MRI-Based Brain Tumor Detection for the African Context

Jideofor J. Umeadi JOHNUMEADI16@GMAIL.COM
Department of Computer Engineering
University of Lagos

Editor: Sakinat Folorunso, Roseline Ogundokun, and Francisca Oladipo

Abstract

The chronic shortage of radiologists along with the predominance of low-field MRI machines in sub-Saharan Africa adversely affects brain tumor diagnosis in the region. This study evaluates the utility of lightweight transfer learning for binary brain tumor detection under these constraints, with a focus on generalisation across a domain shift from Western-heavy training data to actual Nigerian clinical scans. The best results achieved were from a frozen MobileNetV2 backbone, pre-trained on ImageNet, trained on a skull-stripped Kaggle MRI dataset and evaluated externally on the BraTS-Africa cohort—a dataset of tumor-positive scans from six Nigerian diagnostic centres. Skull-stripping preprocessing, applied to mitigate the Clever Hans effect, improved mean external sensitivity from $52.67\% \pm 24.04\%$ to $85.67\% \pm 13.58\%$ across three random seeds, with 100% sensitivity on the best seed run. These results demonstrate that targeted domain alignment through simple preprocessing is a viable way to close the generalisation gap in low-resource settings, without requiring additional compute. However, the model's specificity on healthy Nigerian scans remains uncharacterised, and it should be understood strictly as a high-sensitivity screening tool requiring human oversight rather than a standalone diagnostic system. Ongoing work is focused on curating a representative Nigerian MRI dataset to enable a complete sensitivity/specificity profile under realistic deployment conditions, alongside benchmarking against comparable lightweight architectures and on-device validation.

Keywords: Brain Tumor Detection, MobileNetV2, Transfer Learning, Domain Shift, Skull Stripping, Low-Resource Settings, Africa, BraTS-Africa

1. Introduction

The stunning ability of modern medicine to accurately diagnose and treat complex conditions, from brain tumors to heart disease, relies heavily on the insights provided by medical imaging. Healthcare has been revolutionised by imaging technologies like MRI and CT scans as they provide a non-invasive window into the human body, enabling the precise diagnosis of diseases that were once undetectable.

While this technological advancement has improved the quality of life, its effectiveness is still reliant on human interpretation of generated medical images. Specially trained healthcare professionals (particularly radiologists) are needed to correctly interpret these images and use them for accurate diagnosis of certain diseases (Chen et al., 2023). This dependence creates significant bottlenecks in regions where such professionals are in short supply; particularly in Sub-Saharan Africa. Recent studies estimate that Nigeria has fewer than 700 radiologists serving a population of over 200 million, resulting in a radiologist-to-patient ratio of roughly 1:660,000 (Udam et al., 2025) compared to 50–120 per million in most European countries (Brady et al., 2025). This severe workforce shortage leads

to prolonged "radiology report turnaround time (TAT)," a metric identified in literature as a key indicator of service quality. When TAT is measured in days or even weeks, the diagnostic interval is compromised (Ritchie et al., 2025). In low-resource settings, this scarcity often results in delayed or non-specialist image interpretation, increasing diagnostic risk and uncertainty (Tahir et al., 2022).

This profound gap highlights the need for a rethinking of current diagnostic workflows. Deep Learning, particularly Convolutional Neural Networks (CNNs), could offer a promising pathway to multiply a severely understaffed workforce. However, a review of existing literature reveals a mismatch between current AI research and the reality in Africa. Most state-of-the-art medical imaging models are trained and evaluated on Western datasets (Ledford, 2019) and they often fail to generalise to the heterogeneous, lower-quality imaging environments found in developing regions, a phenomenon known as "domain shift." In addition, high-performing models usually require large compute resources which are unavailable in most Nigerian clinics, making them impressive on paper but undeployable practically.

