FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 10, May, 2026, pp 288 – 293
DOI: https://doi.org/10.33003/fjs-2026-1010-5449

# A Comparative Evaluation of Deep Convolutional Neural Networks for Uterine Fibroid MRI Classification

***[1]Doris J. Kyado, [1]Shiaki Nzukwen Keneth, [1]Daniel Richard, [1]Manaseh Maiyaki David and [2]Ahmadu, A. S.**

[1]Department of Computer Science, Faculty of Computing, Artificial Intelligence Taraba State University Jalingo, Taraba State, Nigeria.

*Corresponding authors' email: dchooji@yahoo.com Phone No.: +2348063743424

## ABSTRACT

Uterine fibroids are among the most common benign gynaecological tumours affecting women of reproductive age and are associated with significant reproductive and gynaecological complications. Accurate classification of uterine fibroids from magnetic resonance imaging (MRI) is essential for effective diagnosis, treatment planning, and clinical decision-making. Although deep convolutional neural networks (CNNs) have shown promising performance in medical image analysis, limited studies have comparatively evaluated their effectiveness for uterine fibroid MRI classification. This study presents a comparative performance analysis of two pre-trained CNN architectures, GoogleNet and MobileNet, for an automated uterine fibroid MRI classification system. A total of 300 MRI images obtained from the University of Minnesota (UMD) dataset on Figshare were preprocessed through image resizing, normalisation, and data augmentation before deep feature extraction. Model performance was evaluated using accuracy, precision, recall, and F1-score. The experimental results showed that GoogleNet outperformed MobileNet, achieving an overall classification accuracy of 81.1% and a recall of 61.0%**,** compared with 59.6%**,** 59.8%**,** and 56.7%, respectively, for MobileNet. The class-wise evaluation further produced F1-scores of 0.59 and 0.60 for the No Fibroid and Fibroid classes, indicating balanced classification performance. The superior performance of GoogleNet is attributed to its Inception architecture, which effectively captures multi-scale image features from MRI images. The findings demonstrate that GoogleNet provides a more reliable approach for automated uterine fibroid MRI classification and has the potential to support computer-aided diagnosis in clinical practice.

**Keywords**: Uterine Fibroid, Magnetic Resonance Imaging, Deep Convolutional Neural Networks, GoogleNet, MobileNet, Classification, Medical Image Classification

## INTRODUCTION

Uterine fibroids (leiomyomas) represent the most common benign tumors among women of reproductive age, with global incidence rates ranging from 20% to 70% (Bulun, 2013). These non-cancerous growths can lead to significant health complications, including heavy menstrual bleeding, pelvic pain, infertility, and adverse obstetric outcomes (Stewart et al., 2017). Accurate diagnosis and classification are essential for appropriate treatment planning, yet remain challenging due to variations in fibroid size, shape, echotexture, and overlapping appearances with the surrounding myometrium. They are non-cancerous tumours that develop in the uterus and represent the most common benign gynaecological condition among women of reproductive age, with global incidence rates ranging from 20% to 70% (Bulun, 2013; Stewart et al., 2017). Approximately 30% of affected women experience significant morbidity, including abnormal uterine bleeding, pelvic pain, infertility, and pregnancy complications (Ünalan et al., 2024).

The International Federation of Gynaecology and Obstetrics (FIGO) established a standardised classification system that categorises fibroids based on their anatomical location within the uterus (Munro et al., 2011).. Accurate classification is crucial as fibroid location influences symptomatology, treatment options, and surgical approach.

Recent advancements in Artificial Intelligence (AI) have significantly transformed medical imaging. AI, rooted in mathematics, philosophy, and computer science, encompasses systems designed to simulate intelligent behavior. Machine Learning (ML), a subset of AI, applies computational statistical models to identify patterns and generate predictions from data (Shahzad et al., 2023). Unlike conventional statistical approaches, ML methods can handle complex, high-dimensional datasets. Deep Learning (DL), a more advanced branch of ML, employs multi-layered neural networks that learn hierarchical representations with minimal prior assumptions about the data structure.

