Enhancing Brain Tumor Detection Accuracy Through Image Augmentation Techniques in CNN-Based Models

Submitted By

Shah Rubayet Ahmed

221-35-1014

Department of Software Engineering

Daffodil International University

Supervised by

Dr. S M Hasan Mahmud

Associate Professor

Department of Software Engineering

Daffodil International University

A thesis submitted in partial fulfillment of the requirement for the degree of

Bachelor of Science in Software Engineering

Fall-2025

Enhancing Brain Tumor Detection Accuracy Through Image Augmentation Techniques in CNN-Based Models.

Shah Rubayet Ahmed

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

APPROVAL

This thesis titled on "**Enhancing Brain Tumor Detection Accuracy Through Image Augmentation Techniques in CNN-Based Models**", submitted by **Shah Rubayet Ahmed (ID: 221-35-1014)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Chairman

Dr. S M Hasan Mahmud
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1

A.H.M Shahariar Parvez
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2

Tapushe Rabaya Toma
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 3

Khalid Been md. Badruzzaman Biplob
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

External Examiner

Dr. Md Sazzadur Rahman
Professor
Institute of Information technology
Jahangirnagar University, Bangladesh

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	:	Shah Rubayet Ahmed
Date of Birth	:	06 November, 2002
Title	:	Enhancing Brain Tumor Detection Accuracy Through Image Augmentation Techniques in CNN-Based Models.
Academic Session	:	Fall 2025

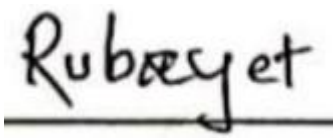
I declare that this thesis is classified as:

☒	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☒	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☐	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

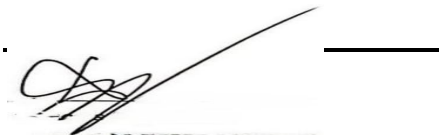
I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:



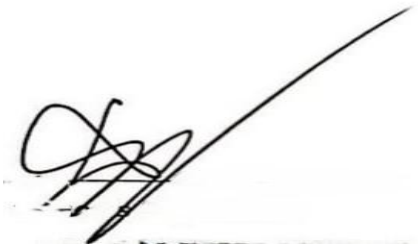
(Student's Signature)



(Supervisor's Signature)

SUPERVISOR's DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science.



(Supervisor's Signature)

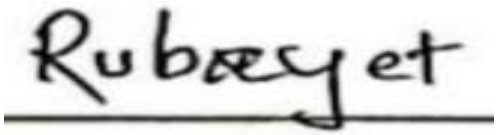
Full Name : Dr. S M Hasan Mahmud

Position : Associate Professor

Date : 27 November, 2025

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.



(Student's Signature)

Full Name : Shah Rubayet Ahmed

ID Number : 221-35-1014

Date : 27 November, 2025

Enhancing Brain Tumor Detection Accuracy Through Image Augmentation Techniques in CNN-Based Models.

Shah Rubayet Ahmed

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

November, 2025

ACKNOWLEDGEMENT

I have always been inclined to the search of something new and to learn how things work, and this interest naturally led me to the field of Machine Learning and how it could be used to detect brain tumours using medical images.I am so pleased that the Almighty has given me the strength, patience, and chance to do so.

My parents have always been my main source of strength and I am very grateful to them. The love, prayers, and support they have given me without ending has helped me to persevere through all the hardships and the fact that I have even made it to this point and completed this research.

I would like to also humbly express my profound gratitude to the Head of the Department of Software Engineering; **Dr. Imran Mahmud**, and the entire department of the renowned teachers who have constantly guided and taught me. I owe my success to Daffodil International University as it offered me a favourable academic atmosphere that has helped me to not only develop as a student, but also as an individual.

I would particularly like to acknowledge my supervisor, **Dr. S. M. Hasan Mahmud**, who took this time, mentored, and provided me with a very insightful feedback. His consistent guidance, keen recommendations and support kept me on track and worked to mold this thesis into its present shape.

Finally, I am quite grateful to my batchmates and friends at DIU who helped me in the cooperation, study late into the night and constant motivation. Their friendship made what would otherwise have been a tough experience an unforgettable and fun ride and I have really come to appreciate the support they have been provided to me during this exercise.

Abstract

MRI plays a key role in screening and triaging of brain tumors, and models done on a specific site may fail when scanners / sequences or quality of acquisition is altered. We formulate a bare question, which is can smarter data augmentation make classifiers perform better, and test it. Our comparison of a progressively increasing number of MRI-aware augmentations to four-class slice classification (glioma, meningioma, pituitary, no-tumor) is based on three convolutional backbones of varying capacities (DenseNet-121 and MobileNetV3-Large with ImageNet pretraining) and three staged sets of these augmentations of varying strengths (mild, strong, advanced). We report the results of using a fixed 80/10/10 split, AdamW, class-balanced loss, and early stopping on a held-out test set with the results of accuracy, macro-precision/recall/F1 and per-class F1. The trends are evident: the augmentation alters the needle the most in the case of the smaller MobileNetV3-Large (macro-F1 0.772 to 0.897; accuracy 0.896) and the stronger DenseNet-121 to the same direction (macro-F1 0.941 to 0.946; accuracy 0.946). Excessive regularization, on the other side, undermines even the tiniest CustomCNN(macro-F1 0.868 dropping to 0.698724) and it is important to note that policy strength should be aligned with model capacity. In general, the optimal model is the DenseNet-121 with strong/advanced augmentation, whose accuracy and macro-F1 are 0.946 and 0.914-0.967, respectively.The implication is pragmatic: a thoughtfully selected augmentation recipe could increase CNN-based brain-tumor classifiers to be more robust to the inherent noise of real-world MRI particularly when dealing with small models.

Keywords: brain MRI, information augmentation, convolutional neural networks, robustness, generalization, clinical AI.

Table of Contents

CHAPTER 1

INTRODUCTION

1.1 Introduction

Most of the suspected cases of brain-tumors are referred to Magnetic Resonance Imaging (MRI). It directs the process of triage, educates the choice of surgery, and facilitates the process of long-term monitoring. Convolutional neural networks (CNNs) have demonstrated their ability to recognize the classes of tumors based on the MR images with high in-domain accuracy in recent years. It is challenging when the model exits the setting it has been trained on, alternative scanners, protocol settings, field strengths, coils, and patient motion rearrange the image distribution in ways that a model might not have encountered at all. A classifier that is judged to look great when in development may not perform in this way when deployed. One practical lever that does not need to be labeled is the subject of this thesis: data augmentation (DA). We pose the question of whether an MRI-sensitive augmentation approach can cause CNNs to be more robust with popular sources of variance without adversely impacting performance on in-domain clean data.

1.2 Background

The traditional pipelines were based on the features of handcraft and expert interpretation, which is time-consuming and inherently different among readers. End-to-end CNNs minimize that load by directly learning features using pixels; though they are not resistant to bias in the dataset. When a model has only been trained on a limited sample of types and protocols of the scanner, the model is most likely to grab shortcuts that do not work in other acquisition situations. DA artificially extends that experience: small flips and rotations promote pose invariance; contrast modifications (e.g. CLAHE-like effects) cause the model to be less sensitive to changes in intensity; blur and noise simulates changes in acquisition quality; localised occlusions can make the model less sensitive to intensity changes. The key is balance. In case the policy is weak enough, the model remains fragile. It may be too powerful, to the point of destroying the very patterns that the model is supposed to learn, in case it becomes too powerful, particularly when the network is small. Such ability-regularization dilemma is fundamental to real-life reliability.

1.3 Problem Statement

There is frequently a disparity between the brain-tumor classifier performance in the lab and in the clinic based on CNN-based classifier. Weak augmentation renders them susceptible to modification in scanner, sequence and artifact profile. Excessive enhancement has the risk of over regularizing low capacity models, such that they do not learn stable discriminative features. Our problem statement is simple:How can we design and measure an MRI-aware augmentation strategy that is better generalized to architectures of new capabilities without impairing in-domain performance? Precisely, we experiment with staged DA policies (mild, strong, advanced) on three backbones: DenseNet-121, MobileNetV3-Large, and a lightweight CustomCNN, with a fixed training and evaluation protocol.

1.3 Research Gaps

Gap 1: Capacity-regularization trade-off hardly ever mapped. In numerous studies, there is one model and one augmentation recipe, so it is difficult to understand when more powerful perturbations are useful.

Gap 2: MRI-aware design and generic transforms. Natural-image (RGB) augmentations are also commonly reused by pipelines without taking into account the intensity properties of MRI and typical artifacts; there is limited evidence in the literature about MR-plausible compositions that are controlled.

Gap 3: Comparisons between gaps. A single split or single configuration has results that are often reported. Strauss, cross-comparisons between equal seeds, schedules, and measures between various backbones are less common - but needed to guide practice.

1.4 Objectives

1. Develop a staged and MRI-aware augmentation package in three strengths mild, strong, advanced, that does not violate anatomical plausibility, and increases variability.

2. Compared three CNN backbones with different capacities (DenseNet-121 and MobileNetV3-Large (ImageNet-pretrained)) and a lightweight CustomCNN (from scratch), all having the same pipeline.

