Vol.21 Iss.1 pp.42-51 International Journal of Scientific Research & Engineering Technology (IJSET)

Balancing Efficiency and Accuracy: An Empirical Study of Image Compression Effects on Machine Learning Classification

Howayda Abedallah Elmarzaki[1], Elbhlul Emhimmed Fgee [2], Rabei Alhadad[3]

1 Faculty of Information Technology, University of Benghazi, Benghazi, Libya

howayda.elmajpri@uob.edu.ly

2 Faculty of Information Technology, University of Gharyan, Tripoli, Libya

bahlul.fgee@academy.edu.ly

3 Faculty of science and Information Technology, University of La Ttrobe, Melbourne, Australia

R.Alhadad@latrobe.edu.au

***Abstract*—In machine learning, especially in medical image analysis where combining prediction accuracy with compute economy is of great significance, choosing an appropriate model is usually difficult. Growing integration of artificial intelligence (AI) into healthcare applications like telemedicine, digital pathology, and clinical diagnostics underscores the need of improving both model performance and data processing. In this view, image reduction becomes fairly important, especially in settings with limited resources, allowing faster transmission and reduced storage needs without obvious loss of diagnostic integrity. This work explores the implications on machine learning model accuracy applied in medical imaging of many photo compression settings. With a dataset of about 7,000 photos linked to Alzheimer's disease, four different rounds of experimental assessments were carried out. Five different machine learning methods were evaluated under different picture resolutions and compression ratios. Among them, logistic regression usually outperformed the others in preserving high accuracy even in significant picture deterioration. The results give pertinent direction for choosing models that show strong resistance to compression artifacts and defining suitable compression limits for AI-based diagnostic systems and telemedicine platforms.**

***Keywords*—Intelligent learning systems, neurodegenerative diseases, clinical image, image compression levels.**

I. INTRODUCTION

In the discipline of machine learning especially in medical uses where early and precise diagnosis may greatly affect patient outcomes, image categorization and predictive modeling are a major focus of research. Comprehensive anatomical and functional knowledge about interior body structures is mostly obtained via magnetic resonance imaging (MRI), computed tomography (CT), X-rays, and positron emission tomography (PET). Crucially relevant for the diagnosis and monitoring of neurological diseases like multiple sclerosis, brain tumors, and Alzheimer's disease, figure 1 shows different forms of brain imaging techniques [1]. Accurate anatomical location—kidneys, liver, uterus, heart—determining the stage of advancement, and stressing the area of interest (ROI) inside the organ allow to identify disease presence by means of these pictures [2]. Using machine learning (ML) approaches, rich imaging data is being progressively applied in automated diagnosis systems. Large-scale datasets from public health databases, open-source picture banks, clinical archives, or other sources might all find place in ML models. To guarantee consistency and increase model generalizability, these datasets encompassing multiple picture formats and resolutions necessitate preprocessing including normalizing, scaling, and augmentation. Sometimes restricted processing resources or storage capacity lead remote healthcare systems or real-time systems to adopt picture compression in particular. Although file size is being reduced, both lossy and lossless compression methods maintain diagnostically significant properties. Most importantly for building strong machine learning models, this not only provides quicker data transport and reduced storage costs but also lets dataset quantities increase [3]. Machine learning methods are meant to uncover patterns and infer correlations

ISSN: 1737-9237

among difficult data on their own without clear rule-based programming. Sometimes the type of the learning process directs the categorization of ML techniques: supervised learning (where models learn from labeled data), unsupervised learning (which needs the discovery of hidden patterns in unlabeled data), and forecasting models (used for time-dependent predictions). This work focuses on supervised learning techniques based on their exhibited performance on challenging classification problems. Five well acknowledged techniques are used in performance evaluation: Support Vector Machine (SVM), known for its high accuracy in high-dimensional spaces [5]; Logistic Regression (LR), a probabilistic model effective for binary classification problems [6]; Decision Tree (DT), which builds interpretable tree-structured decision rules; K-Nearest Neighbor (KNN), a distance-based method relying on proximity to labeled samples [7]; and Artificial Neural Network (ANN), a biologically inspired model capable of learning complex nonlinear patterns through layered architectures [8]. These models are examined extensively under several picture compression settings to evaluate their resistance in maintaining classification performance and their susceptibility to data deterioration. This thorough assessment not only increases knowledge of model robustness under various data quality levels but also offers useful information for the implementation of machine learning solutions in telemedicine, mobile health, and other AI-driven diagnostic platforms where compressed medical imaging is periodically a need.