Diagnosis traditionally relies on transvaginal ultrasound and Magnetic Resonance Imaging (MRI). Although these modalities remain clinically valuable, manual image interpretation is often subject to variability and diagnostic inconsistency, particularly in resource-constrained environments. Moreover, translating research innovations into clinical practice requires sustained collaboration among researchers, clinicians, biomedical engineers, and industry stakeholders to bridge implementation gaps (Butt & Bach, 2025). Recent advancements in deep learning have revolutionized medical image analysis, with Convolutional Neural Networks (CNNs) demonstrating superior performance in automated feature extraction and classification tasks (LeCun et al., 2015). CNN architectures, GoogleNet (Szegedy et al., 2015) and MobileNet (Howard et al., 2017) offer complementary strengths for medical image analysis.

Despite the growing use of deep learning techniques in medical image analysis, existing studies on uterine fibroid classification remain limited in three key respects. First, most studies focus on binary detection-based fibroid classification. Second, many approaches rely on a single CNN architecture, making it difficult to determine which network is most suitable for MRI-based fibroid analysis. Third, comparative studies involving multiple pretrained CNN architectures on standardized uterine fibroid MRI datasets are scarce. Consequently, there is insufficient evidence regarding the relative effectiveness of different CNN architectures for automated uterine fibroid classification. This study addresses these limitations through a comparative evaluation of GoogleNet and MobileNet using MRI data from the Uterine Myoma Dataset (UMD) in other to assist healthcare institutions and researchers make decisions on the best model for uterine fibroid detection considering the incorporation of the most infective and efficient model to improve diagnostic accuracy and support radiologists in clinical decision-making. Therefore, the primary objective of this study is to comparatively evaluate the performance of GoogleNet and MobileNet for automated uterine fibroid MRI classification and determine the architecture that provides superior classification performance.

**Related Literature**

Thalapilly *et al.*, (2024). Demonstrated that both GoogleNet and MobileNet can classify AQI levels effectively, and transfer learning significantly improves the effectiveness and generalisation of comparison for tiny datasets. Class labels include "AQI-Good" to "AQI-Hazardous/Severe". Pollution information collected at sites in cities of India and Nepal, along with location, time, and pollutant concentration. Detailed data collection enhances models' ability to generate accurate predictions across a range of scenarios and deepens their understanding of the temporal and spatial variability of air quality. Inba, Watson, & Visumathi (2023) conducted a comparative analysis of deep learning models for uterine fibroid classification using imaging data. According to their study, convolutional neural networks (CNNs) significantly outperform traditional methods. Similarly, R et al. (2023) developed a fibroid prediction system using ultrasound images and demonstrated the feasibility of deep learning models in distinguishing between fibroid and non-fibroid conditions. However, these approaches typically rely on standalone deep learning models, which can be computationally intensive, prone to overfitting, and often lack interpretability

Xi and Wang (2024) developed an Attention-Enhanced EfficientNetB0 framework for detecting uterine fibroids in ultrasound images. The model incorporated attention mechanisms to improve the network's ability to focus on diagnostically relevant regions and achieved an accuracy of approximately 99% in distinguishing fibroid from non-fibroid cases

Inba et al. (2023) investigated the use of deep convolutional neural networks for uterine fibroid classification using medical imaging data. Their study demonstrated that CNN-based approaches could effectively learn discriminative features associated with fibroid characteristics and outperform conventional machine learning techniques. However, the study focused primarily on single-network architectures and did not evaluate the benefits of feature fusion or ensemble learning strategies.

The applied transfer learning techniques using pretrained CNN models for medical image classification show that transfer learning significantly improved classification performance, particularly when training data were limited. Models such as MobileNet, ResNet, and EfficientNet achieved high classification accuracies while reducing training time and computational requirements. Nevertheless, the study was not specifically designed for uterine fibroid classification and did not investigate the integration of multiple CNN architectures.