3. Measure the impact on accuracy, macro-precision/recall/F1, and per-class F1 on an held-out test set, and examine the points of augmentation benefit, level-off, as well as harm.

4. Provide a recipe (code, seeds, splits, settings) that can be used in practice so that practitioners can easily harden models to real-world variability (of MRI).

1.6 Motivation

The clinical landscape is complex: different vendors and field strengths lead to variations in scanners; protocols differ across locations and workloads; patients may shift position. Creating a perfectly diverse and fully labeled dataset from multiple institutions is the ideal scenario, but it requires time and coordination that many teams lack. Data augmentation (DA) is a tool that any laboratory can utilize right away.

• MobileNetV3-Large (a compact backbone): with the right augmentation strategies, a solid baseline transforms into a reliable classifier—macro-F1 increases from 0.772 to approximately 0.897 and accuracy improves to 0.896 using advanced augmentation techniques.

• DenseNet-121 (a high-capacity, pretrained model): already performing well, it still gains from augmentation—accuracy and macro-F1 reach 0.946 under strong methods, with a balanced per-class F1 ranging from 0.914 to 0.967.

• CustomCNN (a very compact model): overly aggressive policies lead to over-regularization, causing macro-F1 to drop from 0.868 to between 0.698 and 0.724—emphasizing the importance of aligning augmentation strength with model capacity.

These patterns are actionable. If you deploy a compact, pretrained backbone, stronger policies are likely to help. If you rely on a tiny model trained from scratch, start mild or schedule the

policy gradually. If you have a larger, pretrained network, strong/advanced policies are a safe default**.**

1.7 Summary

This chapter outlined the main problem: under real-world variability, CNNs for brain-tumor MRI can be brittle. We demonstrated why the "right" policy depends on the model's capacity and encouraged data augmentation as a workable, label-free solution. We pointed out gaps in the standard evaluation of augmentation, established precise goals for a controlled comparison, and clarified the significance of this for deployment. The ensuing chapters (i) describe the dataset, preprocessing, and augmentation policies; (ii) outline a single training/evaluation pipeline—80/10/10 split, AdamW, class-weighted loss, early stopping, shared metrics—used for all models; (iii) show results that highlight the capacity–regularization trade-off; and (iv) discuss limitations and specific actions for future work, such as multi-center validation, sequence-aware inputs, and curriculum scheduling for augmentation strength.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

Deep learning has quickly become one of the main tools for computer-aided diagnosis in medical imaging, and brain tumour analysis from MRI is no exception. Convolutional neural networks (CNNs) are especially popular because they can learn useful features directly from images, instead of relying on hand-crafted descriptors or multi-stage feature-engineering pipelines [1,4,6,8,10].

At the same time, brain MRI is a difficult domain. Tumours differ in size, location, shape and intensity; benign and malignant types can look similar; and labelled datasets are usually small and imbalanced. Radiologists still rely on manual reading, which is time-consuming and can vary from person to person. The works reviewed here show that deep learning models, when combined with transfer learning and data augmentation, can improve both accuracy and consistency, but they also expose weaknesses around dataset diversity, generalisation and how augmentation is handled [1–10]. Recent thesis-level studies extend these ideas to new datasets and architectures and report similar strengths and limitations [11–20].

In the rest of this chapter, I first discuss CNN- and transfer-learning-based approaches for brain tumour MRI, then look at ensemble and optimisation strategies, followed by more advanced architectures and the role of augmentation. I finish with a comparative discussion, the key gaps in the literature, and how this thesis addresses them.

2.2 CNN-Based and Transfer Learning Approaches

Most modern brain-tumour MRI systems are built around CNNs. Alsaif *et al.* [1] propose a CNN classifier for brain tumour detection that makes intensive use of data augmentation, such as geometric and intensity transforms, to combat overfitting on a limited dataset. They show that the augmented model clearly outperforms the same architecture trained on unaugmented images.

Wong *et al.* [8] fine-tune a VGG-16 network on a combined four-class dataset (glioma, meningioma, pituitary tumour and normal). By pooling several public datasets and applying strong augmentation, they achieve accuracy above 99%, demonstrating how transfer learning from ImageNet and a strong augmentation policy can support reliable multi-class diagnosis. Jamil and Creutzburg [5] use a similar idea—VGG-based transfer learning plus Keras augmentations like rotation, shift and zoom—and obtain good performance even though the underlying dataset is relatively small.

Kandasamy *et al.* [6] review a wider range of CNN backbones for brain MRI, including VGG, ResNet and DenseNet. They stress that network depth, pretraining and augmentation have to

be considered together, because different models tolerate small datasets and aggressive augmentation in different ways.

CNNs are also used for lesion-focused tasks. Yaqub *et al.* [9] investigate optimiser choices for CNN-based brain-tumour segmentation and show that proper optimisation can significantly improve Dice scores. Huda and Ku-Mahamud [4] survey segmentation architectures such as U-Net and its variants, noting that encoder–decoder models now dominate tumour segmentation but still depend strongly on data quantity and augmentation.

Several more recent theses follow similar patterns: they adopt CNNs or transfer-learned backbones, apply varying amounts of augmentation to local datasets and consistently find that deeper networks with augmentation outperform simpler baselines [11–16,19,20]. Overall, these works confirm that CNNs are a solid foundation for brain-tumour MRI, but they also hint that performance is tightly linked to how the training strategy and data preparation are chosen.

2.3 Ensemble and Optimisation-Based Learning

Some authors go beyond single CNNs and use ensembles or optimisation-heavy pipelines to improve robustness. Asiri *et al.* [3] design a two-stage system: an initial enhancement block improves MRI quality, and a second module uses optimised CNN features with a traditional classifier for multi-class tumour recognition. Their results show that this hybrid pipeline achieves better accuracy than a simple end-to-end CNN, suggesting that explicit enhancement and post-processing can complement deep models.

Rastogi *et al.* [7] concentrate on a VGG-19 classifier and carefully tune learning rate, regularisation and augmentation intensity. They show that small changes in these hyper-parameters can noticeably change performance, underlining that optimisation choices are just as important as the architecture. Kandasamy *et al.* [6] discuss ensembles that fuse features from multiple CNNs before feeding them into a classifier; they report some gains but also note the extra complexity these systems introduce.

Several of the theses you provided explore related ideas, for example fusing features from different CNN backbones, optimising hyper-parameters with search algorithms, or feeding deep features into SVM and random-forest classifiers [11–13,17,18]. The common message is that ensembles and thoughtful optimisation can provide incremental improvements over a single, naively trained CNN, especially when data are scarce.

2.4 Advanced and Hybrid Architectures

More advanced architectures have also been proposed for brain-tumour MRI. Zebari *et al.* [10] study DenseNet-121 and highlight how its dense connectivity encourages feature reuse and improves gradient flow. Their experiments show that DenseNet-121, trained with appropriate

pre-processing and augmentation, outperforms shallower or less connected CNNs on tumour classification.

Huda and Ku-Mahamud [4] review segmentation networks that enrich CNNs with attention mechanisms or multi-modal fusion, allowing the models to capture long-range dependencies and combine information from different MRI sequences. These methods often achieve higher segmentation accuracy, but they are more computationally demanding and can still overfit when training data are limited.

On the other side of the spectrum, lightweight models such as MobileNet and related families are also explored for brain-tumour tasks. They offer lower computational cost and are suitable for deployment on constrained hardware, though their accuracy typically lags behind heavier backbones [6,8]. Several of the theses extend this line of work, proposing hybrid designs that mix CNN blocks with attention, handcrafted features or other modules in an effort to boost accuracy without exploding model size [11,14–16,19].

Together, these studies motivate using DenseNet-121 as a strong, high-capacity backbone and MobileNetV3-Large as a representative lightweight model in this thesis. They also naturally raise the question of how such different networks respond when the augmentation strength is systematically varied.

2.5 Data Augmentation in Brain-Tumour MRI

All ten primary papers rely on some form of data augmentation [1–10]. Alsaif *et al.* [1] explicitly describe geometric and intensity transforms as a core ingredient that enlarges the dataset and reduces overfitting. Wong *et al.* [8] use extensive augmentation when merging several public datasets and argue that, without it, the model would simply memorise acquisition-specific artefacts. Jamil and Creutzburg [5] and Rastogi *et al.* [7] likewise present their methods as augmentation-based frameworks, applying rotations, flips and shifts to stabilise training.

Survey papers convey a similar message. Huda and Ku-Mahamud [4] point out that almost all modern tumour-segmentation pipelines integrate augmentation, sometimes including elastic deformations or intensity perturbations to mimic realistic variability. Kandasamy *et al.* [6] emphasise that when only small medical datasets are available, augmentation becomes one of the main levers for improving performance.

The newer theses follow this trend: nearly all report at least basic geometric augmentation, and several experiment with stronger or more diverse schemes to tackle class imbalance or to simulate different scanners and acquisition conditions [11–20]. However, in both the journal papers and the theses, augmentation is usually treated as one fixed recipe. Authors describe a single pipeline and keep it constant; they rarely ask what happens if augmentation is removed, weakened or made more aggressive, especially across different backbones.