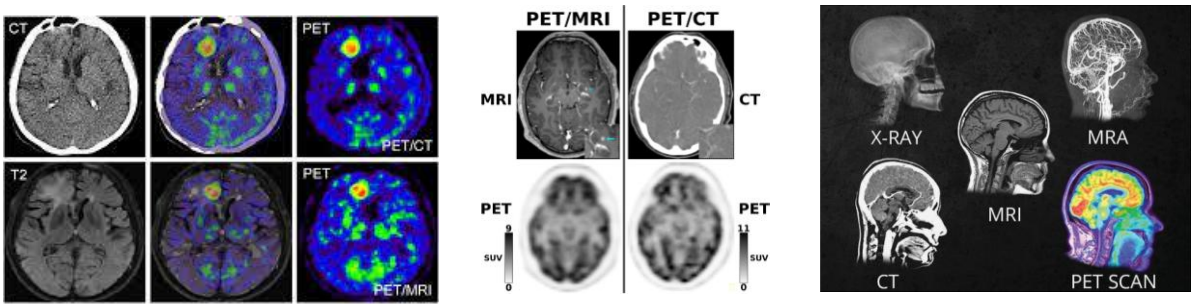

Fig.1.Different Modalities of Brain Imaging

Additionally addressing medical imaging connected to Alzheimer's disease (AD), this study mostly concentrates on brain areas showing pathogenic alterations at many phases of the illness. The shrinking in the hippocampal volume—a region crucial for memory processing—is one of the most obvious neuroanatomical markers of AD. This shrinking is often used by doctors as a disease diagnosis tool; medical imaging indisputably shows [1]. Common clinical symptoms include depressed behaviors, memory loss, and social disengagement aside from structural alterations in the brain. Early identification is vital as studies reveal that the neurodegenerative mechanisms behind AD start much before the onset of obvious cognitive decline, including major memory loss [2]. Figure 2, among the first areas under influence, displays the hippocampus region and provides a major clue used in the Alzheimer's disease diagnosis.

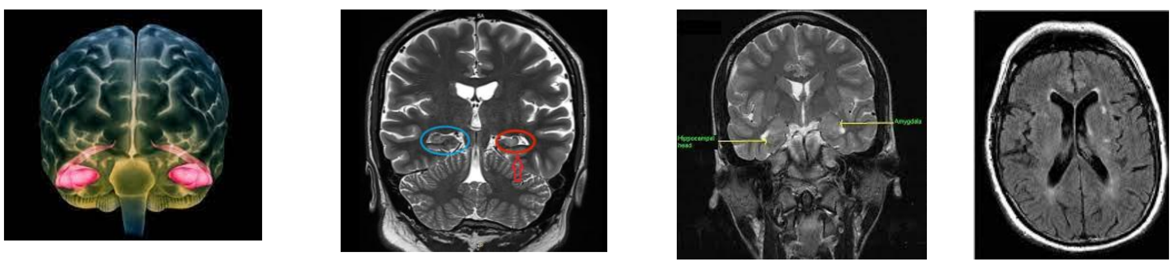

Fig.2. MRI Image (Hippocampus region) of Human Brain

There are three different data sets used in this research, the first is original dataset about the brain of Alzheimer disease with four stages are Non-Demented, Mild Demented, Very Mild Demented, and Moderate Demented based

ISSN: 1737-9237

on Kaggle platform named Alzheimer's Disease(4 classes of images). Are JPEG sized in (176*208) of 5789 different image of brain. The regain of interest ROI is the Hippocampus size. The second and the third datasets are compressed images. As it clearly in experiment section. The paper is structured as follows: Section II presents the problem statement, highlighting key challenges and objectives. Section III introduces the dataset used in this experiment. Section IV reviews literature review to summarizing prior studies and methodologies. Section V demonstrates the study's contribution. The details of the proposed methodology are illustrated in section VI, while section VII discusses experimental results and the comparisons show in VIII. In IX, challenges will be discussed, and in X, the conclusion and future work section is at last presented.