In light of these challenges, this study aims to bridge this gap by evaluating the utility of lightweight computer vision for binary brain tumor detection (tumor present vs. absent) particularly in low-resource settings, with an emphasis on scalable, context-specific deployment. This work investigates the impact of data preprocessing on generalisation in a domain shift. We benchmark the performance of a highly efficient MobileNetV2 architecture when transferred from standard public datasets (Kaggle) to an external dataset representing African patients (BraTS-Africa). The MobileNetV2 model was selected due to its ability to achieve inference times under 50ms on common mobile devices (Sandler et al., 2018). Also, the "Clever Hans" effect—where models appear to be very accurate but are just overfitting to unimportant nuances (Lapuschkin et al., 2019)—was specifically analysed to demonstrate how domain alignment techniques can restore diagnostic confidence.

The methodology details the dataset preprocessing, the lightweight training pipeline, and the rigorous experimentation conducted to ensure robustness in a domain shift. The results section presents a comparative analysis of internal accuracy versus external sensitivity, while the discussion interprets these findings for safe, scalable diagnostics in under-resourced African settings.

2. Related Work

The use of CNNs in the binary detection of brain tumors has been widely studied and recent work has favoured the use of lightweight architectures for classification tasks. MobileNetV2 in particular has been applied to Kaggle-sourced brain MRI data (Yuliawan, 2026), achieving strong internal accuracy but without external validation. Deeper models like DenseNet121 have also been evaluated but tend to overfit with limited training data, suggesting that frozen lightweight backbones are a more stable choice under low-data constraints.

The Clever Hans effect has been identified across a number of models, where models exploit image-specific artifacts like skull contours or acquisition noise rather than clinical features (Wallis and Buvat, 2022). Skull stripping is an established preprocessing step in neuroimaging for isolating brain tissue, though its utility as a domain adaptation strategy—particularly its interaction with shortcut learning in MRI-based classification—has been

further investigated in the context of Alzheimer's disease classification (Tinauer et al., 2025), but its application specifically for African clinical data has not been systematically investigated.

There is limited work in the medical AI field that addresses the same deployment context as this paper: sub-Saharan Africa with predominantly low-field imaging hardware. Adewole et al. (2025) introduced the first curated cohort of brain tumor scans from Nigerian diagnostic centres, representing a significant step in closing this gap; however, existing work leveraging this dataset focuses on segmentation rather than lightweight binary classification. To the best of our knowledge, this study is the first to report external validation of a lightweight binary tumor detector on Nigerian clinical MRI data, with preprocessing guided by domain alignment rather than standard convention.

Table 1: Contextualisation within recent medical imaging literature (2021–2026)

Paper	Dataset(s)	Model	External Valid.?	Light-weight?	African Data?	Domain Shift?	Key Limitation
Yuliawan (2026)	Kaggle	MobileNetV2	No	Yes	No	No	No external validation
Adewole et al. (2025)	BraTS-Africa	U-Net/SOTA	Yes	No	Yes	Yes	Focus on segmentation
Wallis and Buvat (2022)	Nanfang/Public	SVMs	No	Yes	No	Yes	No proposed solution
Tinauer et al. (2025)	ADNI	3D CNN	No	No	No	No	AD task specific
This Work	Kaggle/BraTS-Africa	MobileNetV2	Yes	Yes	Yes	Yes	Lack of healthy scans

3. Methodology

3.1. Datasets and Curation

Two distinct datasets representing the training environment and the real-world deployment environment were used to ensure the model was robust enough for deployment.

Training Dataset (Source Domain): A publicly available brain tumor MRI dataset obtained from Kaggle was employed for training and internal validation of the model. The images were collected from publicly available medical imaging repositories and organised into two separate folders (one for scans with a tumor and the other for healthy scans). However, initial exploratory analysis revealed skull and scalp tissue in the "healthy" class, necessitating skull-stripping preprocessing (see Section 3.2). The raw Kaggle dataset comprised 253 images, split 80/20 into 203 training and 50 validation images. Following skull stripping, 80 images were discarded due to failed contour detection, yielding 173 images (139 train / 34 validation). For internal evaluation of the baseline model, a class-balanced subset of 34 images (17 healthy, 17 tumor) was drawn from the 50-image validation set to ensure fair metric computation.