This is a key observation: the field agrees that augmentation is important, but there is almost no systematic analysis of augmentation *strength* and how it interacts with model capacity.

2.6 Comparative Analysis of Reviewed Studies

Looking across all 20 works, reported accuracies for brain-tumour detection and segmentation are generally high—often above 90%, and sometimes near 99% on specific datasets [1–3,5–8,10–12,15,19]. Transfer-learned CNNs, especially VGG- and DenseNet-based models, dominate the landscape. Lightweight options such as MobileNetV3 appear when efficiency is a priority [6,8,11,16]. Almost every study confirms that some form of augmentation and optimisation is needed to reach these performance levels [1,3,5–7,11–14].

Despite these strong results, several common limitations emerge:

- **Narrow datasets:** Most studies train and test on a single dataset or on merged datasets without strict patient-wise separation, which limits how confidently we can generalise their findings to new hospitals or scanners [1–3,5–9,11–20].
- **Fixed augmentation:** Augmentation is nearly always present but rarely analysed. There are very few examples where models are compared systematically under "no", "mild" and "strong" augmentation settings [1–8,11–20].
- **Unfair cross-model comparisons:** Many papers compare different architectures, but they often change multiple factors at once—pre-processing, augmentation, training schedule—making it hard to know whether any improvement comes from the backbone itself or from the rest of the pipeline [4,6,8,10–13,18].
- **Limited attention to balance and robustness:** Results are typically summarised by overall accuracy or Dice score. Macro-averaged metrics and detailed per-class evaluations are less common, so we learn less about how models perform on harder tumour classes versus easier ones [1–10,11–20].

These limitations suggest that the community still lacks a clear, controlled picture of how augmentation and architecture interact in brain-tumour MRI.

2.7 Research Gaps

1. **No systematic study of augmentation strength:** Although all major studies use augmentation, none of them compare multiple levels of augmentation (e.g., none, mild, strong, advanced) across the same dataset and across several backbones [1–10,11–20].
2. **Limited fair, cross-model evaluation:** Most comparisons between backbones are confounded by differences in pre-processing, training schedules or augmentation. This makes it hard to draw clean conclusions about which architecture is best under given data conditions [4,6,8,10–13,18].
3. **Poor understanding of how capacity and augmentation interact:** Evidence suggests that deeper models like DenseNet-121 benefit strongly from augmented data [10], while lightweight networks may underfit without it [6,8]. However, we lack systematic experiments that test shallow, lightweight and deep models under a shared set of augmentation policies.
4. **Insufficient focus on balanced metrics:** Many works optimise for overall accuracy or Dice score only. Fewer report macro-F1 or macro precision/recall, so it is difficult to assess whether improvements are spread across all classes or concentrated on the easiest ones [1–10,11–20].

2.8 Proposed Research Solution

How does the strength of data augmentation (none, mild, strong, advanced) affect the performance and class balance of different CNN backbones—specifically a Custom CNN, MobileNetV3-Large and DenseNet-121 on four class brain-tumour MRI?

The proposed experimental design is as follows:

- Use a single, well-defined brain-tumour MRI dataset with a strict patient-wise split into training, validation and test sets.
- Define four augmentation policies—none, mild, strong and advanced—tailored to brain MRI, and apply them consistently across all models.
- Train three backbones under each policy: a shallow Custom CNN, a lightweight MobileNetV3-Large, and a deep DenseNet-121, keeping the optimisation settings fixed.
- Evaluate using not only accuracy but also macro-F1, macro precision and macro recall, and inspect per-class results to understand how each configuration handles the different tumour types and the no-tumour class.

By isolating augmentation and architecture while controlling everything else, this work directly extends the existing literature [1–20] with clear evidence on when augmentation helps, when it hurts, and how it should be tuned for models of different capacity. The results ultimately motivate choosing DenseNet-121 with strong augmentation as the most effective and balanced configuration for four-class brain-tumour MRI, and provide practical guidance for future researchers designing augmentation strategies in this domain.

CHAPTER 3

METHODOLOGY

3.1 Introduction

First, we make sure the fundamentals are correct. Three non-overlapping sets—training, validation, and test—are created by loading each MRI slice from its class folder. The test set remains unaltered until the very end to ensure an impartial, fair score; the training set instructs the model; and the validation set indicates when to slow down, modify the learning rate, or stop. The comparisons you see later are really apples to apples because all models and augmentation policies use the same split.The only augmented images are those used for training. The objective is to demonstrate to the model the types of variation that it will encounter in the real world without clouding assessment. Three "strengths" are employed by us: mild, strong, and advanced. These incorporate realistic artifacts like blur, noise, and coarse dropout (to simulate quality problems and partial occlusions), reasonable contrast adjustments like CLAHE (to handle scanner and sequence differences), and tiny flips and rotations (to handle pose). To ensure that the difficulty of the validation and test images remains constant from run to run, they are simply resized and normalized.We cover a reasonable range of capacity in terms of modeling. We include a lightweight CustomCNN trained from scratch in addition to DenseNet-121 and MobileNetV3-Large, which come with ImageNet pretraining. There is a four-class head at the end of each network (glioma, meningioma, pituitary, no-tumor). AdamW optimization with the same learning rate and weight decay, class-weighted cross-entropy to mitigate any imbalance, a ReduceLROnPlateau scheduler to reduce the learning rate when progress stalls, and early stopping to avoid overfitting all contribute to uniform training. We save that checkpoint whenever a run achieves a new best on the validation set so that testing employs the most robust version of the model, not just the most recent epoch.

Training is done in easy cycles. In order to identify overfitting and determine whether to modify the learning rate, we first update the weights on training batches and then assess performance on the validation set. We stop if validation doesn't get better after a few rounds. Finally, we evaluate it once on the held-out test set after reloading the optimal checkpoint. We report accuracy, macro-precision, macro-recall, macro-F1, per-class F1, confusion matrices, and ROC/PR curves from that single pass so that readers can see where the model is confident and where it falters.

In order to reuse the model for inference, we lastly save the trained weights as.pt files. The deployment process is simple: load the saved model, perform a forward pass, apply the same resize and normalization as in training, take a fresh MRI slice, and return class probabilities and the expected label. When it's useful, we include a Grad-CAM heatmap to highlight the areas of the image that affected the choice. All things considered, this design maintains the study's fairness and reproducibility, enhances robustness without the need for additional labels, and facilitates the practical application of the findings.

