanding [24].

In horizontal FL, different entities have data on different individuals, but the types of data are similar. For example, multiple banks may have similar types of customer data (such as spending habits and account balances) but for different sets of customers. In horizontal FL, these data are combined in a way that avoids revealing individual customer data from each bank [24].

Federated transfer learning [16] combines the principles of TL with FL. TL is a method where a model developed for one task is reused as the starting point for a model that will be used on a second task. The federated transfer learning approach is used to improve model performance and reduce the need for extensive communication between the participating entities. It is particularly useful in scenarios where some entities have limited data or when it is necessary to enhance the learning process by leveraging pre-trained models [25]. Each type is suited to different scenarios based on the nature of the data available and the privacy requirements of the participating entities.

The second group addresses methods that are based on the type of dataset used in federated learning and encompasses several applications based on data type.

Structured data, such as tabular data, transaction data, sensor data, and log files, are used in federated learning on structured data [26]. Image data, including photographs, medical images, and satellite images, are utilized in federated learning on image data [27]. Additionally, text data (for example, from emails, chat logs, social media posts, and news articles) serve as input for federated learning on text data [28]. Lastly, time-series data, such as temperature readings or machine vibrations collected over time, are employed in federated learning on time-series data [29]. There are two types of federated learning from the standpoint of data and client diversity. The first is homogeneous federated learning, where data from multiple devices or users within a single group are utilized, and these devices or users exhibit similar characteristics or data distributions [7].

The second type is heterogeneous federated learning, which involves data from multiple devices or users from different groups, and these devices or users may have varying characteristics or data distributions [30].

In summary, in this section, we have discussed various types of FL approaches across different categories, but our focus in this study is federated transfer learning using a homogeneous X-ray image data set and a pre-trained CNN. Our objective is to train five pre-trained models (DenseNet121, RES-NET18, Table 1 overview of Recent Studies on Privacy Measures in FedVGG-Net11, GoogLeNet, and Inception-v3) on an X-ray FL the table summarizes key research from 2021 image dataset in a simulated federated learning environment. and 2022, detailing the datasets used, the privacy measures implemented, and their specific applications in enhancing data privacy in FL systems. accuracy, ensuring data security, and optimizing resource usage in distributed learning environment.

B. RELATED WORK

This section reviews significant contributions in the field of FL, particularly focusing on its application in medical image analysis and compression techniques. The studies discussed below highlight the progress in enhancing model accuracy, ensuring data security, and optimizing resource usage in distributed learning environments.

1. Enhancing Medical Image Analysis Through Federated Learning:

Recent advancements in federated learning applied to medical image analysis are notable for their innovation and impact on healthcare diagnostics. For example, the work discussed in [32] introduces FedFBN, a novel FL framework that improves the classification of medical images across non-identically distributed datasets with incomplete labels. This framework achieves an AUROC of 0.75 on the CheXpert dataset by integrating TL and static batch normalization. Similarly, the study in [33] utilizes FL and TL to enhance breast cancer classification, achieving a remarkable 98% accuracy with a hybrid model comprising FeAvg-CNN and MobileNet across various datasets. Additionally, [34] explores a blockchain-enhanced FL method for training models on heterogeneous, multi-class respiratory medical datasets, achieving up to 88.10% test accuracy. This approach not only secures data privacy and integrity but also innovates with a weight manipulation technique based on local model test accuracy, facilitating secure and transparent collaboration among medical institutions.

2. Advanced Model Compression Techniques in Federated learning environment

Investigations of model compression within federated learning frameworks have addressed the dual challenges of heterogeneous settings and privacy protection. Researchers in [35] discuss an adaptive in-parallel pruning-quantization method that significantly reduces model sizes and transmission times without compromising performance.
This method employs mutual learning and multi-teacher knowledge distillation, enhancing efficiency in resource-constrained environments. The study discussed in [36] details a model compression strategy within an over-the-air federated learning

(OTA-FL) framework, employing pruning and quantization-aware training to reduce model sizes by up to 80%, while maintaining nearly original performance under limited bandwidth and strict resource constraints. Moreover, introduces FedCGS, a framework that integrates conditional generative adversarial networks with singular value decomposition. This method effectively reduces communication loads and minimizes privacy risks, tested on the FMNIST and CIFAR10 datasets to demonstrate competitive accuracy and superior communication efficiency.

3. Addressing Image Compression in Federated Learning

Despite extensive research in federated learning, the application of image compression techniques within this context has been largely unexplored. This study aims to bridge this gap by applying JPEG compression [14] to X-ray images within a federated learning framework. In this work, various experiments employing pre-trained neural network models are conducted to assess the efficacy of image compression in enhancing data transmission efficiency while maintaining model performance in medical diagnostics.

Table 2 encapsulates key contributions to federated learning in medical imaging, where studies in 2023 [32]-[34] focus on frame works for knowledge aggregation, privacy, and accuracy in diagnosis. Subsequent 2024 studies [35] -[37] pivot to model compression, enhancing efficiency and maintaining performance across various datasets.

Our study introduces JPEG image compression to federated learning, targeting X-ray image classification on the CheXpert dataset with a comprehensive suite of performance metrics.

III. METHODOLOGY

This section outlines our research methodology and is subdivided into a Materials section, where we describe the data and tools used, and an Experimental Setup section, which the procedures and configurations for our analyses. This structure ensures a clear understanding of how our study was conducted.

A. MATERIALS

This section provides a detailed description of the dataset's structure, usage, and the specific preprocessing techniques applied to ensure data integrity and accuracy.

1. Dataset Description

The CheXpert dataset [38] encompasses a total of 224,316 radiographic images from 65,240 distinct patients. For our study, we utilized 70% of this dataset, corresponding to 156,390 X-ray images, for training. Additionally, 10% of the dataset, or 22,342 X-ray images, were designated for testing, while the remaining 20%, which included 44,683 images, were reserved for validation in all conducted experiments. Each radiograph in the dataset has been annotated with one or more classifications from a set of 14 possible categories: "No Finding", "Enlarged Cardiomediastinum", "Cardiomegaly", "Lung Opacity", "Lung Lesion", "Edema", "Consolidation", "Pneumonia", "Atelectasis", "Pneumothorax", "Pleural Effusion", "Pleural Other", "Fracture", and "Support Devices". Our study involves the classification of images across all these specified labels.

Table II. Overview of Federated Learning Contributions in Medical Imaging: Comparative Analysis of Recent Studies Highlighting

Ref.	Year	Contribution	Dataset	No. Of Classes	Matric Performance	Compression Image Algorithms Or Compressing Model algorithms
[32]	2023	introduces FedFBN, a federated learning framework that enhances the aggregation of knowledge from distributed non-iid datasets with partial labels by using pretrained networks and freezing batch normalization layers during training.	Chexpert NIH MIMIC	Binary classes	binary cross-entropy loss	N
[33]	2023	integrating transfer learning with federated learning to improve breast cancer classification, achieving enhanced privacy with the FeAvg-CNN + MobileNet model on diverse datasets.	Breast Cancer	9	accuracy, recall, F1-score and the AUC	N
[34]	2023	introduces a FL mechanism that effectively trains and aggregates models on multi-class and heterogeneous respiratory medical data using blockchain technology for enhanced privacy and introduces a weight manipulation technique based on local model test accuracy	COVID-19 Pneumonia	6	Accuracy Precision Recall F1-scor	N
[35]	2024	introduces an adaptive in-parallel pruning-quantization method that enhances model compression and maintains accuracy in heterogeneous federated learning environments.	Various datasets	Not mentined	accuracy rates compression effectiveness	Model Compression
[36]	2024	proposes a compression pipeline combining pruning and quantization-aware training to significantly reduce the model size and computation requirements in an Over-the-Air Federated Learning (OTA-FL) system while maintaining comparable accuracy	CIFAR	Not mentinesd	accuracy	Model compression
[37]	2024	introduces the FedCGS framework that uses conditional GANs and feature compression via singular value decomposition to enhance privacy and communication efficiency in FL.	FMNIST CIFAR10	10	Accuracy communication efficiency	Model parameters compression
This paper	2024	Employing JPEG compression algorithms within a simulated federated learning environment for the classification of X-ray images across 14 labels.	CheXpert	14	Accuracy, recall, F1-score and the AUC Classification report	JPEG image compression

2. Preprocessing

We utilized a data loader to import the CheXpert dataset, which processes the labels and image paths from the CSV files provided. The data loader also addresses label uncertainty by categorizing labels into known pathologies (1), absence of pathologies (0), uncertain cases (-1), and unknown cases (”). We chose to simplify this by adopting a “U-zeroes method”, treating all uncertain cases as an absence (0), thereby resolving ambiguities in label interpretation.