External Validation Dataset (Target Domain): The BraTS-Africa dataset (Adewole et al., 2025), a cohort specifically developed to represent brain tumor populations within Africa, was used as the external validation set to assess the model's ability to generalise to new data. This dataset was accessed via The Cancer Imaging Archive (TCIA) and included clinical scans obtained from six diagnostic centres in Nigeria between January 2010 and December 2022. The images were acquired using 1.5T MRI scanners, which is important because this hardware is representative of what is typically found in low-resource clinical settings. While the original data was provided as 3D volumes across four different

imaging modalities (T1, T1ce, T2, and FLAIR), a 3D-to-2D pipeline was applied to extract informative axial slices (see Section 3.2). It is noted that this cohort consists exclusively of tumor-positive cases; consequently, external specificity could not be evaluated from this dataset. The model's behaviour on healthy Nigerian scans is addressed in Section 6, and expanding this evaluation is the focus of ongoing work described in Section 7.1.

3.2. Preprocessing and Domain Alignment

Skull Stripping: Skull stripping was applied to the training data to remove unimportant cues that were being learned during training. This step made the visual structure of the training data similar to the external dataset and reduced the model's reliance on dataset-specific artifacts—an intervention that reduced raw accuracy but improved generalisation. Earlier models trained on the raw data achieved an artificially high accuracy by exploiting non-brain artifacts rather than learning clinically relevant features, an instance of the Clever Hans effect.

3D to 2D Transformation Pipeline: The external dataset was provided as 3D volumes, so a custom pipeline was developed to adapt the data for a 2D CNN architecture. Axial slices were extracted from each volume, and to ensure meaningful supervision, non-informative slices were filtered out, leaving only slices containing visible tumor regions. All retained slices were converted to PNG format and resized to 224×224 pixels to match the model input resolution.

3.3. Model Architecture

To optimise for deployment on resource-constrained devices, the MobileNetV2 architecture was selected as the backbone and initialised using weights pre-trained on the ImageNet dataset to leverage transfer learning. We adapted the model for binary classification by modifying the final layers as follows:

- A Global Average Pooling layer was applied to condense the spatial feature maps.
- This was followed by a fully connected layer containing 128 neurons with a ReLU activation function to learn non-linear patterns.
- To minimise overfitting, a Dropout layer with a rate of 0.3 was introduced.
- The network concludes with a single output neuron using a sigmoid activation function to produce a binary probability.

3.4. Training Configuration

The model was implemented using TensorFlow/Keras and the Adam optimiser with binary cross-entropy loss was used during training. Data augmentation was applied to increase sample diversity and improve the model's robustness under limited data. The augmentation pipeline included random horizontal flipping, rotations (up to 10°), zooming (up to 10%), contrast adjustments (up to 10%), spatial translations (up to 10% of image dimensions) and Gaussian noise injection. Training was carried out for a maximum of 30 epochs with a batch size of 32 and a learning rate initialised at 5×10^{-4}. Training stability was enforced using

callback-based controls: early stopping based on validation accuracy, adaptive learning-rate reduction and checkpointing of the best-performing weights. To ensure that the model did not just hit a "sweet spot" with a particular initialisation seed, multiple training runs were conducted using different random seed values (23, 42, and 2026). The final model was selected based on a fair trade-off between validation accuracy and sensitivity on external data rather than peak accuracy alone.

All experiments were implemented in Python using the TensorFlow/Keras framework and executed on the Google Colab platform with freely available GPU acceleration.

4. Results

The experiments were conducted across three random seeds (23, 42, and 2026) to evaluate reproducibility and the impact of skull-stripping preprocessing on the frozen MobileNetV2 model's performance. Measured metrics were focused on internal validation (a held-out subset of the Kaggle data) and external validation (BraTS-Africa, 150 tumor-positive slices). The results below are reported as means ± standard deviation, revealing how a minor trade-off on internal accuracy interestingly improved generalisation.