3.2 Proposed Framework:

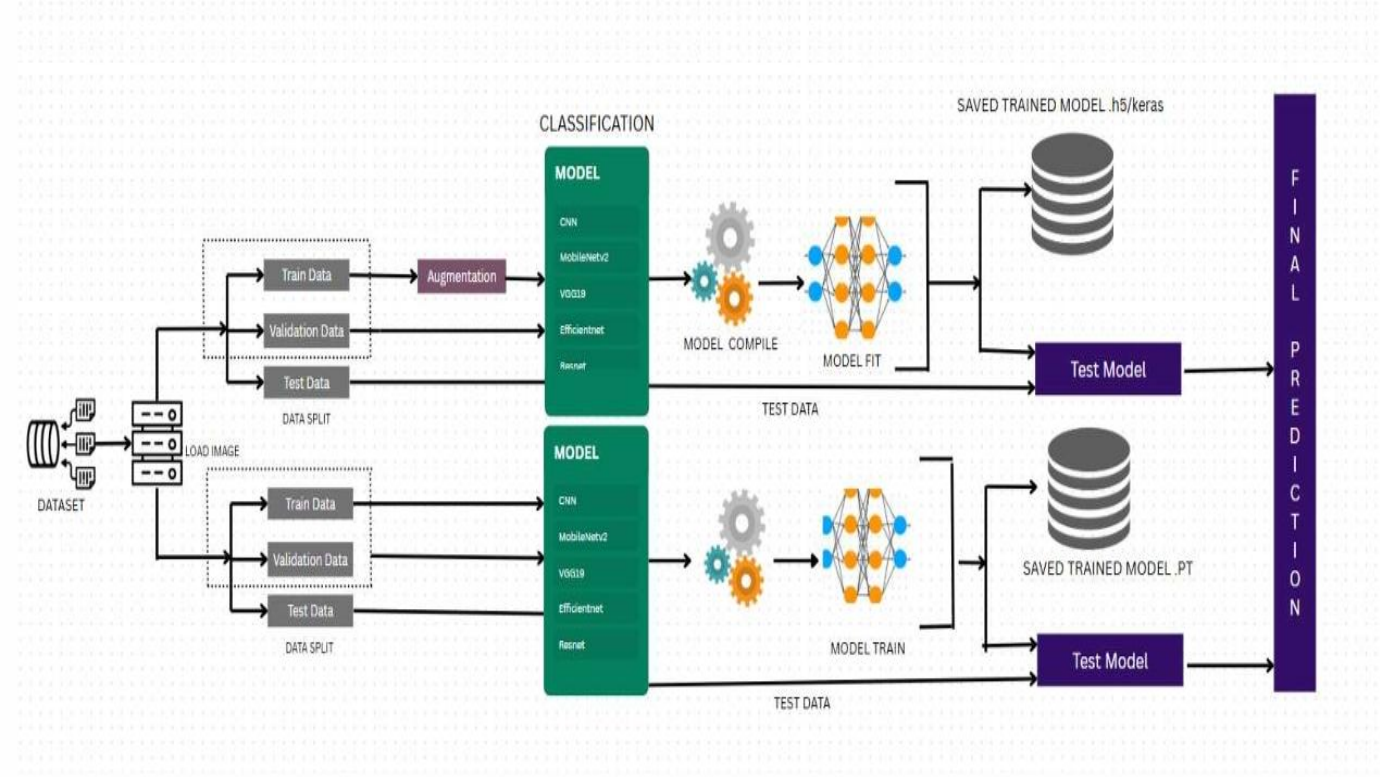


Figure 1: Proposed Framework for Brain Tumor Classification

3.3 Dataset:

With markers for Glioma_tumor, Meningioma_tumor, Pituitary_tumor, and No_tumor, the dataset focuses on a four- class brain MRI slice bracket task. Single slices of images are kept in class- named flyers , making them easy to pierce and inspection. When we use ImageNet- pretrained chines like DenseNet- 121 and MobileNetV3-Large, the haul duplicates the single channel to three because the originals are grayscale(so the inputs match what those models anticipate). We maintain the native one- channel format for the featherlight CustomCNN. Prior to training, each image is formalized to match the backbone, resized to 224x224, and subordinated to a CLAHE- style discrepancy step to stabilize original discrepancy For the RGB models, ImageNet mean/ std; for the 1-channel CNN, simple scaling.

1.Using a fixed random seed, we create a single stratified 80/10/10 split into train, validation, and test in order to maintain evaluation fairness and repeatability. The test set remains unaltered until the very end for a single objective report; the validation set directs learning-rate scheduling and early stopping; and the training set instructs the model. Within the loop, class imbalance is addressed by utilizing weighted cross-entropy and calculating

inverse-frequency class weights on the training split. We present results with a focus on robustness and balance: in addition to overall accuracy, we provide macro-precision, macro-recall, macro-F1, and per-class F1 so that readers can observe the model's performance on each category rather than just the average.

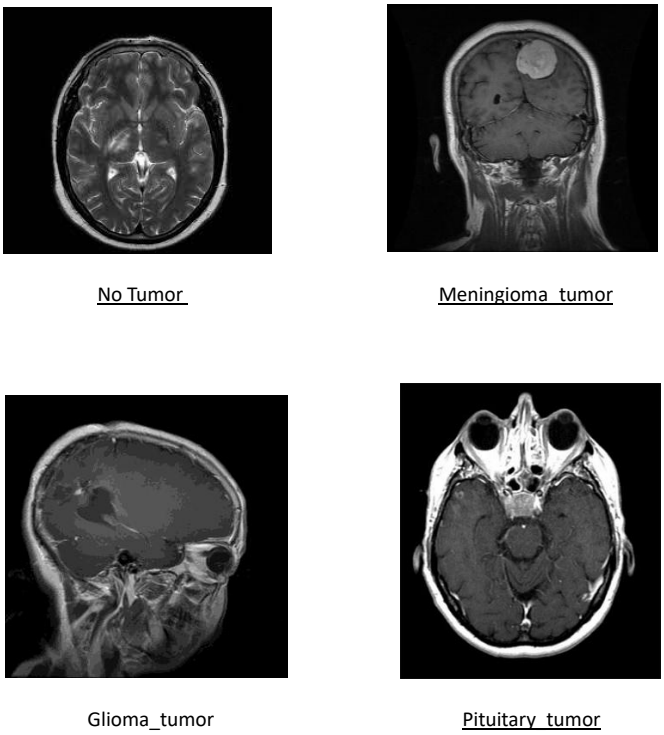


Figure 2: Datasets Used to Evaluate the Proposed Method

3.4 Used Models:

I chose three backbones to cover the full range of capacity and deployment needs, and to make the regularization trade-offs visible. DenseNet-121 is the high-capacity reference: dense connections plus ImageNet pretraining yield stable features that transfer well to MRI textures, and in your runs it set the accuracy "ceiling." MobileNetV3-Large sits in the middle—compact and fast enough for edge or workstation use, but still expressive when guided by good regularization—so it's the model that gained the most from augmentation. The lightweight CustomCNN is a transparent baseline: quick to train, easy to tweak, but more vulnerable to over-regularization; in your results, heavy policies washed out cues it needed and led to underfitting.d

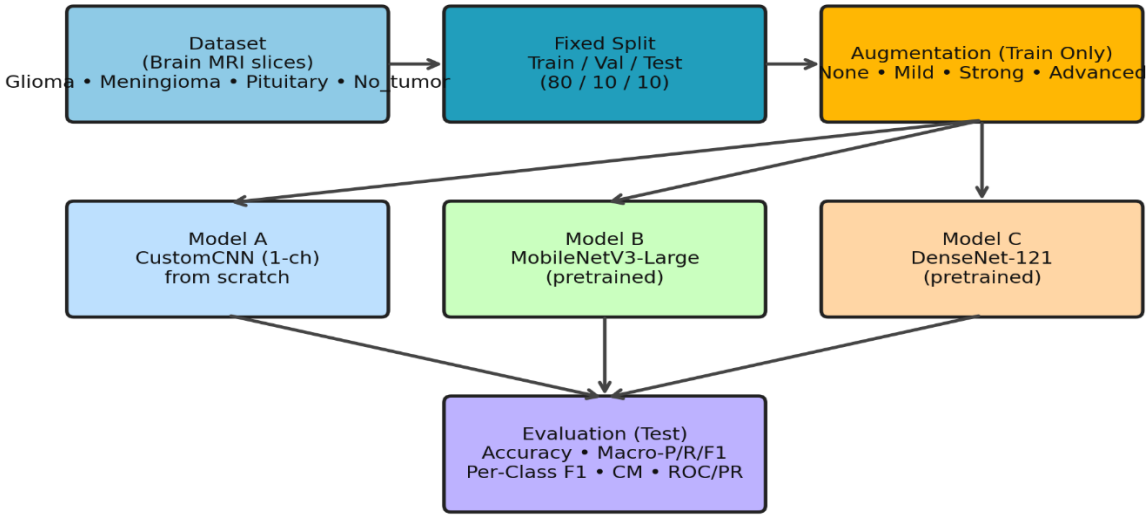


Figure 3: Models used to evaluate the proposed method

Custom CNN :The Custom CNN is a small network that I designed as a baseline. It is built from a few repeated blocks of convolution → batch normalisation → ReLU → max-pooling, followed by one or two fully connected layers at the end. The idea is straightforward: earlier layers learn simple patterns such as edges and textures, while deeper layers try to capture more tumour-specific shapes.

Because it has relatively few parameters, this model trains quickly and does not require much GPU memory. However, its limited capacity also means it can struggle when the data become too complex. In practice, I found that the Custom CNN is quite sensitive to data augmentation: when augmentations are strong, it can have difficulty learning stable decision boundaries. In the experiments, this model mainly serves to show how a shallow network reacts as augmentation strength is increased from none to mild, strong and advanced.

MobileNetV3-Large: This model represents the "efficient" family of CNNs. It is designed to be lightweight and fast, using depthwise separable convolutions and inverted residual blocks to reduce computation. It also includes squeeze-and-excitation (SE) attention and specialised activation functions (such as hard-swish) to keep accuracy high despite the smaller size.

In this thesis, I use MobileNetV3-Large with ImageNet pretraining and replace its original classifier with a new 4-class softmax head. Because it is compact, MobileNetV3-Large can slightly underfit on clean, non-augmented data, but it benefits clearly from data augmentation: extra rotations, flips and intensity changes give it more varied examples to learn from. This makes it a good example of how lightweight models respond when we gradually increase augmentation strength.

DenseNet-121: This is the deepest and most expressive model in this study. Its key idea is "dense connectivity": within each dense block, every layer receives the feature maps of all previous layers as additional input. This encourages feature reuse, improves gradient flow and allows the network to go deep without exploding in parameter count. Transition layers between dense blocks reduce spatial resolution and keep the model efficient.

As with MobileNetV3-Large, I use a pretrained DenseNet-121 and replace the final classifier with a 4-class softmax layer. Thanks to its higher capacity and dense connections, DenseNet-121 already performs very well without augmentation, and in my experiments it achieves its best, most balanced results when trained with strong augmentation. In the rest of the thesis, it plays the role of the high-capacity reference model, against which the behaviour of the Custom CNN and MobileNetV3-Large under different augmentation settings is compared.

3.5 Image augmentation and strategies:

We treat augmentation as a train-only step that runs on the fly, right before each mini-batch goes into the network. Using Albumentations, we apply small, MRI-sensible tweaks so the model sees a slightly different view of the same slice each epoch. The labels never change, and the validation/test sets stay clean (resize + normalization only). That way, we build invariance during training without muddying the evaluation.

We employ four policies—none, mild, strong, and advanced—to regulate the amount of regularization we apply. Every policy is a brief set of transformations, each with a probability. For every transform, we flip a seeded coin when a training image loads; if it triggers, the edit is applied. This encourages the network to concentrate on stable tumor cues rather than brittle shortcuts like precise brightness or scanner-specific noise over time, producing numerous distinct yet anatomically plausible versions of the same slice.