B. EXPERIMENTS SETUP

Our research methodology was executed within a simulated federated learning framework, utilizing the Google Colab Pro+ environment to conduct several experiments.

1. Stage one: optimizing the training duration for our models

In the process of optimizing the training duration for our models, we initially focused on the VGG-Net11 architecture. Given that training could extend to an excessive 20 hours, we experimented with varying the number of epochs, rounds, and clients specifically for this model. The insights gained from these trials were used to inform our approach to other pre-trained models.

The results indicated that an increase in epochs adversely affected the AUC accuracy; hence we standardized the epoch count to three across all models. Since the essence of federated learning hinges on utilizing a substantial number of clients, we maintained a minimum of five clients per model. Interestingly, we observed that a reduction in the number of clients could potentially enhance AUC accuracy.

Consequently, we established two primary experiments based on our preliminary findings:

- Investigating pre-trained models without compression, using five clients and ten rounds.
- Implementing all pre-trained models with JPEG lossy compression, also with five clients and ten rounds.

The outcomes of these experiments formed the basis for our analysis. Table 3 substantiates the narrative described above, presenting data from the CheXpert dataset with various configurations of the VGG-Net11 model. It shows the relationship between the number of clients, rounds, epochs, and the resultant AUC scores.

2. Stage two: Training on Original X-ray Images

- Preliminary experiments:

  Our research undertook extensive experimentation on a range of pre-compression pre-trained models to discern the influence of various numbers of rounds and clients on model performance. It was observed that a greater number of rounds positively correlates with AUC accuracy. Consequently, we planned to scale the number of rounds up to ten. This scaling was constrained by the operational limits of Google Colab Pro+, which ceases training at the 24-hour mark, leading us to adjust the number of rounds accordingly. These preliminary experiments, which we refer to as “pre-experiments” served as the foundation for our more extensive, substantive work. The data presented in Table 4 corroborate our findings. The experiments were carried out on the CheXpert dataset using different pre-trained models with a fixed number of clients (five) and epochs (three), and the AUC accuracy was recorded. The VGG11 model achieved an AUC of 0.88, Inception v3 reached 0.79, ResNet18

scored 0.83, DenseNet121 obtained 0.80, and Google Net recorded 0.78. These results support our decision to adjust the number of rounds in our next experiments**.**

- Primary experiments:

Our experiments in this section involved training five distinct pre-trained neural network architectures (DenseNet121, RES-NET18, VGG-Net11, GoogLeNet, and Inception-v3) on the original CheXpert dataset. These models, which had been previously trained on the ImageNet dataset, were fine-tuned on X-ray images distributed across five simulated clients. This setup aimed to establish baseline performance metrics for each model without any modifications to the images.

3. Stage three: Training on Compressed X-ray Images:

   In the third stage, we applied a JPEG compression algorithm to the X-ray images before training the same neural networks. This step introduced a quality setting of 60 using the Python Imaging Library (Pillow), converting each image to grayscale and compressing it using the JPEG format. The "optimize = True" parameter was also employed to potentially further reduce file sizes. This compression aimed to assess the impact on the models' performance, facilitating a direct comparison with the results from uncompressed images.

4. Federated Learning Configuration and Parameters:

   We consistently configured the federated learning environment across both stages using the following parameters: a learning rate ("lr") of 0.0001, Adam optimizer betas ("betas") at (0.9, 0.999), epsilon ("eps") at '1e-08', and a weight decay ("weight decay") of 0. The dataset-specific configurations such as batch size ("trBatchSize"), number of classes ("nnClassCount"), and maximum epochs ("trMaxEpoch") were also set. Each client processed 31278 images over 10 communication rounds ("com rounds"), ensuring comprehensive participation in the learning process.

5. Training Procedure and Evaluation:

   The training procedure comprised local training on client datasets and subsequent global model aggregation via federated averaging. This process was devoid of the complexities of secure data transfer, as the federated learning was simulated, and data privacy was intrinsically maintained. Performance evaluation was implemented using "sklearn.metrics" along with the "classification report" function to derive key metrics such as accuracy, precision, recall, F1-score, and AUROC. Despite the constraints of Google Colab Pro+, which limits continuous operation to 24 hours, we achieved a high accuracy rate of 99.95% AUC. The AUC metric was particularly chosen for its comprehensive ability to measure the model's discriminative power at various threshold levels without being affected by class imbalance, offering a singular performance measure applicable for comparative analysis and essential for real-world deployment.

Table III: displays the results of applying the VGG model .

Dataset	Pre-trained model	No. of clients	No. of rounds	No. of epochs	AUC
CheXpert	VGG-Net11	5	2	10	0.76
		2	2	3	0.90
		5	1	3	0.79
		2	4	3	0.97
		5	10	3	0.98

Table IV: This table presents the AUC scores for various pre-trained models

Data set	Pre-trained Models	FL parameter		AUC
		No. of clients	No. of rounds	
CheXpert	VGG11	5	2	0.88
	Inception v3	5	2	0.79
	ResNet18	5	2	0.83
	DeneNet121	5	2	0.80
	GoogleNet	5	2	0.78

IV. RESULTS

In this section, we present the performance outcomes of our study, which evaluates the effectiveness of five different pre-trained neural network models under two distinct conditions: before and after the application of image compression. The results are divided into two main categories, uncompressed and compressed image results, to distinctly illustrate the impact of compression techniques on model performance and training efficiency. This comparison sheds light on the trade-offs between accuracy and training duration, providing a deeper understanding of the influence of image compression on federated learning models in the context of medical image classification. We discuss these findings with an emphasis on the balance between maintaining high diagnostic accuracy and optimizing computational resources.

1. Uncompressed Image Results:

- VGG11: Achieved an AUC of 0.98, with a training time of 18 hours.

-Inception v3: Reached an AUC of 0.97, taking the same amount of time as VGG11, which is 18 hours.

-GoogLeNet: Obtained a lower AUC of 0.93, but was significantly faster, requiring only 7 hours.

-DenseNet121: Reported an AUC of 0.96 with a training time of 15 hours.

-ResNet18: Matched the highest AUC of 0.98 among the models and was the quickest, with a training time of only 7 hours.

2. Compressed Image Results:

-VGG11: After compression, the AUC slightly increased to 0.99, but the model required more time to train, totaling 20 hours.

-Inception v3: Maintained an AUC of 0.97 post-compression and saw a reduction in training time by 3 hours, making it 15 hours.

Table V: Comparative Performance of Pre-Trained Models on CheXpert Dataset Before and After Applying 60 % JPEG Compression.