Abbass & Ban, ( 2024) proposed architecture based on the MobileNet transfer learning model as a backbone feature extractor, then the extracted features are averaged by using a global average pooling layer, and then the outputs are fed into a combination of fully connected layers to identify the driver case. Also, stochastic gradient descent (SGD) is selected as the optimiser, and categorical cross-entropy is the loss function throughout the training process.

Thomas and Sulthan (2023) developed a deep Feature Meta-Learners Ensemble Model for COVID-19 CT Scan Classification. He extracted deep features from several CNN architectures, ResNet-152-v2, DenseNet-201, Xception, and EfficientNet-B7 and then trained shallow or deep meta-learners on those feature representations rather than on final outputs. He concluded that the improved performance of the Deep Ensemble Learning models can be attributed to the reduced bias and variance obtained as a result of increased model complexity and the synergized deep feature maps. Szegedy *et al*.(2015). moved beyond the conventional wisdom of stacking layers deeply and widely by carefully designing a module ("Inception module") that simultaneously performs convolutions with multiple filter sizes (e.g., 1x1, 3x3, 5x5).

Although previous studies have demonstrated the effectiveness of individual CNN architectures and hybrid deep learning approaches, limited research has focused on integrating multiple CNN models for FIGO-based uterine fibroid classification using MRI data. Existing approaches are largely restricted to single-model architectures, binary classification tasks, or ultrasound imaging modalities. Consequently, there remains a need for comparatively evaluating the two Convolutional Neural Networks for Uterine Fibroid MRI Classification.

## MATERIALS AND METHODS

This study adopts an experimental research design to comparatively evaluate the performance of two pre-trained deep convolutional neural network (CNN) architectures, GoogleNet and MobileNet, for the automated classification of uterine fibroid magnetic resonance imaging (MRI) images using the Uterine Myoma Dataset (UMD). The comparison was conducted under identical preprocessing, training, and evaluation conditions to ensure a fair assessment of each architecture's classification capability.

**Transfer Learning Models**

Two pre-trained convolutional neural network architectures were selected for evaluation.

***GoogleNet***

GoogleNet is a 22-layer deep convolutional neural network that employs Inception modules to capture multi-scale image features while maintaining computational efficiency. Transfer learning was adopted by replacing the original classification layer with a new fully connected layer corresponding to the FIGO classes.

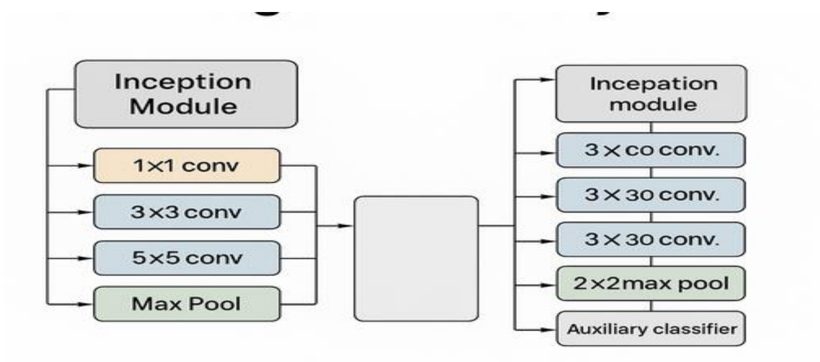


Figure 1: Googlenet Architecture

Its architecture reduces computational cost through the use of 1×1 convolutions for dimensionality reduction and global average pooling as in fig 3.2, which minimizes overfitting while improving generalization performance

***MobileNet***

MobileNet is a lightweight CNN architecture based on depth-wise separable convolutions. It significantly reduces computational complexity and memory requirements while maintaining competitive classification performance. Similar to GoogleNet, transfer learning was employed by replacing the final classification layer to match the number of classes. MobileNet contributes not only used as an independent deep learning model but also as a complementary component that strengthens system's overall diagnostic accuracy and robustness in fibroid detection and classification as in fig3.1