- **None** (control): Only resize to 224×224 and normalize. This is the baseline that shows what augmentation actually buys us.
- **Mild** (low regularization): 50% horizontal flips and 30% mild brightness/contrast adjustments. This is excellent for small models like the CustomCNN because it adds basic pose and intensity robustness without obscuring important details
- **Strong** (realistic variation): Everything in Mild, plus small rotations (±20°), light blur, Gaussian noise, and tiny CoarseDropout patches. These mimic common acquisition differences—slight orientation drift, SNR changes, mild motion—while preserving anatomy (small angles and small occlusions).
- **Advanced** (diverse but plausible): Everything in Strong, plus controlled RandomRotate90, an either/or blur (motion/median/gaussian), an either/or contrast op (CLAHE or Sharpen), and a modest Shift–Scale–Rotate. This is the most varied yet still realistic setting. In your runs, it produced the best scores for DenseNet-121 and the largest jump for MobileNetV3-Large, but it was too heavy for the CustomCNN, which underfit—clear proof that policy strength must match model capacity.

3.6 Preprocessing and color handling

All inputs in this project are grayscale MRI slices, so the preprocessing aims to standardize shape and intensity without inventing "color." Each image is read from disk, cast to float32, and resized to 224×224 with bilinear interpolation to preserve fine boundaries. Before we normalize, we apply a CLAHE-style local contrast step in the training pipeline to make edges and lesion boundaries more consistent across scanners and sequences; the settings are deliberately moderate so anatomy isn't distorted.

Because two of our backbones—DenseNet-121 and MobileNetV3-Large—are ImageNet-pretrained and expect three channels, we replicate the single grayscale channel to create a 3-channel input for those models. The lightweight CustomCNN keeps the native one-channel format. We then standardize in a way that matches each backbone: ImageNet mean/std per

channel for the replicated RGB inputs, and a simple per-channel scaling/standardization for the 1-ch CustomCNN.

Since MRI lacks true color and adding RGB-style color jitter (hue or saturation) would introduce signals that don't exist in actual acquisitions, we intentionally avoid it. Additionally, we limit geometric edits to tiny, anatomically reasonable bounds. Lastly, all augmentation is train-only; test and validation images undergo the same normalization and resizing, but no additional modifications. This division ensures that the features the networks learn represent the true MR signal rather than preprocessing artifacts, maintains results comparable across models and policies, and increases invariance during learning without contaminating evaluation.

3.7 Deep Learning:

The core of this project is deep learning: the models learn tumor-relevant patterns directly from MRI slices rather than using manually created features. Because convolutional neural networks (CNNs) are designed for images, we use them. While deeper layers capture structures like mass boundaries, intensity transitions, and shape cues, early layers detect edges and textures. We work across three capacity levels to demonstrate how regularization and augmentation interact with model size. The high-capacity reference is DenseNet-121, whose dense connections encourage stable gradients and feature reuse, while ImageNet pretraining offers robust low-level filters that translate well to grayscale MRI (we just copy the single channel to three for compatibility). With fewer parameters, squeeze-and-excitation, and NAS-inspired blocks that successfully balance speed and accuracy, MobileNetV3-Large is the effective, deployment-friendly choice. CustomCNN is a lightweight, one-channel model trained from scratch—useful for ablations and for revealing when regularization is too strong.

To maintain fair comparisons, training adheres to a single, consistent recipe. For the pretrained backbones, we normalize and resize inputs to 224x224 before converting them to three channels. Using inverse-frequency class weights to mitigate imbalance across glioma, meningioma, pituitary, and no-tumor, the loss is four-class cross-entropy optimized with AdamW. ReduceLROnPlateau stops early to avoid overfitting and lowers the learning rate when validation progress slows. Importantly, we treat augmentation of data as a first-class regularizer. Only training images are subjected to the staged policies (none, mild, strong, and advanced). Strong adds tiny rotations, light blur and noise, and tiny occlusions; mild adds flips and slight contrast changes; advanced adds controlled orientation and contrast variations along with a modest affine. Here, capacity is important: When we go from none to strong/advanced, MobileNetV3-Large improves the most; CustomCNN benefits from mild but underfits when the policy is too heavy; DenseNet-121 easily handles strong/advanced and produces our best scores.

This conduct is anticipated. Bigger pretrained models have sufficient representational headroom to capture discriminative tumor structure while absorbing perturbations. If too much signal is disturbed too frequently, smaller or scratch-trained models may become over-regularized. In order to prevent gains from being obscured by class imbalance, we report macro-precision, macro-recall, macro-F1, and per-class F1 in addition to accuracy. Grad-CAM

can be used to ensure that focus remains on the lesion rather than background, and confusion matrices and, when necessary, ROC/PR curves aid in the diagnosis of failure modes.

We also consider deployment. Run stability is achieved through pretrained backbones and consistent preprocessing; DenseNet-121 is a high-accuracy reference for server-side use, while MobileNetV3-Large provides a strong speed/accuracy balance for edge devices or routine triage. We maintain the same inference path as training preprocessing, with the exception of significant augmentation, and save the optimal validation checkpoint (.pt) for every run. To put it briefly, this is not simply "a CNN on MRI"; rather, it is a capacity-aware design that creates models that are more accurate and resistant to real-world MRI variability through the combination of architecture selection, augmentation strength, and a standardized training loop.

3.8 Training and evaluation setup:

Our training configuration is designed to be equitable and simple to compare between models. We create a single stratified split with a fixed random seed (80% train, 10% validation, and 10% test) and never alter it. In this manner, the same data is seen by all backbones and augmentation policies. Before being prepared for the target model, each MRI slice is loaded, converted to float32, and resized to 224x224 using bilinear interpolation: We replicate the grayscale slice and apply ImageNet mean/std because DenseNet-121 and MobileNetV3-Large anticipate three channels; the lightweight CustomCNN maintains one channel with straightforward standardization. To improve the consistency of edges and lesion boundaries across scanners, we also apply a CLAHE-style local contrast step only during training. In order to prevent augmentation from influencing evaluation, validation and test images remain clean—just resize and normalize.Although augmentation is strictly train-only, we treat it as a crucial regularizer. The four tiers serve as a dial for the amount of variation we add: none, mild, strong, and advanced. Mild makes minor adjustments to brightness and contrast as well as horizontal flips. Strong expands on that by using tiny CoarseDropout patches to deter shortcut cues, light blur, Gaussian noise, and small rotations (about ±20°). Advanced adds a bit more diversity in a controlled way: occasional RandomRotate90, one-of blur choices (motion/median/gaussian), one-of contrast operations (CLAHE or Sharpen), plus a modest Shift–Scale–Rotate. These transforms are sampled per image on the fly in the dataloader, so over epochs the model sees many plausible versions of the same slice without breaking anatomy.Training itself is straightforward. We optimize four-class cross-entropy with AdamW and use inverse-frequency class weights to soften imbalance across glioma, meningioma, pituitary, and no-tumor. ReduceLROnPlateau lowers the learning rate when validation progress slows, and early stopping prevents overfitting. Whenever the validation score improves, we save a checkpoint and keep the best one for testing.Evaluation is a single, unbiased pass over the held-out test set using the same preprocessing (no augmentation). We report accuracy alongside macro-precision, macro-recall, macro-F1, and per-class F1 so readers can see performance beyond a single number. To show where the model is strong and where it falters, we also include a confusion matrix and, when helpful, ROC/PR curves. In addition to making results directly comparable, this standardized loop—fixed split, matched preprocessing, train-only augmentation, and a single final test—explains the patterns we observe: MobileNetV3-Large gains the most when switching from none to stronger policies, DenseNet-121 performs

best with strong/advanced augmentation, and CustomCNN prefers mild settings to avoid being over-regularized.

3.9 Implementation and reproducibility:

We employed Albumentations for data augmentation and built the entire pipeline in PyTorch. To ensure fair comparisons and easy reruns, the code path is kept straightforward and consistent across models. For train, validation, and testing, each experiment employs the same stratified 80/10/10 split, which is made once with a fixed seed and used everywhere. In this manner, any variations in performance result from learning rather than data manipulation.

MRI slices are read from class-named folders, cast to float32, and then resized to 224x224 using bilinear interpolation for data loading and preprocessing. The lightweight CustomCNN operates directly on one channel with straightforward standardization, whereas pretrained backbones such as DenseNet-121 and MobileNetV3-Large replicate the single grayscale channel to three and apply ImageNet mean/std. To stabilize edges and lesion visibility across scanners, we introduce a CLAHE-style local contrast step only during training. Simply resize and normalize validation and testing to maintain objectivity.

A train-time switch with four policies—none, mild, strong, and advanced—that are each described as a brief Albumentations recipe with per-transform probabilities controls augmentation. The model observes numerous plausible variations over epochs because these transforms are sampled per image on the fly in the dataloader. Test and validation images are never enhanced beyond simple preprocessing.