CheXpert dataset with FL with 5 clients and 10 rounds						
Pre rained models		VGG11	Inception v3	GoogleNet	DenseNet121	ResNet18
Before compress the images	**AUC**	0.98	0.97	0.93	0.96	0.98
	No. of hour	18	18	7	15	7
After compression	**AUC**	0.99	0.97	0.93	0.97	0.98
	No. of hour	20	15	8	16	8

- GoogleNet: Showed no change in AUC, remaining at 0.93, with a slight increase in training time to 8 hours.

-DenseNet121: Had a slight improvement in AUC to 0.97 and an additional hour of training, totaling 16 hours.

-ResNet18: Sustained its AUC at 0.98 and experienced a minor increase in training time to 8 hours.

These results provide insights into the trade-offs between model accuracy and training efficiency. The data suggest that while compression can enhance or maintain the discriminative power of certain models (with VGG-Net11 showing improvement and Inception v3 and DenseNet121 showing stable performance), it may also lead to increased training times for some models. This highlights the importance of considering both the quality of predictive performance and computational resources when deploying federated learning models in practice.

Table 5 provides a comparative analysis of the performance of the five pre-trained neural network models using the CheXpert dataset within an FL setup with 5 clients and 10 rounds. The performance of the models is evaluated based on the AUC and the time taken (in hours) for the models to train before and after image compression at a 60% compression ratio.

Furthermore, we generated classification reports to provide detailed performance measures of our models; these reports offer insights into accuracy, precision, recall, and F1 scores for each class within the dataset. These reports were instrumental in comparing the models' performance across the two experimental stages (before and after the application of image compression) in the context of federated learning. For an in-depth examination of these performance metrics, comprehensive tables and figures have been included in the Appendix.

V. DISCUSSION

The discussion delves into the implications of these findings, considering both the model performance and the practical aspects of deploying such models in real-world scenarios. Table 6 presents a comparison of techniques and results from various studies, including ours, to contextualize the advancements our approach offers.

A. Impact of Image Compression on Model Performance:

Our study demonstrates that image compression at a 60% ratio can variably affect the performance of different pretrained models.

Notably, VGG-Net11 showed a remarkable improvement in AUC from 0.98 to 0.99 after compression, suggesting that some models may extract more relevant features from compressed images. Conversely, GoogLeNet and ResNet18 maintained their AUC, indicating resilience to the loss of information typically associated with compression. This highlights the potential of certain models to operate efficiently even with reduced data quality, an essential consideration in bandwidth-limited environments.

B. Training Efficiency versus Accuracy:

An intriguing observation from our results is the trade-off between training efficiency and model accuracy. While ResNet18 and GoogLeNet were quick to train, their AUCs were lower compared to VGG-Net11 and Inception-v3, which took longer. This trade-off is crucial in settings where time and computational resources are limited. For instance, in medical imaging diagnostics, where rapid results are particularly important, a balance between accuracy and speed is paramount.

C. Federated Learning Context:

In the context of federated learning, where data privacy and decentralized learning are prioritized, our results provide a nuanced perspective. The variations in training times and AUCs post-compression suggest that different models may be more suitable for different federated learning scenarios. Models with shorter training times and relatively high accuracy, such as ResNet18, might be preferable in time-sensitive settings, while those with higher accuracies but longer training times could be reserved for scenarios where precision is more critical.

D. Implications for Model Selection in Practice:

The findings from this study underscore the importance of model selection based on specific use-case requirements. For practitioners and researchers, this implies a need to weigh the pros and cons of each model in relation to the specific constraints and objectives of their federated learning setup.

E. Limitations and Considerations:

While our study provides valuable insights, it is important to acknowledge its limitations. The exclusive use of the CheXpert dataset and a fixed compression ratio could potentially limit the broader applicability of our results. Future research seeking to enhance the robustness and generalizability of our findings should incorporate a variety of datasets and explore different compression ratios. Additionally, our experiments were conducted using Google Colab Pro+, which imposes a constraint of a 24-hour runtime limit, potentially affecting the extent of our experimental exploration.

VI. CONCLUSION

This study's investigation of the effects of image compression on various pre-trained neural network models within a federated learning framework, using the CheXpert dataset, reveals significant variations in model performance. Key findings indicate a crucial trade-off between training efficiency and accuracy, notably with some models, such as VGG-Net11, showing improved accuracy post-compression. These insights are vital for federated learning applications, especially in resource-constrained environments such as healthcare. While highlighting the importance of tailored model selection, the study also opens avenues for future research, particularly in extending the analysis to diverse datasets and exploring varied compression techniques. Ultimately, this research underscores the complex balance between efficiency and accuracy that is

necessary for effective federated learning implementations.

Table VI: Performance Comparison of Different Learning Techniques

Ref.	Techniques	Results
[3]	DL	97%
[39]	TL	97%
[40]	DL	96%
[7]	FL	94%
This paper	FL	99.95%

ABBREVIATIONS

1. DL - Deep Learning
2. CNN - Convolutional Neural Networks
3. CNN - Convolutional Neural Networks
4. TL - Transfer Learning
5. FL - Federated Learning
6. JPEG - Joint Photographic Experts Group
7. AUC - Area Under the Curve
8. Non-IID - Non-Independent and Identically Distributed
9. DOS - Distance-based Outlier Suppression
10. COPOD - Copula-based Outlier Detection
11. FL - Federated Learning
12. DP - Differential Privacy
13. NFDP - Noise-Free Differential Privacy
14. ISO - International Organization for Standardization
15. AUROC - Area Under the Receiver Operating Characteristic Curve

REFERENCES

[1] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708, 2017.

[2] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan,D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., "Chexnet: Radiologist-level pneumonia detection on chest xrays with deep learning," arXiv preprint arXiv:1711.05225, 2017.

[3] J.Irvin, P.Rajpurkar, M.Ko, Y.Yu, S.Ciurea-Ilcus, C.Chute,H. Marklund,B. Haghgoo, R. Ball, K. Shpanskaya, et al., "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proceedings of the AAAI conference on artificial intelligence, vol. 33, pp. 590–597, 2019.

[4] K. T. Islam, S. N. Wijewickrema, A. Collins, and S. J. O'Leary, "A deep transfer learning framework for pneumonia detection from chest x-ray images.," in VISIGRAPP (5: VISAPP), pp. 286–293, 2020.

[5] T. Rahman, M. E. Chowdhury, A. Khandakar, K. R. Islam, K. F. Islam, Z. B. Mahbub, M. A. Kadir, and S. Kashem, "Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray," Applied Sciences, vol. 10, no. 9, p. 3233, 2020.

[6] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., "Imagenet large scale visual recognition challenge," International journal of computer vision, vol. 115, pp. 211– 252, 2015.

[7]J.Ziegler,B.Pfitzner,H.Schulz, A.Saalbach,and B.Arnrich, "Defending against reconstruction attacks through differentially private federated learning for classification of heterogeneous chest x-ray data," Sensors, vol. 22, no. 14, p. 5195, 2022.

[8] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial intelligence and statistics, pp. 1273–1282, PMLR, 2017.

[9] R. Kumar, W. Wang, C. Yuan, J. Kumar, H. Qing, T. Yang,A. A. Khan, et al., "Blockchain based privacy-preserved federated learning for medical images: a case study of covid-19 ct scans," arXiv preprint arXiv:2104.10903, 2021.

[10] A. Fettah, R. Menassel, and A. Gattal, "Machine learning for medical image analysis: A survey," in International Conference on Advanced Intelligent Systems for Sustainable Development, pp. 148–164, Springer, 2022.