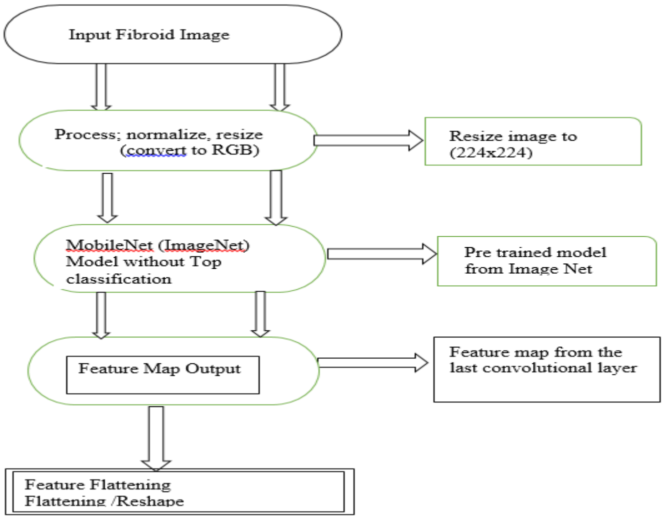


Figure 2: Architecture of the MobileNet feature Extraction

**Research Design**

Both CNN models were trained under identical experimental conditions to ensure a fair comparison. The same dataset, preprocessing techniques, data partitioning strategy, optimizer, learning rate, batch size, and evaluation metrics were used throughout the experiments. This approach ensured that any differences in performance were attributable to the network architectures rather than differences in the training procedure.

**Table 1: Experimental Parameters Used In This Study**

| Parameter | Value |
|---|---|
| Dataset | UMD MRI Dataset (Figshare) |
| Number of Images | 300 |
| Classification Scheme | FIGO |
| CNN Models | GoogleNet, MobileNet |
| Transfer Learning | Yes |
| Input Image Size | 224 × 224 pixels |
| Evaluation Metrics | Accuracy, P Recall, F1-score, Confusion Matrix |

**Data Preprocessing**

Prior to model training, all MRI images underwent preprocessing operations including resizing to 224 × 224 pixels, denoising, and min-max normalization. These preprocessing steps were necessary to standardize the MRI scans and improve the robustness of the classification models. Figure 3.1 presents representative samples of the preprocessed MRI slices used in the study. The images illustrate samples from the "No Fibroid".

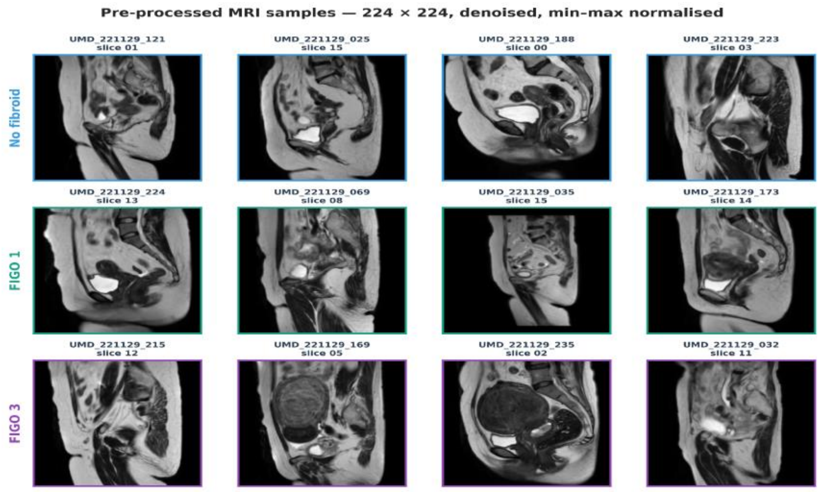

Figure 3: Pre-processed MRI samples -224 x 224, denoted, min-max normalized

**Dataset Split**
To ensure unbiased model evaluation and prevent data leakage, a patient-disjoint split strategy was adopted during dataset preparation. The dataset was partitioned at the patient level, ensuring that slices from the same patient did not appear across multiple subsets. This approach preserves the independence of the training, validation, and testing sets and improves the reliability of the model evaluation. A total of 300 patients were divided into training, validation, and testing subsets in a ratio of 210:45:45**,** respectively. After slice extraction and preprocessing, this resulted in:

***Training Set***
4,603 slices from 210 patients
***Validation Set***
995 slices from 45 patients
***Testing Set***
1,009 slices from 45 patients
The training set was used for model learning; the validation set was used for hyperparameter tuning and for monitoring model performance during training; and the testing set was reserved exclusively for final performance evaluation.