4.1. Internal Performance (Kaggle Dataset)

On the internal validation set, the frozen MobileNetV2 admittedly performed better without skull-stripping, demonstrating strong performance with a mean validation accuracy of $84.00\% \pm 4.00\%$, while it achieved a mean validation accuracy of $70.35\% \pm 5.88\%$ with stripping. Sensitivity remained high (mean $82.35\% \pm 5.88\%$ with stripping), while specificity was moderate (mean $47.06\% \pm 5.88\%$ with stripping), so the model accepts a higher rate of false positives to minimise the risk of missing a brain tumor—a primary necessity for a screening tool in underserved regions. Accuracy on the full internal set was strong, averaging $91.31\% \pm 1.97\%$ with stripping.

Table 2: Internal Validation Metrics (Mean ± SD Across 3 Seeds for Frozen MobileNetV2)

Preprocessing	Val Acc. (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUPRC
Without Stripping	84.00 ± 4.00	91.33 ± 3.75	90.20 ± 3.40	68.00 ± 12.49	85.00 ± 5.00	0.92 ± 0.04
With Stripping	70.35 ± 5.88	91.31 ± 1.97	82.35 ± 5.88	47.06 ± 5.88	78.00 ± 4.32	0.89 ± 0.05

4.2. External Performance (BraTS-Africa Dataset)

External validation revealed a significant domain shift, one which the baseline model could not handle. The frozen MobileNetV2 achieved a mean sensitivity of $85.67\% \pm 13.58\%$ with stripping, compared to $52.67\% \pm 24.04\%$ without. Since the external BraTS-Africa cohort consists only of scans with tumors, metrics dependent on true negatives (Specificity, Precision, AUC-ROC) are not applicable. Sensitivity is therefore relied on as the definitive performance metric for this dataset. The absence of healthy control scans in the external cohort means that specificity under real Nigerian deployment conditions could not be fully