[11] A. E. Cetinkaya, M. Akin, and S. Sagiroglu, "A communication efficient federated learning approach to multi chest diseases classification," in 20216 the International Conference on Computer Science and Engineering (UBMK), pp. 429–434, IEEE, 2021.

[12] S. A. E. M. Nasri, I. Ullah, and M. G. Madden, "Compression scenarios for federated learning in smart manufacturing," Procedia Computer Science, vol. 217, pp. 436–445, 2023.

[13] Y. Zhu, H. Cao, Y. Ren, W. Wang, B. Wang, M. Hu, B. Cai,W. Wang, et al., "Defending privacy inference attacks to federated learning for intelligent iot with parameter compression," Security and Communication Networks, vol. 2023, 2023.

[14] R. D. Evans, L. Liu, and T. M. Aamodt, "Jpeg-act: accelerating deep learning via transform-based lossy compression," in 2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA), pp. 860–873, IEEE, 2020.

[15] J. Konecny`, H. B. McMahan, F. X. Yu, P. Richtarik, A. T.´ Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, vol. 8, 2016.

[16] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 2, pp. 1–19, 2019.

[17] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated learning for mobile keyboard prediction," arXiv preprint arXiv:1811.03604, 2018.

[18] S. Reddi, Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Konecny`, S. Kumar, and H. B. McMahan, "Adaptive feder-ˇ ated optimization," arXiv preprint arXiv:2003.00295, 2020.

[19] X. Yin, Y. Zhu, and J. Hu, "A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions," ACM Computing Surveys (CSUR), vol. 54, no. 6, pp. 1–36, 2021.

[20] P. Liu, X. Xu, and W. Wang, "Threats, attacks and defenses to federated learning: issues, taxonomy and perspectives," Cybersecurity, vol. 5, no. 1, p. 4, 2022.

[21] N. Alkhunaizi, D. Kamzolov, M. Taka´c, and K. Nandaku-ˇ mar, "Suppressing poisoning attacks on federated learning for medical imaging," in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 673–683, Springer, 2022.

[22] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," Proceedings of Machine learning and systems, vol.2, pp. 429–450, 2020.

[23] Y. Di and Y. Liu, "Mfpcdr: A meta-learning-based model for federated personalized cross-domain recommendation," Applied Sciences, vol. 13, no. 7, p. 4407, 2023.

[24] J.Xu, B.S.Glicksberg, C.Su, P.Walker, J.Bian, andF.Wang, "Federated learning for healthcare informatics," Journal of healthcare informatics research, vol. 5, pp. 1–19, 2021.

[25] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "Fedhealth: A federated transfer learning framework for wearable healthcare," IEEE Intelligent Systems, vol. 35, no. 4, pp. 83–93, 2020.

[26] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, et al., "Advances and open problems in federated learning," Foundations and trends® in machine learning, vol. 14, no. 1–2, pp. 1–210, 2021.

[27] F. A. KhoKhar, J. H. Shah, M. A. Khan, M. Sharif, U. Tariq, and S. Kadry, "A review on federated learning towards image processing," Computers and Electrical Engineering, vol. 99, p. 107818, 2022.

[28] M. Liu, S. Ho, M. Wang, L. Gao, Y. Jin, and H. Zhang, "Federated learning meets natural language processing: A survey," arXiv preprint arXiv:2107.12603, 2021.

[29] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang, and M. S. Hossain, "Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach," IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6348–6358, 2020.

[30] M. Noble, A. Bellet, and A. Dieuleveut, "Differentially private federated learning on heterogeneous data," in International Conference on Artificial Intelligence and Statistics, pp. 10110–10145, PMLR, 2022.

[31] L. Sun and L. Lyu, "Federated model distillation with noisefree differential privacy," arXiv preprint arXiv:2009.05537, 2020.

[32] P. Kulkarni, A. Kanhere, P. H. Yi, and V. S. Parekh, "Optimizing federated learning for medical image classification on distributed non-iid datasets with partial labels," arXiv preprint arXiv:2303.06180, 2023.
[33] Y. N. Tan, V. P. Tinh, P. D. Lam, N. H. Nam, and T. A. Khoa, "A transfer learning approach to breast cancer classification in a federated learning framework," IEEE Access, vol. 11, pp. 27462–27476, 2023.
[34] A. A. Noman, M. Rahaman, T. H. Pranto, and R. M. Rahman, "Blockchain for medical collaboration: A federated learning-based approach for multi-class

APPENDIX

This Appendix contains supplementary materials supporting our research, including detailed Tables that provide quantitative data and Figures that visually represent findings, facilitating deeper insights into our study's outcomes.

•Tables from 7 to 16 comprehensive classification performance Analysis of VGG11, Inception, DenseNet, and Res AUC, precision, recall, and F1-score evaluations for each class.

•Figures from 1 to 5 : represent ten heat maps which provi models like ResNet and VGG11 maintain robust performance after compression, others such as GoogleNet and DenseNet display more fluctuation across metrics. This variance highlights the nuanced effects of image compression on different models and underscores the need for careful consideration of model selection in clinical settings where image compression is utilized.

Table VII: VGG11 Classification Report Result before Compression

Data	Pre Trined Model	Class No.	Class Name	Classification Report For Each Class				
				AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	VGG11 Classification Report Result before Compression	1	No finding	0.99	0.91	0.85	0.85	2204
		2	Enlarged Cardiomediastinum	0.99	0.88	0.65	0.75	2337
		3	Cardiomegaly	0.97	0.91	0.84	0.87	3483
		4	Lung Opacity	0.98	0.89	0.91	0.91	11136
		5	Lung Lesion	0.96	0.90	0.69	0.78	1069
		6	Edema	0.98	0.89	0.90	0.90	6387
		7	Consolidation	0.98	0.89	0.80	0.83	4153
		8	Pneumonia	0.97	0.86	0.78	0.81	2439
		9	Atelectasis	0.97	0.84	0.82	0.85	6780
		10	Pneumothorax	0.98	0.88	0.77	0.82	2322
		11	Pleural Effusion	0.98	0.95	0.89	0.92	9782
		12	Pleural Other	0.99	0.91	0.67	0.78	653
		13	Fracture	0.98	0.92	0.62	0.74	966
		14	Support Devices	0.98	0.93	0.93	0.93	11750
			micro avg		0.91	0.86	0.88	65461
			macro avg		0.90	0.79	0.84	65461
			weighted avg		0.91	0.87	0.88	65461
			samples avg		0.86	0.84	0.83	65461
			AUC mean				0.98	

Table VIII: VGG11 Classification Report Result after Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	VGG11 Classification Report Result After Compression	1	No finding	0.99	0.92	0.95	0.93	2204
		2	Enlarged Cardiomediastinum	0.99	0.94	0.88	0.91	2337
		3	Cardiomegaly	0.99	0.91	0.95	0.93	3483
		4	Lung Opacity	0.99	0.96	0.96	0.96	11136
		5	Lung Lesion	0.96	0.90	0.92	0.91	1069
		6	Edema	0.99	0.93	0.96	0.94	6387
		7	Consolidation	0.99	0.93	0.93	0.93	4153
		8	Pneumonia	0.99	0.88	0.96	0.91	2439
		9	Atelectasis	0.99	0.93	0.94	0.94	6780
		10	Pneumothorax	0.99	0.93	0.92	0.92	2322
		11	Pleural Effusion	0.99	0.97	0.96	0.97	9782
		12	Pleural Other	0.99	0.82	0.96	0.89	653
		13	Fracture	0.99	0.94	0.88	0.91	966
		14	Support Devices	0.99	0.98	0.97	0.97	11750
			micro avg		0.94	0.95	0.95	65461
			macro avg		0.92	0.94	0.93	65461
			weighted avg		0.94	0.95	0.95	65461
			samples avg		0.92	0.93	0.92	65461
			AUC			0.99		