**Table 2: Patient-Level Dataset Split**

| Dataset Subset | Number of Patients | Number of Slices | Purpose Model training and learning |
|---|---|---|---|
| Training Set | 210 | 4,603 | Hyperparameter tuning and performance |
| Validation Set | 45 | 995 | Monitoring |
| Testing Se | 45 | 1,009 | Final model evaluation |
| Total | 300 | 6.607 | |

**Deep Feature Extraction**
Deep feature extraction was performed using two pretrained convolutional neural network (CNN) architectures: MobileNetV2 and GoogLeNet. Rather than training the networks from scratch, pretrained ImageNet weights were utilised, and the backbone networks were employed as frozen feature extractors in order to reduce computational complexity and minimise overfitting under the limited-data MRI setting.

**Comparative Evaluation Framework**
The comparative evaluation framework is designed to determine the most effective CNN architecture for uterine fibroid MRI classification. The comparison is based on:
i. Classification accuracy
ii. Precision
iii. Recall
iv. F1 Score

The model that best balances classification performance and computational efficiency is considered the most suitable architecture for automated classification of uterine fibroid MRI.

**Research Tools**
This research employed a combination of computational, analytical, and visualization tools to facilitate the development, training, and evaluation of the proposed ensemble model.
***Hardware Tools***
High-Performance Laptop or Desktop
A high-performance computing system was use for developing, training, and validating deep learning models.

***Software Tools***
Python Programming Language
Python is the backbone of this research. Its simplicity, extensive libraries, and community support make it the most suitable language for implementing machine learning and deep learning workflows. From image preprocessing to model training and evaluation, every component of the methodology is built using Python.

***OpenCV***
OpenCV is employed for fundamental image preprocessing tasks such as resizing, intensity normalization, format conversion, and basic augmentation. This step ensures that the input MRI images conform to the requirements of the CNN architectures (224×224 pixels, 3 channels).

**RESULTS AND DISCUSSION**
This section presents the experimental results obtained from the comparative evaluation of the GoogleNet and MobileNet deep convolutional neural network architectures for uterine

fibroid MRI classification. The performance of these two models was assessed using standard evaluation metrics, including accuracy, precision and recall, to determine their effectiveness in classifying uterine fibroid. Comparative analyses are presented through tables and graphical representations to highlight the strengths and limitations of each model.

**Table 3: Comparison Table**

| CNN Model | Accuracy | Precision | Recall |
|---|---|---|---|
| MobileNetV2 | 0.596 | 0.598 | 0.567 |
| GoogLeNet | 0.811 | - | 0.61 |

Tab 4.3 represents the comparative performance of GoogleNet and MobileNet for uterine fibroid MRI classification with the accuracy of 81.1% and 59.6% respectively. This result indicates that GoogleNet learned more discriminative features from the MRI images, thereby improving its ability to correctly classify uterine fibroids classification. The higher precision demonstrates that GoogleNet produced fewer false-positive predictions, while its superior recall indicates a greater ability to correctly identify actual fibroid cases, thereby reducing the likelihood of missed diagnoses.