II. PROBLEM STATEMENT

Clinical diagnosis and decision-making depend much on medical imaging; in this case, the quality of compressed pictures greatly affects diagnostic usefulness. Compression techniques have become ever more crucial for improving data storage efficiency and computer processing as medical imaging technology develop dramatically. These compression techniques could, however, compromise image quality, therefore affecting the prediction ability of machine learning systems in medically sensitive uses. The exact compression level that preserves best algorithmic accuracy is yet unknown, which makes it difficult to balance computational economy with diagnostic accuracy.

III. DATASETS

This work makes advantage of three different datasets. Comprising actual brain scans linked with Alzheimer's disease, the main dataset consists in four clinical stages: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. Source from the Kaggle platform, this dataset titled Alzheimer's Disease (4 classes of photos) has 5,789 JPEG images with a 176x 208 pixel resolution. The hippocampal region of interest (ROI) in these pictures is the one most of interest as its size is a major determinant of illness development. Apart from the original dataset, two additional datasets were generated by varying degrees of picture compression in order to assess how lowered image quality affected model performance. The experimental part describes specifics of the compression methods and settings.

IV. LITERATURE REVIEW

Jo et al. [1] looked at how picture compression affected deep learning models' performance in mammography classification—malignant or non-malignant? With an AUROC of 0.87 and an AUPRC of 0.75, the study indicated that convolutional neural networks (CNNs) attained ideal performance at a compression level of 5 KB. Performance fell significantly, nevertheless, with increasing compression ratios. Moreover, saliency map analysis shown improved agreement with radiologist annotations in pictures with low compression. Although the research also looked at image quality using Peak Signal-to-- Noise Ratio (PSNR), it did not include comprehensive performance information for any one model.

Poyser et al. [2] investigated how performance of deep neural networks varies with different degrees of lossy picture compression. Their results revealed a notable decline in model accuracy when JPEG quality (quantization) dropped below a certain level. The study neglected conventional machine learning algorithms or medical imaging data even though it gave insightful analysis of the resilience of deep learning models under lossy compression and retraining methods.

Using deep learning architectures, Yijiang et al. [3] evaluated how JPEG and JPEG2000 compression formats affected breast cancer picture segmentation. Unlike research centered on classification, this effort focussed on segmentation performance and confined to a single trial design without cross-model comparisons. Performance criteria including the F1 score and AUROC were used in evaluation.

Emphasizing its applicability for clinical interpretation, Urbaniak and Ilona Anna [4] investigated the sensitivity of deep learning models to the loss of high-frequency visual material arising from compression. The

ISSN: 1737-9237

paper focused on the use of JPEG and JPEG2000 compression in deep learning but did not really address classification problems or make use of CT, MRI, or PET scan datasets.

Kurmukov et al. [5] investigated, using U-Net models, how picture compression affected segmentation performance over several domains. The study found that performance was constant even with compressive ratios as high as 20:1. It stayed focused on segmentation rather than classification, while including JPEG2000 compression and imaging modalities like CT and MRI.

V. CONTRIBUTION

This article presents perceptual analysis of how picture compression affects machine learning model performance in medical imaging applications. Underlining the natural trade-off between compression ratio and classification accuracy helps notably in settings with limited processing capability to provide advice on suitable compression limits to ensure diagnostic reliability. Furthermore, the results reveal the obvious opposition to compression-induced degradation different approaches—more precisely, logistic regression and artificial neural networks. Effective data management and consistent prediction accuracy are quite critical in telemedicine and AI-assisted diagnostic systems; thus, the results offer a good direction for the selection of suitable models.

VI. METHODOLOGY

The research method consists of multiple Stages as it illustrates in next points and shows in fig 3.