Table IX: Inception Classification Report Result Before Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	Inception Classification Report Result Before Compression	1	No finding	0.98	0.86	0.83	0.85	2204
		2	Enlarged Cardiomediastinum	0.95	0.71	0.74	0.73	2337
		3	Cardiomegaly	0.97	0.88	0.80	0.83	3483
		4	Lung Opacity	0.94	0.89	0.85	0.87	11136
		5	Lung Lesion	0.97	0.84	0.62	0.72	1069
		6	Edema	0.97	0.87	0.89	0.88	6387
		7	Consolidation	0.95	0.75	0.81	0.78	4153
		8	Pneumonia	0.96	0.69	0.80	0.74	2439
		9	Atelectasis	0.95	0.82	0.84	0.83	6780
		10	Pneumothorax	0.98	0.94	0.67	0.78	2322
		11	Pleural Effusion	0.97	0.90	0.93	0.91	9782
		12	Pleural Other	0.98	0.81	0.58	0.67	653
		13	Fracture	0.97	0.79	0.71	0.75	966
		14	Support Devices	0.97	0.94	0.90	0.92	11750
			micro avg		0.86	0.85	0.85	65461
			macro avg		0.83	0.78	0.78	65461
			weighted avg		0.86	0.87	0.85	65461
			samples avg		0.83	0.82	0.81	65461
			AUC mean				0.97	

Table X: Inception Classification Report Result After Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	Inception Classification Report Result After Compression	1	No finding	0.98	0.89	0.76	0.82	2204
		2	Enlarged Cardiomediastinum	0.98	0.80	0.67	0.74	2337
		3	Cardiomegaly	0.96	0.83	0.87	0.85	3483
		4	Lung Opacity	0.98	0.82	0.95	0.88	11136
		5	Lung Lesion	0.97	0.89	0.57	0.70	1069
		6	Edema	0.97	0.90	0.86	0.88	6387
		7	Consolidation	0.95	0.81	0.75	0.78	4153
		8	Pneumonia	0.96	0.80	0.71	0.75	2439
		9	Atelectasis	0.95	0.84	0.83	0.84	6780
		10	Pneumothorax	0.98	0.82	0.82	0.82	2322
		11	Pleural Effusion	0.97	0.90	0.94	0.94	9782
		12	Pleural Other	0.98	0.79	0.63	0.63	653
		13	Fracture	0.97	0.79	0.68	0.68	966
		14	Support Devices	0.97	0.96	0.85	0.85	11750
			micro avg		0.87	0.85	0.86	65461
			macro avg		0.85	0.78	0.81	65461
			weighted avg		0.87	0.85	0.86	65461
			samples avg		0.83	0.82	0.81	65461
			AUC mean				0.97	

Table XI: Dense Net Classification Report Result Before Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	Dense Net Classification Report Result Before Compression	1	No finding	0.98	0.81	0.76	0.80	2204
		2	Enlarged Cardiomediastinum	0.94	0.81	0.67	0.67	2337
		3	Cardiomegaly	0.97	0.90	0.87	0.78	3483
		4	Lung Opacity	0.93	0.84	0.95	0.86	11136
		5	Lung Lesion	0.97	0.74	0.57	0.71	1069
		6	Edema	0.96	0.80	0.86	0.85	6387
		7	Consolidation	0.94	0.81	0.75	0.73	4153
		8	Pneumonia	0.95	0.68	0.71	0.72	2439
		9	Atelectasis	0.93	0.80	0.83	0.80	6780
		10	Pneumothorax	0.97	0.81	0.82	0.80	2322
		11	Pleural Effusion	0.96	0.92	0.94	0.88	9782
		12	Pleural Other	0.97	0.79	0.63	0.59	653
		13	Fracture	0.97	0.87	0.68	0.86	966
		14	Support Devices	0.96	0.94	0.85	0.90	11750
			micro avg		0.85	0.81	0.83	65461
			macro avg		0.82	0.73	0.77	65461
			weighted avg		0.85	0.81	0.83	65461
			samples avg		0.81	0.78	0.78	65461
			AUC mean				0.96	

Table XII: Dense Net Classification Report Result Before Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	Dense Net Classification Report Result After Compression	1	No finding	0.98	0.85	0.73	0.79	2204
		2	Enlarged Cardiomediastinum	0.94	0.74	0.69	0.71	2337
		3	Cardiomegaly	0.97	0.75	0.86	0.81	3483
		4	Lung Opacity	0.94	0.87	0.87	0.87	11136
		5	Lung Lesion	0.97	0.70	0.71	0.71	1069
		6	Edema	0.96	0.79	0.91	0.84	6387
		7	Consolidation	0.94	0.76	0.74	0.75	4153
		8	Pneumonia	0.95	0.80	0.67	0.73	2439
		9	Atelectasis	0.94	0.91	0.60	0.73	6780
		10	Pneumothorax	0.97	0.85	0.72	0.78	2322
		11	Pleural Effusion	0.96	0.90	0.89	0.90	9782
		12	Pleural Other	0.97	0.77	0.59	0.67	653
		13	Fracture	0.97	0.74	0.68	0.71	966
		14	Support Devices	0.96	0.90	0.92	0.91	11750
			micro avg		0.85	0.82	0.83	65461
			macro avg		0.81	0.76	0.78	65461
			weighted avg		0.85	0.82	0.83	65461
			samples avg		0.81	0.80	0.79	65461
			AUC mean				0.96	

Table XIII: ResNet18 Classification Report Result before Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	**ResNet18 Classification Report Result before Compression**	1	No finding	0.99	0.84	0.91	0.87	2204
		2	Enlarged Cardiomediastinum	0.98	0.87	0.80	0.83	2337
		3	Cardiomegaly	0.99	0.97	0.78	0.86	3483
		4	Lung Opacity	0.97	0.93	0.88	0.91	11136
		5	Lung Lesion	0.99	0.95	0.72	0.82	1069
		6	Edema	0.98	0.94	0.87	0.91	6387
		7	Consolidation	0.97	0.85	0.84	0.84	4153
		8	Pneumonia	0.98	0.85	0.86	0.85	2439
		9	Atelectasis	0.97	0.90	0.86	0.88	6780
		10	Pneumothorax	0.99	0.94	0.84	0.89	2322
		11	Pleural Effusion	0.98	0.97	0.87	0.92	9782
		12	Pleural Other	0.99	0.87	0.81	0.84	653
		13	Fracture	0.99	0.74	0.89	0.81	966
		14	Support Devices	0.98	0.94	0.94	0.94	11750
			micro avg		0.87	0.90	0.90	65461
			macro avg		0.85	0.87	0.87	65461
			weighted avg		0.87	0.90	0.90	65461
			samples avg		0.85	0.85	0.85	65461
			AUC mean				0.98	