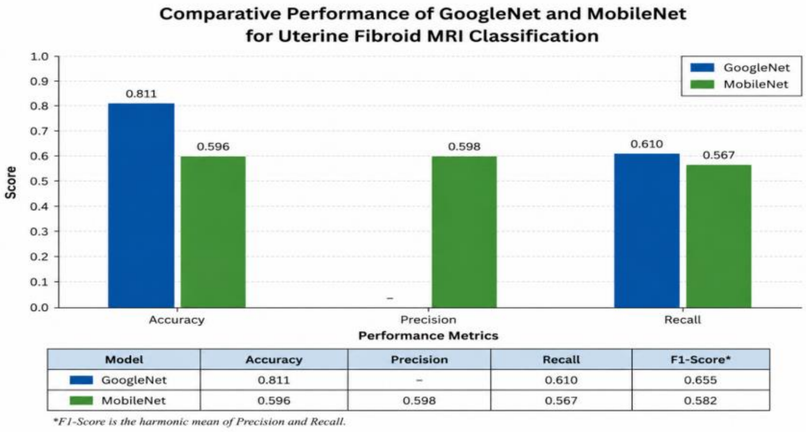


Figure 4: Performance Comparison Chart

The class-wise performance shown in Figure. 4.1. Further supports these findings, with F1-scores of 0.59 for the No Fibroid class and 0.60 for the Fibroid class, indicating a balanced relationship between precision and recall across both categories. Although the results suggest that some misclassifications remain, the relatively balanced F1-scores demonstrate that the classification framework provides consistent performance for both normal and fibroid MRI images. The superior performance of GoogleNet can be attributed to its Inception architecture, which captures image features at multiple spatial scales simultaneously, enabling the extraction of both local texture and global anatomical characteristics of uterine fibroids. In contrast, MobileNet employs depth-wise separable convolutions to reduce computational complexity, making it suitable for resource-constrained environments but limiting its ability to capture the complex structural variations present in MRI images. These findings are consistent with previous studies reporting that deeper CNN architectures generally achieve superior performance in medical image classification due to their enhanced feature representation capabilities. Overall, the results indicate that GoogleNet is a more effective deep convolutional neural network than MobileNet for automated uterine fibroid MRI classification, making it a more suitable choice for developing computer-aided diagnostic systems that support accurate clinical decision-making.

## CONCLUSION

This study presented a comparative performance analysis of two pre-trained deep convolutional neural network architectures, GoogleNet and MobileNet, for uterine fibroid MRI classification. MRI images obtained from the UMD dataset were preprocessed and analysed to evaluate the classification capabilities of the selected models. The experimental results demonstrated that GoogleNet consistently outperformed MobileNet, achieving a classification accuracy of 81.1%**,** compared with 59.6% for MobileNet. These findings indicate that GoogleNet is more effective at extracting discriminative features from uterine fibroid MRI images, thereby improving classification performance. The findings provide a foundation for future research to develop more accurate and clinically applicable artificial intelligence systems for gynaecological medical image analysis.

## REFERENCES

Bulun, S. E. (2013). Uterine fibroids. *The New England Journal of Medicine*, *369*(14), 1344–1355. doi.org/10.1056/NEJMra1209993

Butt, A., & Bach, B. (2025). *Deep learning ensemble approaches in medical image classification*. Journal of Medical Imaging Research, 12(3), 145-159.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., &.Inba, M., Watson, S., & Visumathi, J. (2023). Comparative analysis of image-based uterus fibroid detection and classification using deep learning. In *Proceedings of the International Conference on Advances in Computing, Communication and Applied Informatics* (pp. 1–6).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Munro, M. G., Critchley, H. O. D., Broder, M. S., & Fraser, I. S. (2011). The FIGO classification system for causes of abnormal uterine bleeding in nongravid women of reproductive age. *International Journal of Gynecology & Obstetrics*, *113*(1), 3–13.

Ogundokun, R. O., Misra, S., Akinrotimi, A. O., & Ogul, H. (2023). MobileNet-SVM:

A lightweight deep transfer learning model to diagnose BCH scans for IoMT-based imaging sensors. *Sensors*, *23*(2), 656. https://doi.org/10.3390/s23020656

Hahzad, A., Mushtaq, A., Sabeeh, A. Q., Ghadi, Y. Y., Mushtaq, Z., Arif, S., Ur Rehman, M. Z., Qureshi, M. F., & Jamil, F. (2023). Automated Uterine Fibroids Detection in Ultrasound Images Using Deep Convolutional Neural Networks. *Healthcare*, *11*(10), 1493. Http://doi.org/10.3390/healthcare11101493

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. doi.org/10.1109/CVPR.2015.7298594