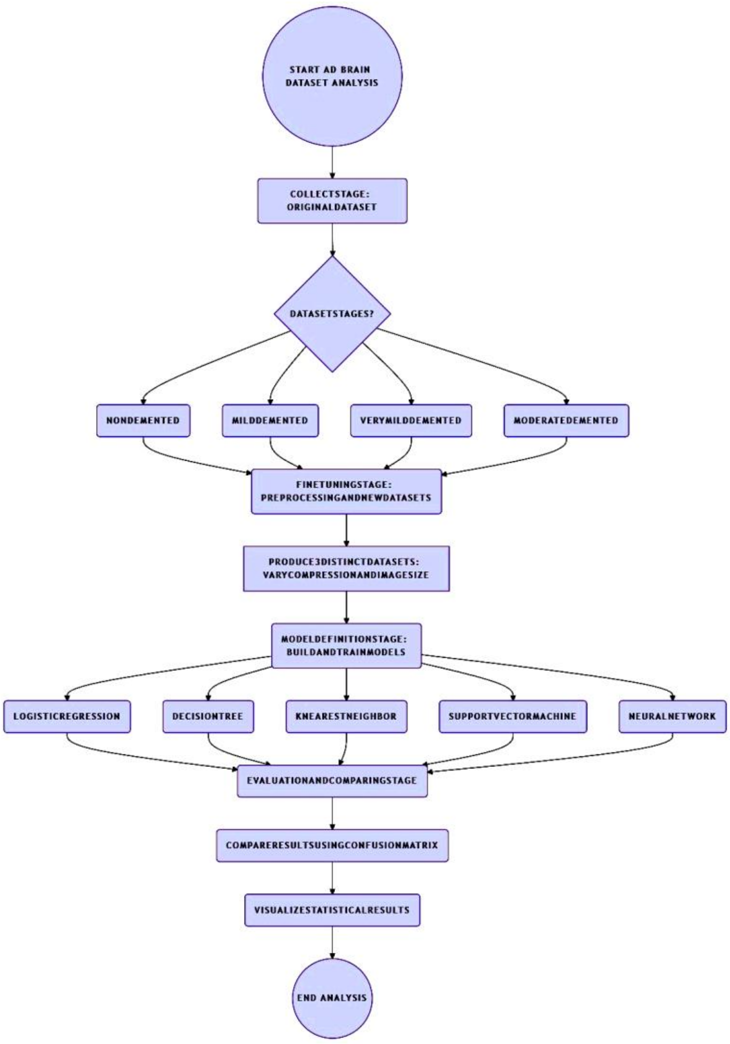

Fig.3.The Methodology With Stages

ISSN: 1737-9237

Collect Stage: includes original data on the four stages of AD brain development: non-demented, mildly demented, very mildly demented, and moderately demented.

Stage of Fine-Tuning: Pre-processing and the creation of new data sets are taken into account at this stage. These are predicated on altering the original datasets in terms of image size and compression quality in order to create three different datasets for the tests. Table I shows the three types of datasets have variance in size and image compression ratio. Table I and Fig 4 shows the datasets used in this study

TABLE I. THE USED DATASETS

Experiment No	Dataset Name	Image No	Image Type	Image Size	Compression Ratio(CR)/Quality
1	Alzheimer-Original-176-208	5789	JPEG	176*208	Acceptable quality for model's Training
2	Alzheimer-320-240	5789	JPEG	320*240	50% Quality
3	Alzheimer-90-10075	5789	JPEG	100*75	0% Quality (Degraded)