Table XIV: ResNet18 Classification Report Result AFTER Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
Chexpert:5 Clients And 10 Rounds	**ResNet18 Classification Report Result AFTER Compression**	1	No finding	0.99	0.73	0.96	0.96	2204
		2	Enlarged Cardiomediastinum	0.98	0.89	0.77	0.77	2337
		3	Cardiomegaly	0.99	0.96	0.78	0.78	3483
		4	Lung Opacity	0.97	0.97	0.79	0.79	11136
		5	Lung Lesion	0.99	0.88	0.75	0.75	1069
		6	Edema	0.98	0.97	0.75	0.75	6387
		7	Consolidation	0.97	0.87	0.83	0.83	4153
		8	Pneumonia	0.98	0.90	0.76	0.76	2439
		9	Atelectasis	0.97	0.90	0.85	0.85	6780
		10	Pneumothorax	0.99	0.84	0.90	0.90	2322
		11	Pleural Effusion	0.98	0.95	0.92	0.92	9782
		12	Pleural Other	0.99	0.93	0.72	0.72	653
		13	Fracture	0.99	0.90	0.78	0.78	966
		14	Support Devices	0.98	0.97	0.88	0.88	11750
			micro avg		0.93	0.84	0.88	65461
			macro avg		0.90	0.82	0.85	65461
			weighted avg		0.93	0.84	0.88	65461
			samples avg		0.89	0.82	0.84	65461
			AUC mean				0.98	

Table XV: GoogLeNet Classification Report Result before Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	**GoogLeNet Classification Report Result before Compression**	1	No finding	0.97	0.90	0.60	0.72	2204
		2	Enlarged Cardiomediastinum	0.89	0.68	0.49	0.57	2337
		3	Cardiomegaly	0.95	0.74	0.78	0.76	3483
		4	Lung Opacity	0.89	0.82	0.80	0.81	11136
		5	Lung Lesion	0.93	0.67	0.52	0.59	1069
		6	Edema	0.94	0.91	0.62	0.74	6387
		7	Consolidation	0.89	0.45	0.86	0.59	4153
		8	Pneumonia	0.92	0.68	0.57	0.62	2439
		9	Atelectasis	0.89	0.67	0.80	0.73	6780
		10	Pneumothorax	0.95	0.72	0.75	0.73	2322
		11	Pleural Effusion	0.96	0.84	0.93	0.88	9782
		12	Pleural Other	0.95	0.61	0.55	0.58	653
		13	Fracture	0.94	0.56	0.65	0.60	966
		14	Support Devices	0.95	0.93	0.86	0.89	11750
			micro avg		0.76	0.78	0.77	65461
			macro avg		0.73	0.70	0.70	65461
			weighted avg		0.79	0.78	0.77	65461
			samples avg		0.74	0.750.72		65461
			AUC mean			0.93		

Table XVI: GoogLeNet Classification Report Result after Compression

Data	Pre Trined Model	Class No.	Class Name	AUC	Precision	Recall	F1-Score	Support
CheXpert: 5 Clients And 10 Rounds	**GoogLeNet Classification Report Result after Compression**	1	No finding	0.97	0.91	0.60	0.72	2204
		2	Enlarged Cardiomediastinum	0.89	0.55	0.63	0.59	2337
		3	Cardiomegaly	0.95	0.72	0.79	0.76	3483
		4	Lung Opacity	0.89	0.78	0.86	0.82	11136
		5	Lung Lesion	0.94	0.61	0.59	0.60	1069
		6	Edema	0.95	0.89	0.70	0.78	6387
		7	Consolidation	0.89	0.73	0.54	0.62	4153
		8	Pneumonia	0.92	0.73	0.53	0.61	2439
		9	Atelectasis	0.89	0.79	0.64	0.71	6780
		10	Pneumothorax	0.96	0.70	0.79	0.74	2322
		11	Pleural Effusion	0.95	0.82	0.93	0.87	9782
		12	Pleural Other	0.94	0.46	0.57	0.51	653
		13	Fracture	0.93	0.72	0.45	0.55	966
		14	Support Devices	0.95	0.85	0.94	0.89	11750
			micro avg		0.79	0.78	0.78	65461
			macro avg		0.73	0.68	0.70	65461
			weighted avg		0.79	0.78	0.78	65461
			samples avg		0.75	0.75	0.73	65461
			AUC mean				0.93	