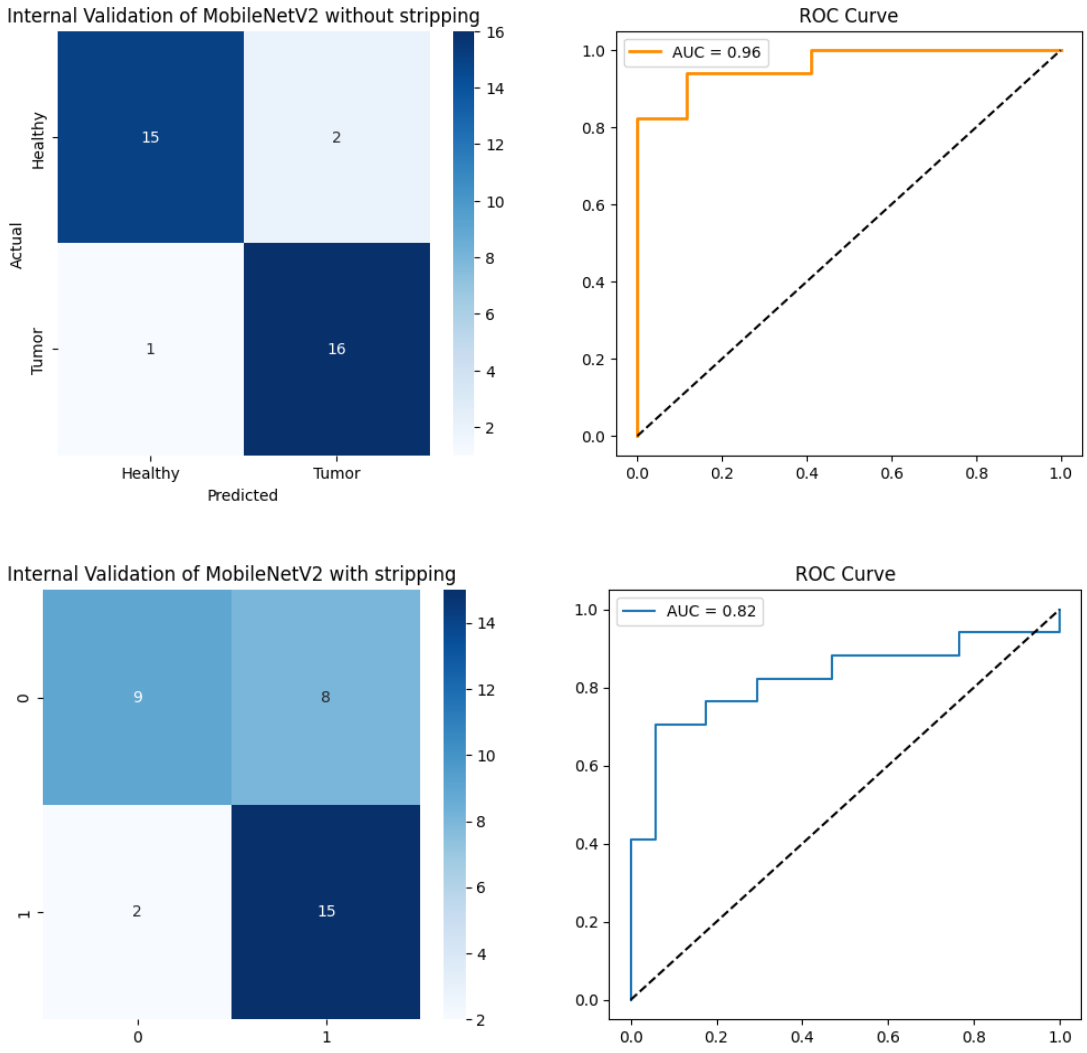

Figure 1: Internal Validation Performance Comparison (best run, seed 2026). **(a)** Baseline Model (No Skull-Stripping): Demonstrates high classification metrics and high specificity (15/17 healthy cases correctly identified). While visually impressive, this performance on the source domain stands in sharp contrast to the model's failure on external data, suggesting overfitting to dataset artifacts. **(b)** Proposed Model (Skull-Stripping): Shows a reduction in AUC (0.82) and specificity (9/17 healthy cases identified), reflecting the deliberate "safety-first" trade-off. Despite the increase in false positives (8 cases), the model maintains robust sensitivity (15/17 tumors detected), ensuring that life-threatening pathology is not missed during screening.

assessed in this study; this gap is acknowledged and directly motivates the data curation work described in Section 7.1.

Table 3: External Validation Metrics (Mean ± SD Across 3 Seeds for Frozen MobileNetV2)

Preprocessing	Sensitivity (%)
Without Stripping	52.67 ± 24.04
With Stripping	85.67 ± 13.58

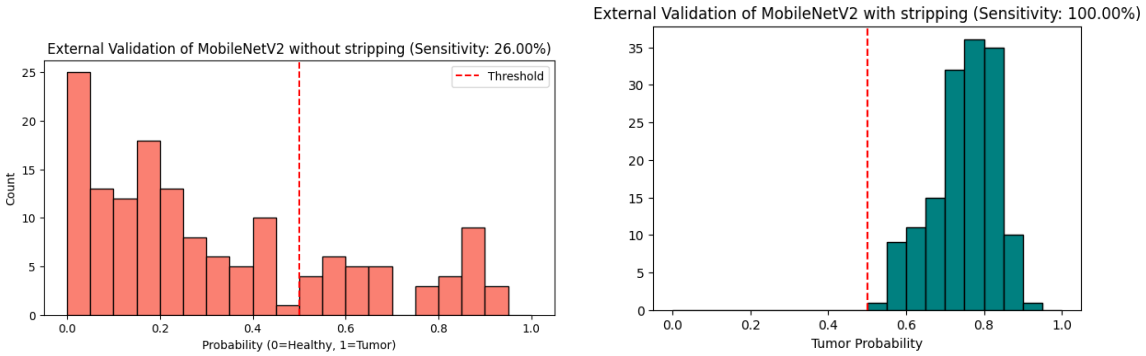

Figure 2: External Generalisation on BraTS-Africa Dataset (best run, seed 2026). **(a)** Baseline Model (No Skull-Stripping): The distribution of predictions is heavily skewed towards the left (probability of tumor < 0.5), resulting in a dangerously low sensitivity (26.00%). **(b)** Proposed Model (Skull-Stripping): The distribution shifts entirely to the tumor range (probability > 0.5), achieving 100.00% sensitivity on the best run.