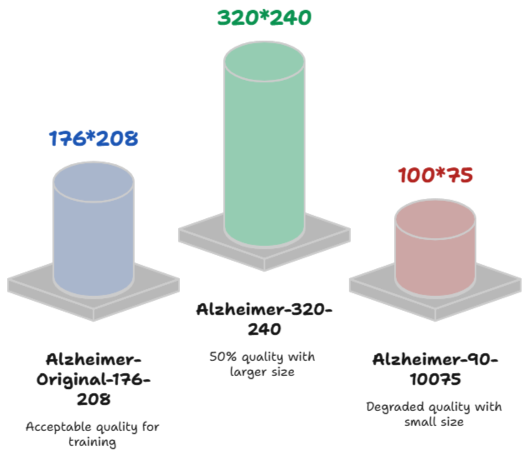

Fig 4. Datasets Used in Three Experiments

A representative sample from the original dataset is shown in Fig. 5 along with corresponding images that have been compressed to different degrees. While the second sample shows the extreme degradation brought on by maximum compression, which lowers the quality to 0%, the first sample shows the image quality following moderate compression at 50%. This demonstrates how severely compression affects visual clarity and image fidelity.

a b a c

Fig 5. (a) Sample of Original Dataset, (b) Image After Compression 50% and (c) Degraded Image /Zero Quality

ISSN: 1737-9237

VII. EXPERIMENTS

Three experiments based on different datasets make up this study, as indicated in Table 2. Every experiment follows the procedures outlined below.

1. Import the readable folder containing the dataset.
2. Image embedding, which involves converting each image to vectors with features.
3. To divide the dataset into training and testing (the 10-fold cross-validation method was used in this work).
4. The Learning of the Model
5. To use the Confusion Matrix to assess each model's performance in each experiment based on F-Score, Accuracy, Recall, and Precession.

VIII. RESULTS AND COMPARISON

A. **Results**

The Artificial Neural Network performed the best, with an accuracy of 0.870, followed by KNN at 0.836, according to the results of the first experiment, which are shown in Table II and Fig5. While Logistic Regression and Decision Tree performed poorly, with Decision Tree being the least effective (0.593), SVM displayed moderate results (0.754). These results highlight the superiority of ANN for classification tasks in this particular context.

Model	Accuracy	F-score	Precision	Recall
Logistic- Regression	0.716	0.716	0.716	0.716
Decision -Tree	0.593	0.592	0.592	0.593
K-Nearest Neighbour	0.836	0.834	0.836	0.836
Support Vector Machine	0.754	0.753	0.759	0.754
Artificial Neural Network	0.870	0.870	0.870	0.870

TABLE II. THE RESULTS OF THE FIRST EXPERIMENT OF FIVE MODELS

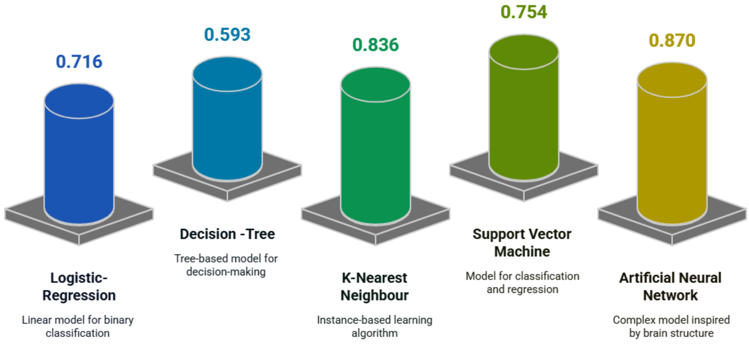

Fig 6. Results of First experiment/ Performance of Classification Algorithms

In the second experiment as it shown in table III and Fig 7, the results show that all models performed poorly, with Logistic Regression obtaining the highest accuracy (0.463) but the lowest precision (0.336) and F-score (0.361). Although they performed poorly overall, KNN and SVM demonstrated marginally higher precision (0.251 and 0.388, respectively). On every metric, ANN and Decision Tree performed the worst. These results show how difficult it was to obtain accurate classification in both experiments.