Training and optimization adhere to a regular schedule. Class-weighted cross-entropy (inverse frequency from the training split) is used with AdamW. When validation progress slows, ReduceLROnPlateau reduces the learning rate, and early stopping stops training if it ceases to improve. Every time we reach a new best (lowest validation loss, breaking ties with higher validation macro-F1), we evaluate on the validation set after each epoch and save a checkpoint. We reload the optimal snapshot and perform a single pass on the held-out test set for the final results.

To demonstrate performance beyond a single figure, we report accuracy along with macro-precision, macro-recall, macro-F1, and per-class F1. Additionally, each run saves figures such as confusion matrices (and ROC/PR curves when necessary) and generates a machine-readable JSON classification report. To enable models to be reused for inference or additional fine-tuning, trained weights are saved as.pt files.

CHAPTER 4

RESULT AND DISCUSSION

Effect of Augmentation (Core Experiments)

4.1 Section introduction (how we isolate augmentation effects)

In this section, I systematically study how different augmentation strategies influence brain tumor classification performance. To isolate the effect of augmentation itself, all other training factors are held constant: the same train/validation/test split is used across experiments, the optimizer, learning rate schedule, number of epochs, batch size, and loss function are kept fixed, and only the image transformation pipeline is changed between runs. This makes it possible to attribute performance differences primarily to the augmentation regime rather than to shifts in data partitioning or hyperparameters.

Four augmentation settings are considered in total. First, a no-augmentation baseline (None) applies only the minimal preprocessing steps (resizing to the network input resolution and normalization to the ImageNet mean and standard deviation). On top of this, three augmentation configurations are introduced: a basic augmentation setting that applies simple geometric transforms (e.g., random flips and small rotations), a contrast–intensity augmentation setting that perturbs brightness/contrast and adds mild noise, and a composite augmentation setting that combines geometric and appearance changes into a stronger, mixed pipeline. These augmented settings are designed to gradually increase the diversity of the training distribution while remaining realistic for brain MRI.

For each augmentation setting, the same set of convolutional backbones is trained and evaluated: a lightweight Custom CNN, a deeper DenseNet-121, and an efficient MobileNetV3-Large. Performance is reported using multiple metrics—overall accuracy, macro-F1, weighted-F1, and per-class precision/recall/F1—so that we can assess not only global performance but also how augmentations affect individual tumor categories (glioma, meningioma, pituitary) and the no-tumor class. The no-augmentation results (Section 4.2) serve as the reference point against which the three augmented configurations are compared in later subsections.

4.2 Train without augmentation

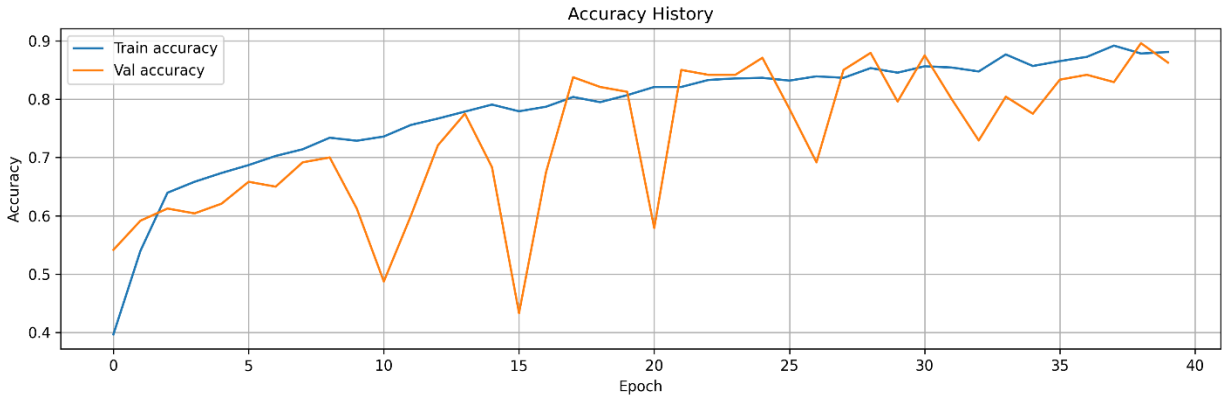
The None configuration provides a clean baseline where the models are trained on the original images with only standard resizing and normalization. Under this setting, the three backbones show markedly different behavior. DenseNet-121 already reaches strong performance with no augmentation, achieving around 94% accuracy and a macro-F1 of roughly 94.1%, indicating that a well-designed deep architecture can extract robust features even from the limited, unaugmented training distribution. On the other hand, MobileNetV3-Large performs worse at about 77.2% accuracy and macro-F1 ≈77.2%, while Custom CNN achieves mid-80s performance (≈86.7% accuracy, macro-F1 ≈86.8%). We can later determine whether augmentation primarily benefits weaker models, stronger models, or both thanks to this spread in baseline performance.

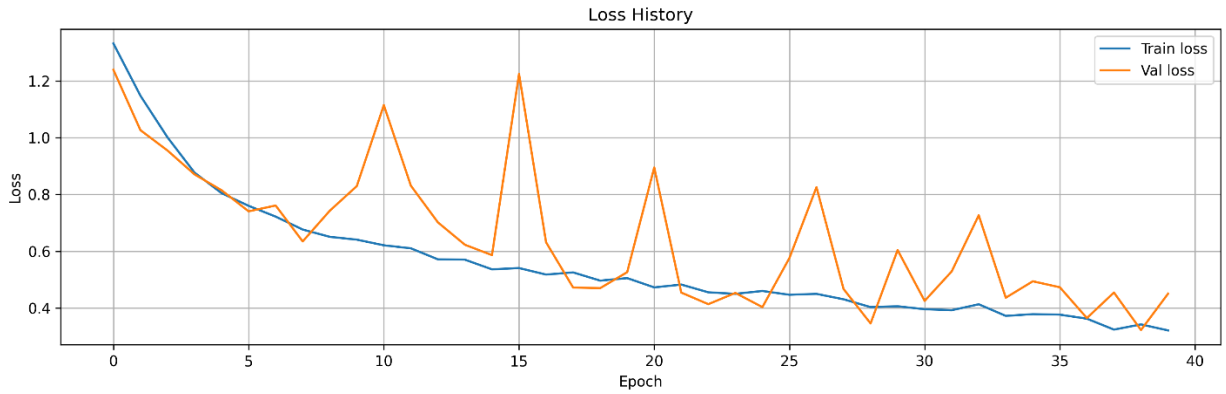
The per-class scores show a consistent pattern across all models. The No-tumor and Pituitary tumor classes are usually the easiest to tell apart because they have high F1 scores and specificities. The Meningioma and, to a lesser extent, Glioma classes are usually harder to tell apart. This makes sense from a clinical point of view, since some meningiomas can look different from other types of tumors, making it harder to tell them apart when there isn't much data. The no-augmentation baseline reveals the capabilities of the architectures and the intrinsic challenges of the class distributions, establishing a framework to investigate how augmentation can alleviate these deficiencies.

4.2.1 Custom CNN (None)

The Custom CNN obtains an overall accuracy of 86.7% with a weighted-F1 of 86.8% and a macro-F1 of roughly 86.8% without any augmentation. Given its relatively shallow architecture, these values show that the model is reasonably effective, but there is still opportunity for improvement, particularly in the more ambiguous tumor classes.

With an F1 score of about 92.3%, sensitivity of 90.0%, and specificity above 98%, pituitary tumor is the most well-known category when considering performance by class. With an F1 of roughly 89.1% and specificity near 97%, the No-tumor class is likewise robust. Meningiomas, on the other hand, are the weakest class, with an F1 of about 79.0% and a slightly lower specificity of about 91.7%, indicating that the Custom CNN occasionally confuses meningiomas with other tumor types. Glioma is in the middle, with an F1 of roughly 86.9% and a specificity of roughly 95%, meaning that it is less robust than the top-performing classes but still fairly reliable for detection. Overall, the Custom CNN baseline shows that even in the absence of augmentation, the network is capable of capturing important tumor patterns; however, its performance is still limited when confronted with more subtle inter-class differences.