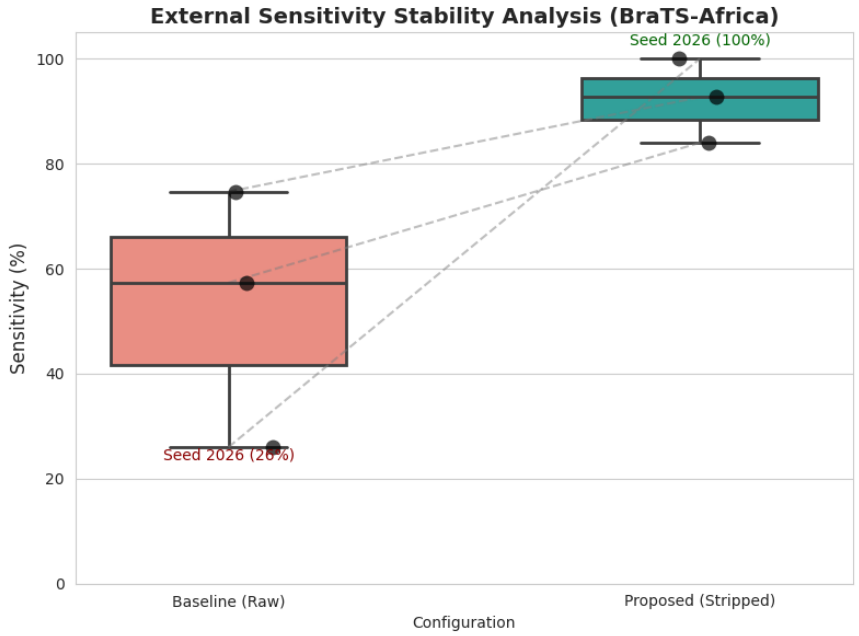

Figure 3: Paired Stability Analysis of External Sensitivity. The plot compares the sensitivity of the Baseline (Raw) and Proposed (Stripped) models across three random seeds (23, 42, and 2026). The baseline model exhibits severe volatility; note the catastrophic failure of Seed 2026 (26% sensitivity). The proposed model demonstrates superior robustness, narrowing performance variance significantly. The steepest improvement (Seed 2026) reflects a jump from 26% to 100% sensitivity, highlighting the critical role of artifact removal in model convergence.

5. Discussion

The results obtained illustrate the potential effectiveness of transfer learning with MobileNetV2 as the frozen base model for binary brain tumor detection in low-resource settings, particularly when enhanced by skull-stripping preprocessing. The model's fairly high performance on internal validation (sensitivity $82.35 \pm 5.88\%$ with stripping) and improved external generalisation with stripping (mean sensitivity $85.67 \pm 13.58\%$) as opposed to $52.67\% \pm 24.04\%$ without stripping, highlight its potential as a triage tool in Nigerian clinics, where radiologist shortages are directly linked to prolonged TAT and high misdiagnosis rates (up to 40% for non-radiologists). Skull stripping improved the model's generalisation by removing unnecessary artifacts, ensuring it prioritised clinically relevant features, as illustrated by the high sensitivity score of the best seed run (seed 2026, 100%).

However, the model's internal specificity of $47.06 \pm 5.88\%$ with skull stripping requires careful consideration as it implies that nearly one in two healthy patients could be incorrectly flagged for specialist review and it currently poses a risk of compounding the very bottleneck the tool aims to solve. This specificity figure reflects the model's source-domain performance only; its false positive rate on healthy Nigerian scans under real deployment remains unknown due to the tumor-only composition of the BraTS-Africa cohort. The model should therefore be strictly understood as a high-sensitivity screening tool that requires human oversight, not a standalone diagnostic system.