ISSN: 1737-9237

Model	Accuracy	F-score	Precision	Recall
Logistic- Regression	0.463	0.361	0.336	0.463
Decision -Tree	0.101	0.117	0.140	0.101
K-Nearest Neighbour	0.247	0.249	0.251	0.247
Support Vector Machine	0.259	0.287	0.388	0.259
Artificial Neural Network	0.208	0.201	0.194	0.208

TABLE III. RESULTS OF THE SECOND EXPERIMENT OF FIVE MODELS

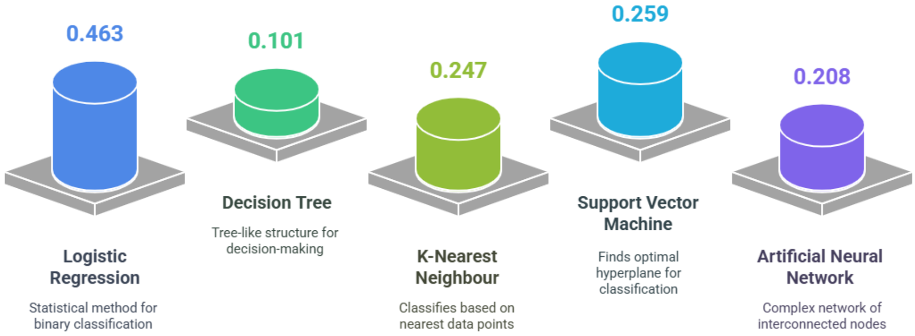

Fig 7. Results of Second experiment/ Performance of Classification Algorithms

The third experiment's results indicate that all models performed below par. The best accuracy (0.455) was demonstrated by logistic regression, which also had a moderate F-score (0.383) and precision (0.363). SVM (accuracy: 0.223) and KNN (accuracy: 0.251) performed worse overall but had marginally higher precision. The weakest models were ANN (accuracy: 0.226) and Decision Tree (accuracy: 0.121).as table iv illustrates and Fig8.

Model	Accuracy	F-score	Precision	Recall
Logistic- Regression	0.455	0.383	0.363	0.455
Decision -Tree	0.121	0.138	0.162	0.121
K-Nearest Neighbour	0.251	0.270	0.293	0.251
Support Vector Machine	0.223	0.257	0.387	0.223
Artificial Neural Network	0.226	0.217	0.209	0.226

TABLE IV. THE RESULTS OF THE THIRD EXPERIMENT OF FIVE MODELS

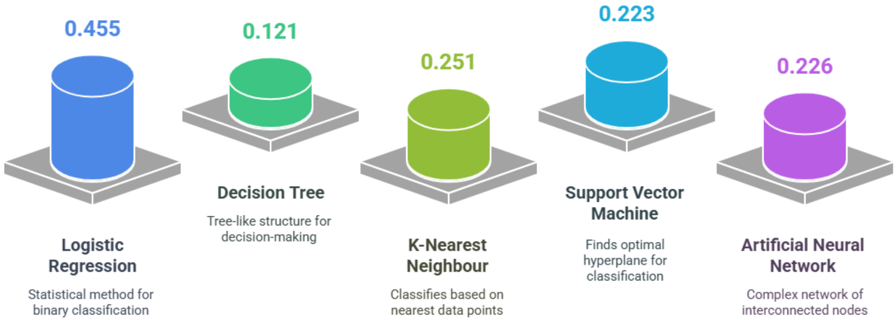

Fig 8. Results of Third experiment/ Performance of Classification Algorithms

ISSN: 1737-9237

B. Compression between model's performance

Significant differences in the robustness and sensitivity to image compression are revealed by comparing the five machine learning models—Logistic Regression, Decision Tree, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Artificial Neural Network (ANN)—under three different experimental conditions. ANN and KNN demonstrated their strong ability to learn from high-resolution medical images by achieving the highest accuracies (87.0% and 83.6%, respectively) in the first experiment using the original, uncompressed dataset. Decision Tree performed the worst (59.3%), followed by SVM and Logistic Regression. However, in the second and third experiments, the accuracy of all models significantly decreased as image compression increased. Although ANN performed better on the original dataset, it drastically declined to 20.8% and 22.6%, respectively, while logistic regression demonstrated the highest resilience, maintaining moderate performance under both compression levels (46.3% and 45.5%). Under high compression, both KNN and SVM showed comparable sensitivity, dropping to about 25% accuracy. With accuracy below 13% in both tests, the Decision Tree model fared the worst under compression. These results imply that while sophisticated models like artificial neural networks (ANN) perform well with high-quality inputs, simpler models like logistic regression might be more resilient to quality degradation, which would make them more appropriate for low-resource compressed medical imaging applications as it illustrated in Fig9.