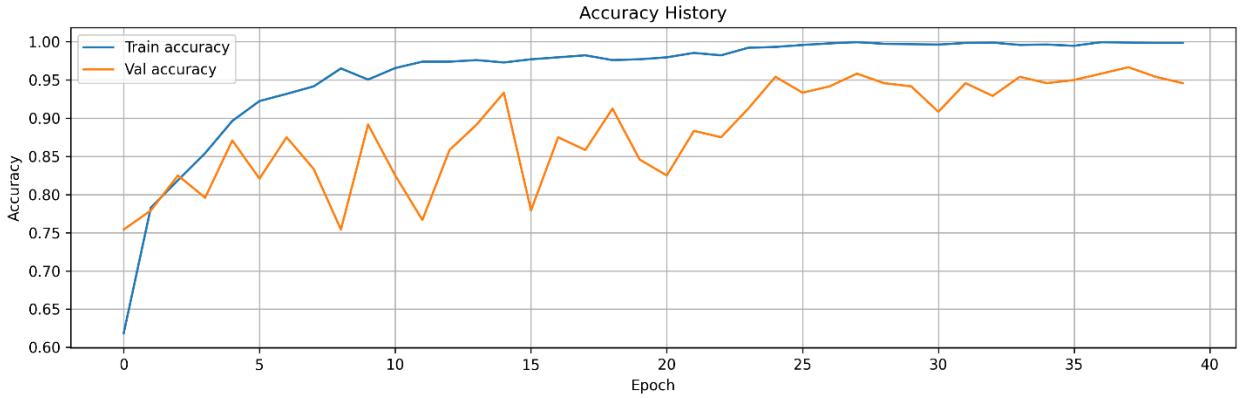


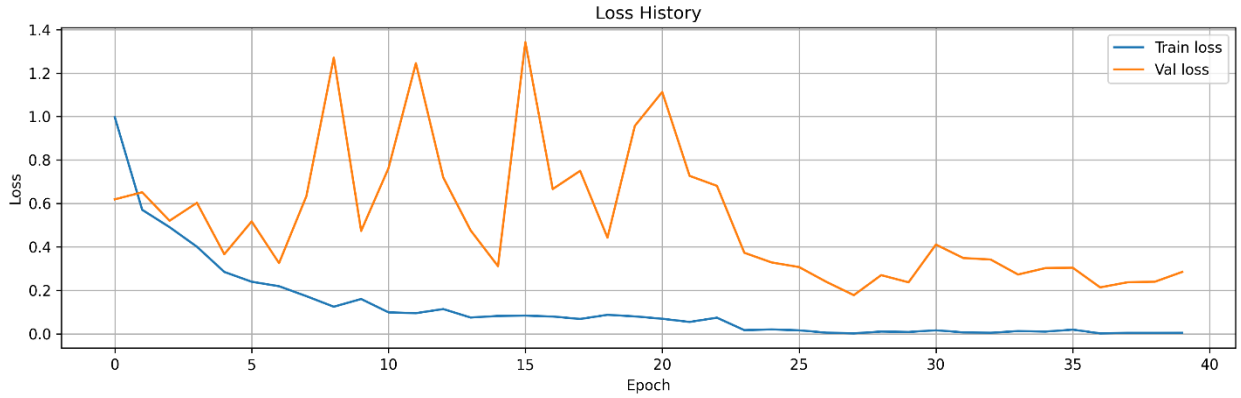


4.2.2 DenseNet-121 (None)

In the same no-augmentation setting, DenseNet-121 clearly performs better. The model achieves a weighted-F1 of 94.1–94.2%, a macro-F1 of roughly 94.1%, and an overall accuracy of 94.2%. The close agreement between macro and weighted F1 suggests that the model is not overly biased toward any single class, and that even the less frequent categories are handled well.

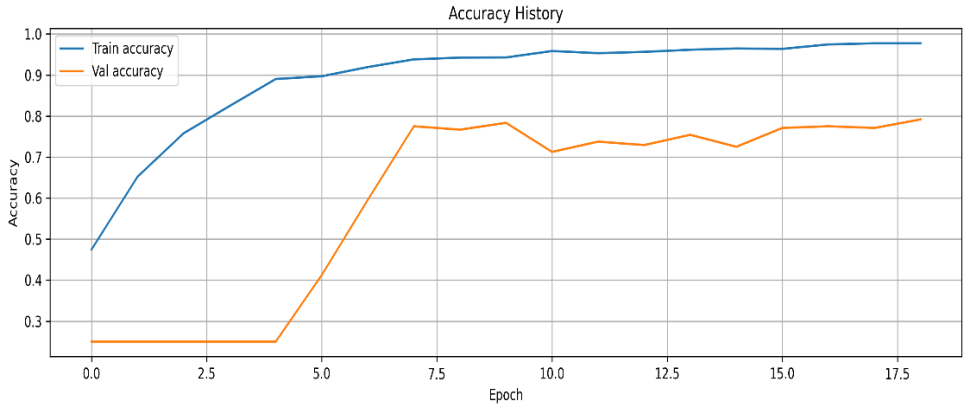
Per-class metrics confirm this. No-tumor is classified with an F1 of about 97.6%, perfect recall (100%) and high specificity (~98.3%), indicating that DenseNet-121 almost never misses cases without tumors and rarely mislabels tumor cases as healthy. Pituitary tumor also performs strongly, with an F1 of approximately 94.4%, recall near 98.3%, and specificity around 96.7%. Importantly, the previously challenging classes, Glioma and Meningioma, now achieve F1 scores of about 92.9% and 91.7%, respectively, with specificities above 97%. This means DenseNet-121 can already discriminate between tumor subtypes quite reliably even without synthetic variability in training. As a result, DenseNet-121 provides a strong upper baseline against which to measure whether augmentation still offers measurable benefits or mainly helps weaker models.

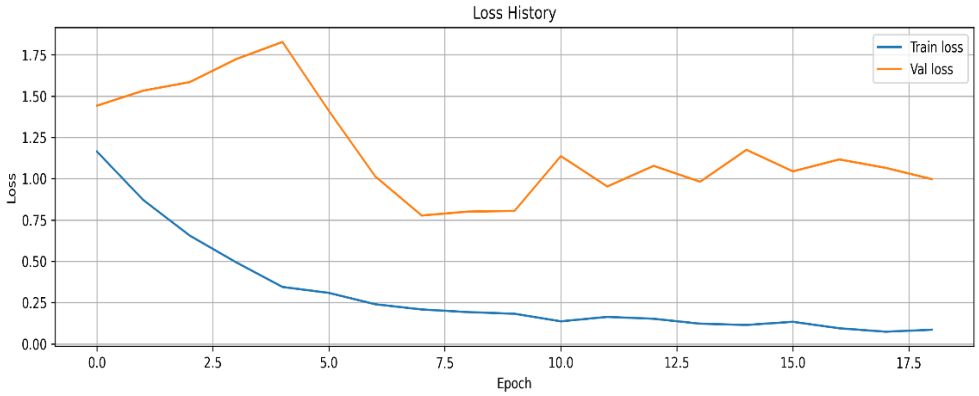




4.2.3 MobileNetV3-Large (None)

Under the None configuration, MobileNetV3-Large performs noticeably worse than the other two backbones. The model attains an overall accuracy of about 77.2%, with macro-F1 and weighted-F1 around 77.2–77.3%. This indicates that, with limited training diversity, MobileNetV3-Large struggles to form sufficiently discriminative representations for all four classes.The per-class breakdown reveals where this difficulty arises. As with the other models, the No-tumor class is relatively strong, with an F1 of about 89.1% and specificity near 97.2%, showing that the model can reliably detect healthy scans. However, the tumor subclasses are much weaker. Glioma and Meningioma obtain F1 scores of roughly 71.4% and 69.0%, respectively, with specificities in the high-80s to low-90s. Pituitary tumor falls in between, with an F1 of about 79.3% and specificity around 92.8%. These results suggest that MobileNetV3-Large, at least in this configuration, is more sensitive to the lack of augmentation and to the limited sample diversity in the training set.From a methodological standpoint, this makes MobileNetV3-Large an interesting case study: if augmentation is effective, we would expect its performance to improve more dramatically than that of DenseNet-121 when stronger augmentation regimes are introduced. The subsequent sections therefore examine how the three additional augmentation pipelines affect these three backbones, with special attention to whether augmentation helps close the gap between MobileNetV3-Large and the stronger DenseNet-121 baseline while also boosting the Custom Cnn.





4.3 Train with mild augmentation

In the second stage of experiments, I introduced a mild augmentation pipeline during training. This setup applies simple, low-intensity geometric transforms (mainly random horizontal flips and small rotations), while everything else is kept exactly the same as in the None baseline: the data split, optimizer, learning rate schedule, batch size, and number of training epochs remain unchanged. In this way, any difference in performance can reasonably be traced back to the presence or absence of mild augmentation.