While 100% external sensitivity on the best seed run demonstrates effective tumor flagging, the external validation cohort consisted exclusively of tumor-positive cases, meaning the model's specificity on real Nigerian healthy scans could not be fully characterised—a gap directly reflected in the source-domain specificity concern raised above. Architectural simplicity and targeted data preparation outperformed more complex approaches in this low-data environment (253 raw images; 173 retained post-stripping). Fine-tuned MobileNetV2 and custom CNN models were explored as potential configurations but excluded after consistently failing to generalise to the external dataset. In the case of fine-tuning, unfreezing the pretrained backbone on such a small training set likely caused catastrophic forgetting—where the model overwrites pretrained representations with dataset-specific patterns—resulting in 0% external sensitivity. This collapse may have been accelerated by the absence of regularisation techniques such as layer-wise unfreezing, and the negative outcome reinforces the decision to adopt a fully frozen backbone as the final configuration.

These findings offer practical insights for deploying scalable AI solutions in areas with fragmented healthcare data, such as Sub-Saharan Africa. By utilising a frozen model optimised for fast inference (under 50ms) (Sandler et al., 2018), this approach enables edge deployment in rural areas with limited power and internet connectivity, serving as a much needed initial interpretation tool where specialist review is usually unavailable.

The deployment of this tool raises important ethical considerations. The risk of overreliance is particularly acute in settings where specialist oversight is already limited. Furthermore, the model's representation of healthy brain tissue is entirely derived from Western imaging data—local healthy Nigerian scans were absent from both training and external validation—raising structural concerns about diagnostic equity that extend beyond demographic mismatch alone. Any clinical integration should therefore be preceded by local

validation on representative data and accompanied by structured human oversight protocols.

6. Limitations

While these results are promising, a number of limitations should be addressed. Firstly, the small training dataset—203 images (baseline) and 139 images (skull-stripped)—contributed to some seed-dependent variations in performance depending on how the data was split, a common issue in low-resource machine learning. Furthermore, because the external BraTS-Africa cohort consisted exclusively of tumor-positive cases, the model's specificity under real Nigerian deployment conditions could not be assessed. This is a material limitation: a screening tool with unknown specificity cannot be fully characterised for clinical safety, particularly given the already-elevated false positive rate observed on the source domain ($47.06\% \pm 5.88\%$). Expanding the external evaluation to include healthy Nigerian scans is the immediate priority of ongoing work. Although preprocessing improved the model's external performance, it slightly reduced internal accuracy, and the model was not tested on actual mobile hardware to confirm its speed in the field.

7. Conclusion & Future Work

This study demonstrates that standard transfer learning can often be misleading, attaining high accuracy by overfitting to imaging artifacts rather than relevant features. It has been illustrated that addressing a domain shift through simple preprocessing can offer a scalable, deployable solution for underserved regions, potentially multiplying understaffed workforces as it prioritises safety metrics (external sensitivity) over high internal accuracy.

7.1. Future Work

The most immediate priority is expanding the Nigerian MRI dataset to include a substantial number of both tumor-positive and healthy low-field scans. The current external validation was limited to tumor-positive BraTS-Africa cases, leaving the model's specificity on locally acquired healthy scans uncharacterised at scale. Curating a dataset large enough to contribute both to training and to a held-out external test set will enable the first complete sensitivity/specificity profile under realistic 0.3T deployment conditions, and will allow low-field healthy brain appearances to be incorporated into the training distribution directly.

Beyond data expansion, future research should explore domain adaptation techniques that train the model to be invariant to scanner hardware differences, focusing exclusively on pathological features regardless of acquisition quality. Benchmarking against comparable lightweight architectures such as EfficientNet-Lite under the same domain shift conditions would also clarify whether the observed gains are architecture-specific or preprocessing-driven. Future evaluations should move from slice-level to patient-level analysis to further enhance clinical realism, and on-device latency testing on representative Nigerian clinical hardware remains necessary to confirm real-world deployment feasibility. Finally, implementing model ensembles or training on larger datasets could help stabilise performance across seeds, moving closer to a reliable AI-driven diagnostic tool specifically adapted to Sub-Saharan Africa.

Data Availability and Reproducibility

The Kaggle brain MRI dataset used for training is publicly available at https://www.kaggle.com/datasets/arwabasal/brain-tumor-mri-detection. The BraTS-Africa external validation dataset is accessible via The Cancer Imaging Archive (TCIA) at https://www.cancerimagingarchive.net/collection/brats-africa/. The full experimental pipeline, comprising the 3D-to-2D slice extraction notebook and the training and evaluation notebook, is publicly available at https://drive.google.com/drive/folders/1nAHtlBPu5dOM6DlPGPeZzpZhjZrgaq59.