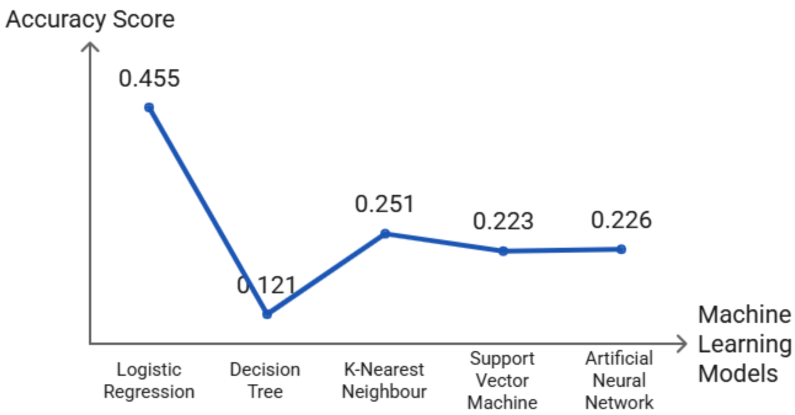

Fig 9. Three experiments comparing machine learning algorithms

IX. CONCLUSION

The experimental findings unequivocally show that different levels of image compression have a significant impact on the classification performance of machine learning models used in medical imaging, especially when diagnosing Alzheimer's disease. Although the degree of this impact varies among algorithms, models' capacity to correctly classify disease stages declines as image quality deteriorates due to compression. Notably, models with a relatively high degree of predictive accuracy, like Logistic Regression and Artificial Neural Networks, demonstrated increased robustness under moderate compression conditions. Decision Trees and K-Nearest Neighbor algorithms, on the other hand, showed a noticeable drop in performance, indicating a reduced ability to withstand the loss of image detail that is usually connected with compressed formats. The importance of optimizing compression ratios to balance storage and transmission efficiency with diagnostic precision is highlighted by these findings, which highlight the crucial trade-off between image compression and model efficacy. This balance is particularly important in settings with limited resources, like telemedicine platforms and remote healthcare settings, where compressed image data is required due to storage and bandwidth limitations. Additionally, the findings offer useful advice for medical professionals and those creating AI-based diagnostic tools, making it easier to choose

ISSN: 1737-9237

models that are accurate and resistant to compression artifacts, thus promoting dependable and expandable medical imaging workflows.

X. FUTURE WORK

Beyond compression, future studies will look into a wider range of image degradations, such as the effects of noise interference, blurring, decreased resolution, and other typical distortions that can occur during image acquisition, transmission, or storage. Understanding how these factors affect deep learning architecture performance is crucial, especially in high-stakes diagnostic scenarios like the early detection of neurological disorders and other critical pathologies. These factors are highly relevant in real-world clinical settings where imaging conditions are frequently suboptimal. Furthermore, broadening the dataset to include more clinical conditions and medical imaging modalities (such as CT, MRI, PET, and ultrasound) will improve the generalizability and resilience of machine learning models in a range of healthcare settings. More inclusive and flexible diagnostic tools may result from this increased representation. Future research will also examine the development and application of hybrid frameworks that combine cutting-edge deep neural networks with image-quality-aware preprocessing methods. Especially in settings with limited bandwidth or resources, such architectures may be able to dynamically adjust to different image quality levels, increasing diagnostic precision and computational effectiveness. These paths will help modern medical imaging and telehealth systems develop AI-powered solutions that are dependable, scalable, and equitable.

REFERENCES

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.
[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
[3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.
[4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in *Proc. ECOC'00*, 2000, paper 11.3.4, p. 109.
[5] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.
[6] (2002) The IEEE website. [Online]. Available: http://www.ieee.org/
[7] M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: http://www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran/
[8] *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.
[9] "PDCA12-70 data sheet," Opto Speed SA, Mezzovico, Switzerland.
[10] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.
[11] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.
[12] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.

ISSN: 1737-9237