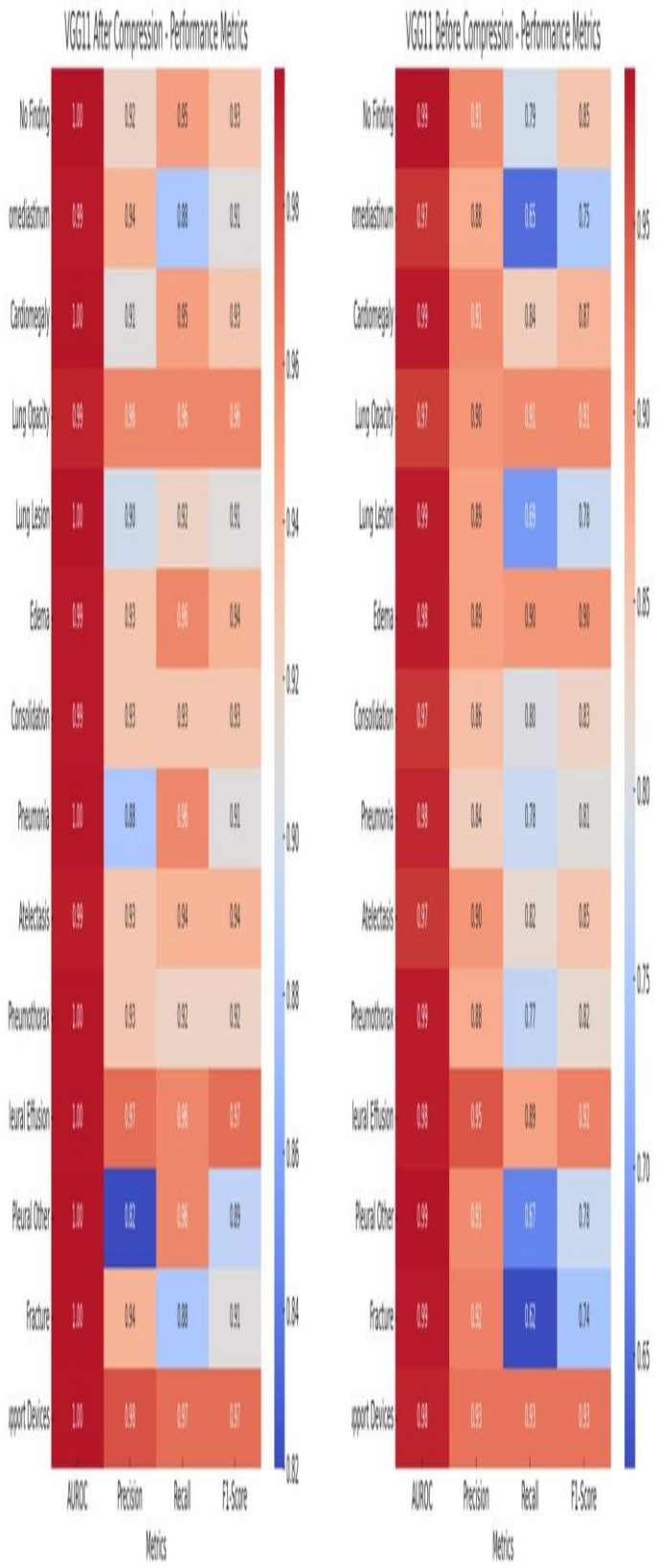

Figure 1:VGG Heat-maps Before and After apply the JPEG algorithm compression

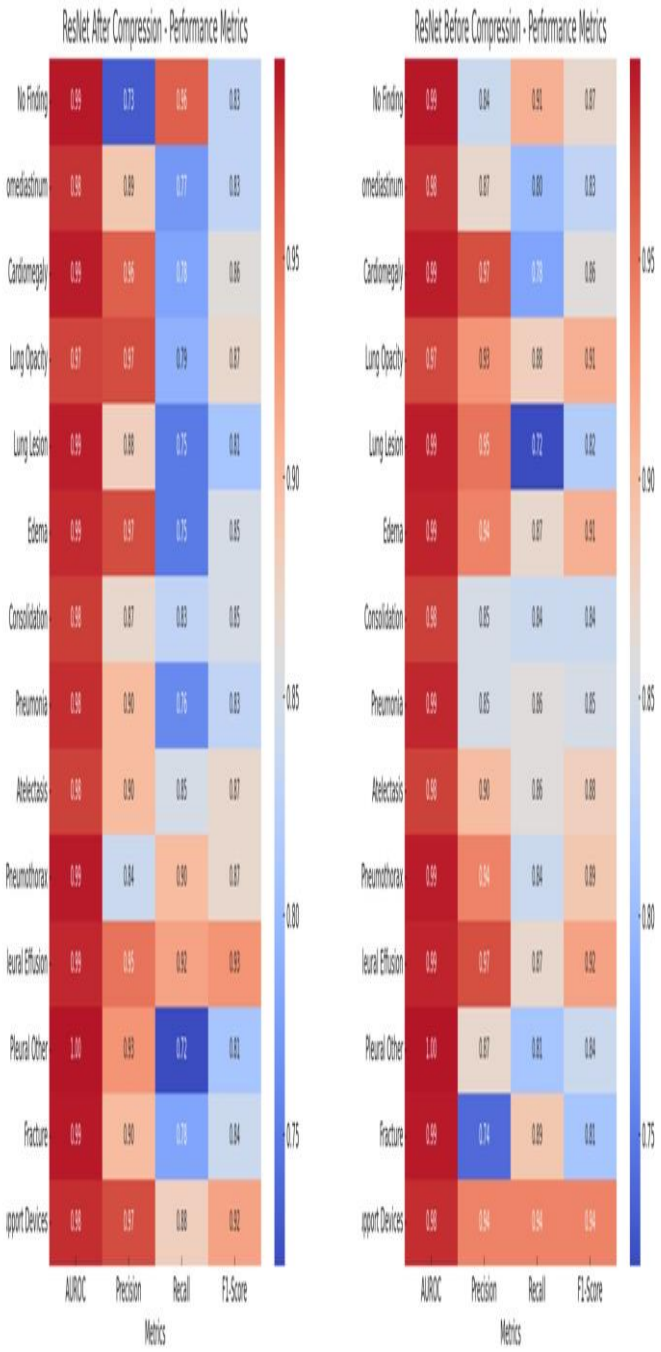

Figure 2: ResNet Heat-maps Before and After apply the JPEG algorithm compression

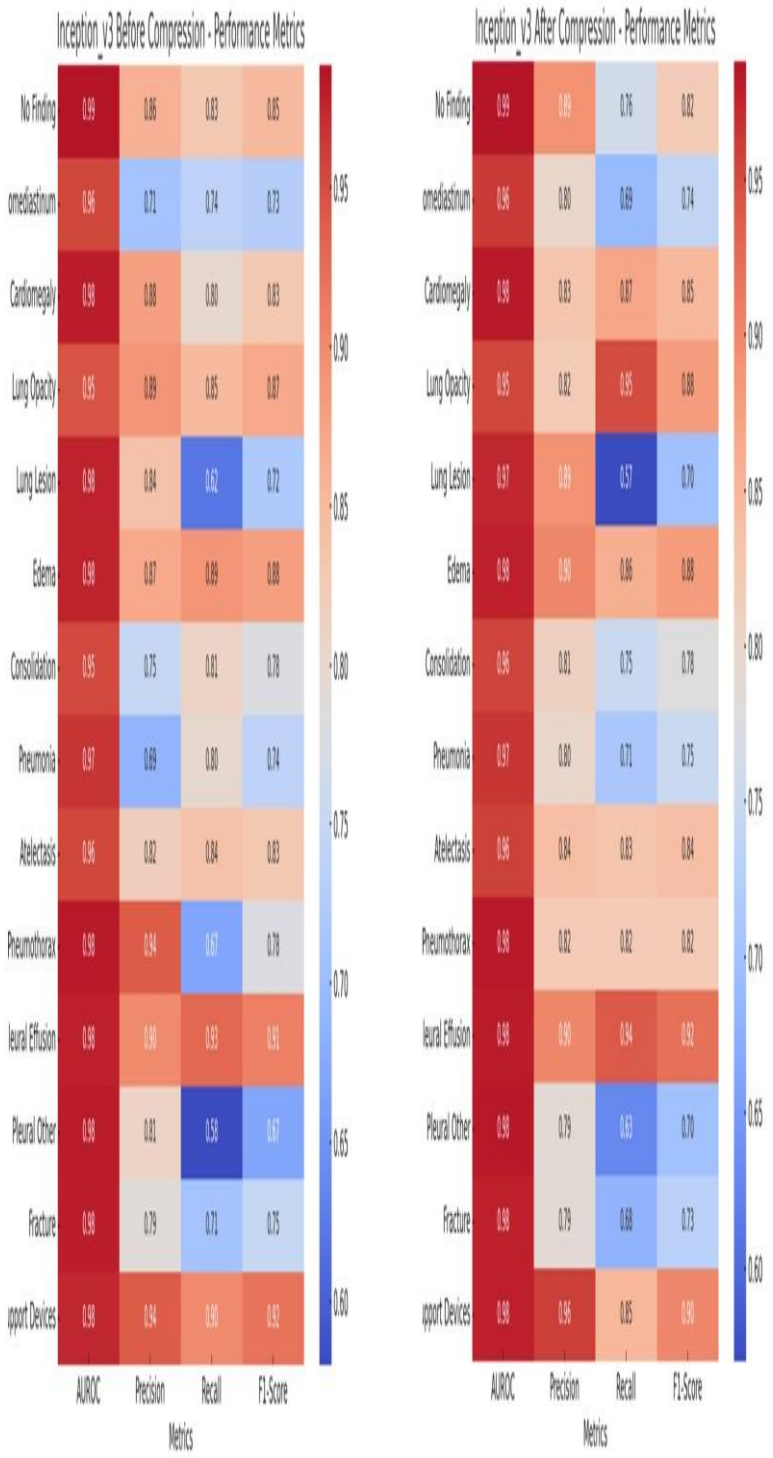

Figure 3 : Inception-v3 Heat-maps Before and After apply the JPEG algorithm compression