Acknowledgments

The author would like to thank the creators of the BraTS-Africa dataset for making this valuable resource available for research, and the University of Lagos for institutional support. The author declares no conflict of interest.

References

M. Adewole, J.D. Rudie, A. Gbadamosi, D. Zhang, C. Raymond, J. Ajigbotoshso, O. Toyobo, K. Aguh, O. Omidiji, R. Akinola, M.A. Suwaid, A. Emegoakor, N. Ojo, C. Kalaiwo, G. Babatunde, A. Ogunleye, Y. Gbadamosi, K. Iorpagher, M. Onuwaje, B. Betiku, J. Cakmak, B. Menze, U. Baid, S. Bakas, F. Dako, A. Fatade, and U.C. Anazodo. The BraTS-Africa dataset: Expanding the brain tumor segmentation data to capture African populations. *Radiology: Artificial Intelligence*, 7(4):e240528, 2025. doi: 10.1148/ryai.240528.

A.P. Brady, C. Loewe, B. Brkljacic, G. Paulo, M. Szucsich, and M. Hierath. Guidelines and recommendations for radiologist staffing, education and training. *Insights into Imaging*, 16(1), 2025. doi: 10.1186/s13244-025-01926-6.

J. Chen, Z. Gandomkar, and W. Reed. Investigating the impact of cognitive biases in radiologists' image interpretation: A scoping review. *European Journal of Radiology*, 166:111013, 2023. doi: 10.1016/j.ejrad.2023.111013.

S. Lapuschkin, S. Wäldchen, A. Binder, G. Montavon, W. Samek, and K.-R. Müller. Unmasking clever Hans predictors and assessing what machines really learn. *Nature Communications*, 10(1), 2019. doi: 10.1038/s41467-019-08987-4.

H. Ledford. Millions of black people affected by racial bias in health-care algorithms. *Nature*, 574(7780), 2019. doi: 10.1038/d41586-019-03228-6.

B. Ritchie, L. Summerville, M. Sheng, M. Choi, S. Tirumani, and N. Ramaiya. Impact of turnaround time in radiology: The good, the bad, and the ugly. *Current Problems in Diagnostic Radiology*, 2025. doi: 10.1067/j.cpradiol.2025.04.018.

M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. MobileNetV2: Inverted residuals and linear bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018. doi: 10.1109/cvpr.2018.00474.

M.Y. Tahir, M. Mars, and R.E. Scott. A review of teleradiology in Africa – towards mobile teleradiology in Nigeria. *South African Journal of Radiology*, 26(1), 2022. doi: 10.4102/sajr.v26i1.2257.

C. Tinauer, M. Sackl, R. Stollberger, R. Schmidt, S. Ropele, and C. Langkammer. Skull-stripping induces shortcut learning in MRI-based Alzheimer's disease classification, 2025.

N.G. Udam, P.A. Obaje, A.C. Emegoakor, E.L. Edache, E.O. Obaje, A.O. Oladosu, A. AbdulKadir, D.T. Nneji, O.D. Odusola, A.C. Ofem, A. Farate, F.S. Bachi, M.O. Nwabunike, A. Edward, and K.C. Nwaogu. Factors influencing the choice of radiology as a specialty among clinical-year medical students in Nigeria: A multi-center cross-sectional study. *BMC Medical Education*, 25(1), 2025. doi: 10.1186/s12909-025-08067-w.

D. Wallis and I. Buvat. Clever Hans effect found in a widely used brain tumour MRI dataset. *Medical Image Analysis*, 77:102368, 2022. doi: 10.1016/j.media.2022.102368.

K. Yuliawan. Advancements in brain tumor classification: Leveraging MobileNet-V2 and DenseNet121 for high-precision prediction. *Journal of Embedded System Security and Intelligent Systems*, 7(1):92–104, 2026. doi: 10.59562/jessi.v7i1.2608.