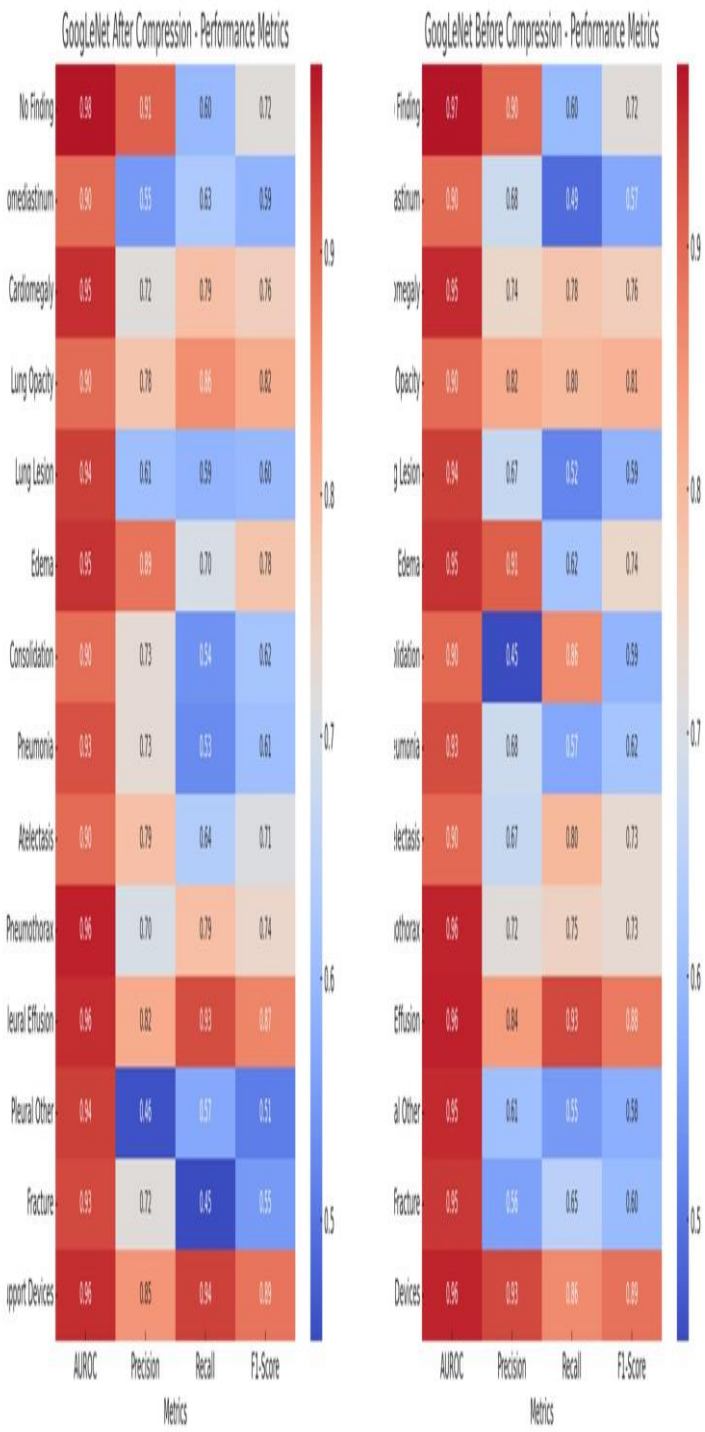

Figure 4: GoogLeNet Heat-maps Before and After apply the JPEG algorithm compression

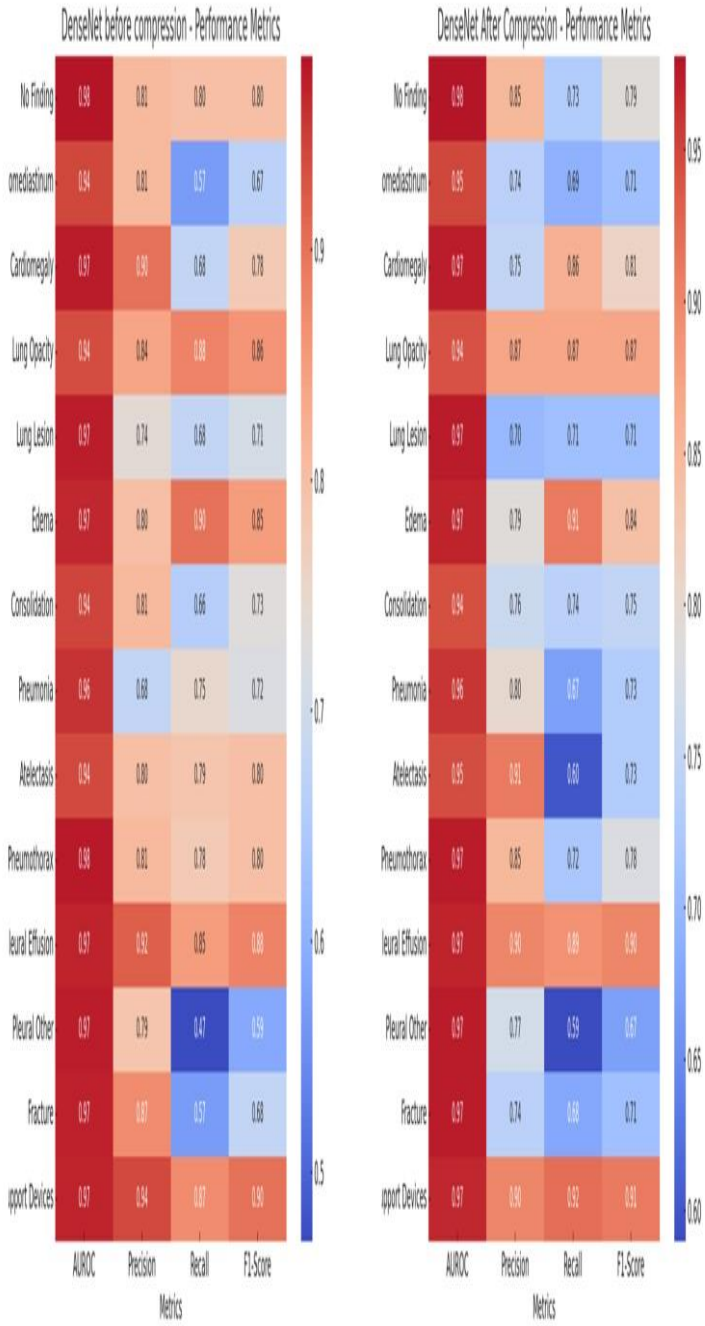

Figure 5: DenceNet Heat-maps Before and After apply the JPEG algorithm compression

تحسين التعلم الاتحادي لتصنيف الصور الطبي: دراسة مقارنة للنماذج المدربة مسبقا على صور الأشعة السينية المضغوطة

سوسن طالع الوادعي [1,4] ، أماني طارق جمال [1,5] ، سمر الخريجي [2] ، لمياء الرفاعي[3]

[123]قسم علوم الحاسبات، كلية الحاسبات وتقنية المعلومات، جامعة الملك عبد العزيز، جدة ، المملكة العربية السعودية

[4]قسم الهندسة الكهربائية كلية الهندسة بشبرا ، جامعة بنها، مصر

Salwadaie@stu.kau.edu.sa, Atjamal@kau.edu.sa, Salkhuraiji@kau.edu.sa, Lamia.alrefaai@feng.bu.edu.eg

المستخلص – حدث ظهور التعلم الآلي، وخاصة التعلم العميق، ثورة في العديد من المجالات، بما في ذلك التشخيص الطبي. تستغل هذه الدراسة إمكانات التعلم الاتحادي أو التعلم التعاوني لمعالجة مخاوف الخصوصية وقيود الوصول إلى البيانات المتأصلة في التصوير الطبي. نستكشف أداء خمسة نماذج للشبكات العصبية المدربة مسبقًا DENSENET121وRESNET18 و VGGNET11و GOOGLENET و INCEPTION_V3، على مجموعة بياناتCheXpert ضمن بيئة محاكاة لتعلم الاتحادي او التعلم التعاوني. ويؤكد البحث على تحسين مدة التدريب وتطبيق ضغط الصور JPEG للحصول على الكفاءة والسرعة أثناء جولات الاتصال. تشتمل منهجيتنا على تحليل مقارن لأداء النماذج قبل وبعد الضغط وتقييم المنطقة الواقعة تحت المنحنى وتقدير وقت التدريب . تشير النتائج إلى أن ضغط الصور يمكن أن يحافظ على أداء النموذج أو يحسنه بينما يؤثر أيضًا على وقت التدريب، مما يؤكد المفاضلة بين دقة النموذج والكفاءة الحسابية.

الكلمات المفتاحيةـ – التعلم الاتحادي أو التعاوني، خوارزمية الضغط JPEG، قاعدة البيانات CheXpert ، بيئة المحاكاة.