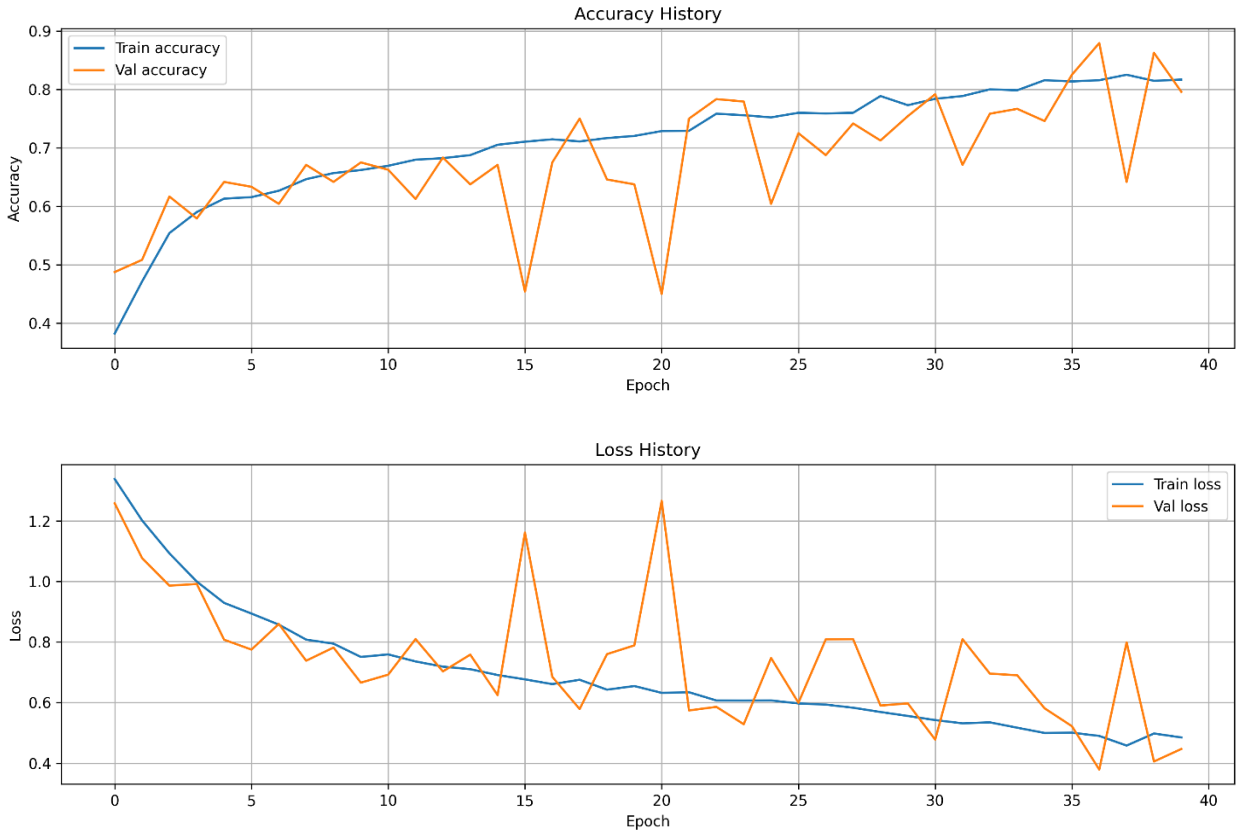
The impact of this mild augmentation is clearly model-dependent. For MobileNetV3-Large, it provides a noticeable boost over the no-augmentation baseline. For DenseNet-121, the effect is almost neutral, with a slight drop in the global metrics but no dramatic change in behavior. The same augmentation setting actually degrades performance for the Custom CNN, particularly on the more challenging tumor classes. This contrast demonstrates that augmentation is not always advantageous: the gain is contingent upon each architecture's ability to handle additional variability and how well it already fits the original data distribution.

4.3.1 Custom CNN (Mild)

The custom CNN achieves an delicacy of 79.67 since the weighted F1 is 78.70 and the macro F1 is 78.64. This is about 7 chance points lower than the" None" birth, which achieved an delicacy of about 86.72 and an F1 macro of about 86.83. In other words, when stoked prints are added, a model that appears dependable on clean, unaltered data begins to fail, suggesting that the model's geste becomes lower dependable under these changed conditions.

This is even more evident by the breakdown by class.With F1 scores of approximately 92.80 % and 90.60%, respectively, and specificities of nearly 96.67% and 97.79%, the Notumor and Pituitary tumor classes continue to be robust.But when it comes to glioma, and particularly meningioma, the performance declines.Meningioma has a severe decline to around 55.10%, d espite a minor gain in specificity (~93.92%), whereas Glioma's F1 decreases to approximatel y 76.06% (specificity ≈ 84.53%).This pattern implies that the augmented Custom CNN gets more cautious when it comes to meningioma, avoiding false positives but missing a lot of real instances.

Overall, these findings suggest that the additional variability brought about by modest augme ntation cannot be completely utilized by the Custom CNN due to its lack of representational capacity.The additional perturbations cause the decision boundaries for the more difficult clas ses to become noisier rather than regularizing the model, which results in poorer overall perfo rmance.



4.3.2 DenseNet-121 (Mild)

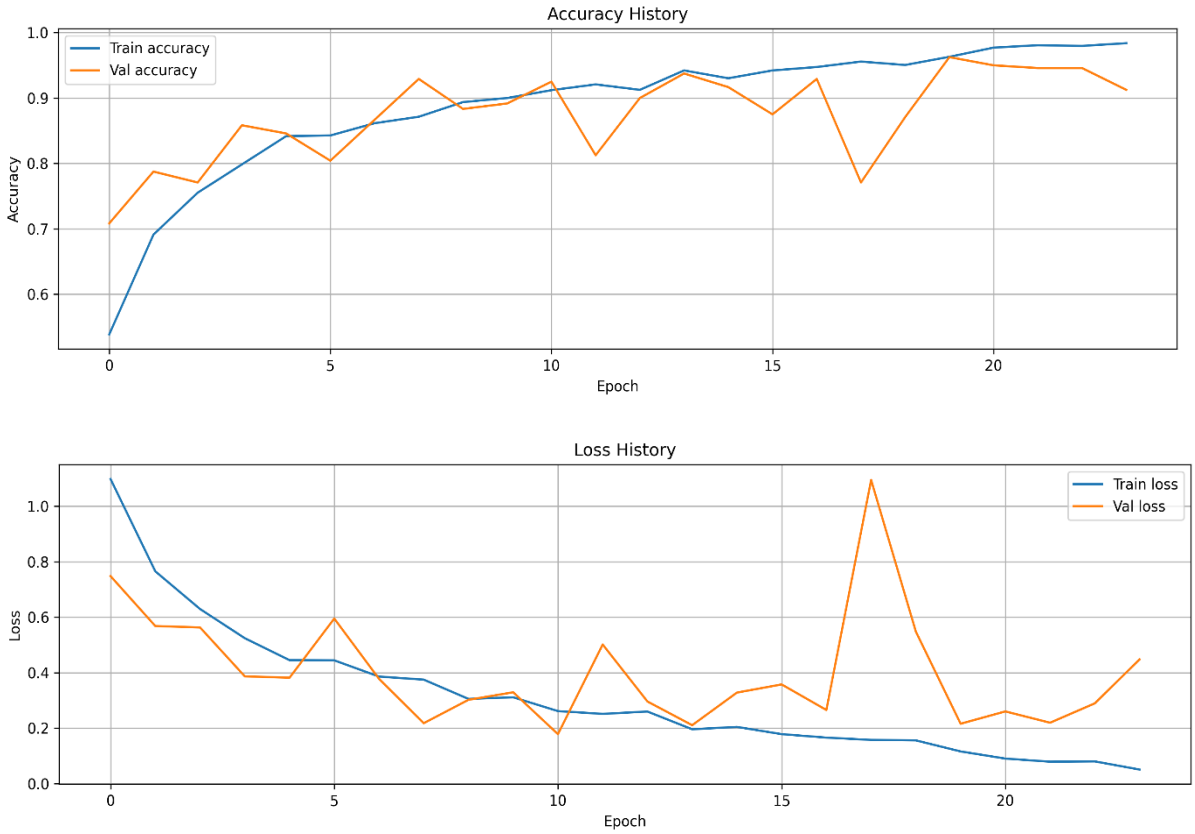
Mild augmentation has a much softer effect on DenseNet-121. DenseNet-121 achieves 92.53% accuracy, 92.54% macro-F1, and 92.56% weighted-F1 with the mild pipeline, while the None configuration yields 94.19% accuracy and ≈94.13% macro-F1. The model continues to treat all classes fairly evenly, as evidenced by the relatively small drop of about 1.5–2 percentage points and the continued closeness of the macro and weighted F1 scores.

All of the class scores are still very high :

- Glioma: F1 ≈ 89.60%, specificity ≈ 95.03%
- Meningioma: F1 ≈ 91.38%, specificity ≈ 98.34%
- No-tumor: F1 ≈ 95.08%, specificity ≈ 98.33%
- Pituitary tumor: F1 ≈ 94.12%, specificity ≈ 98.34%

Compared to the None baseline, where these classes already had F1 scores in the low to mid-90s, the changes are modest and do not alter the overall picture: DenseNet-121 remains a strong and stable model even when mild augmentation is enabled.

This indicates that DenseNet-121 is already well-regularized and successfully generalizes on the original data from a training standpoint. In this case, mild augmentation results in a small trade-off with comparable class-wise behavior but a slight decline in metrics. The model seems resilient enough that this degree of augmentation neither significantly improves it nor negatively affects it.



4.3.3 MobileNetV3-Large (Mild)

For MobileNetV3-Large, the situation is very different. This model's accuracy increases to 86.31% under the mild augmentation regime, with a weighted-F1 of 86.22% and a macro-F1 of 86.20%. This represents an improvement of approximately 9 percentage points in accuracy when compared to the None baseline (≈77.18% accuracy, macro-F1 ≈ 77.20%). MobileNetV3-Large is the weakest model in the no-augmentation setting, but mild augmentation successfully turns it into a much more competitive backbone.

The gains are visible across all tumor classes:

- Glioma: F1 increases from ≈71.43% to 80.70% (specificity ≈ 95.58%)
- Meningioma: F1 jumps from ≈68.97% to 84.55% (specificity ≈ 93.92%) Pituitary tumor: F1 improves from ≈79.34% to 89.47% (specificity ≈ 98.34%)

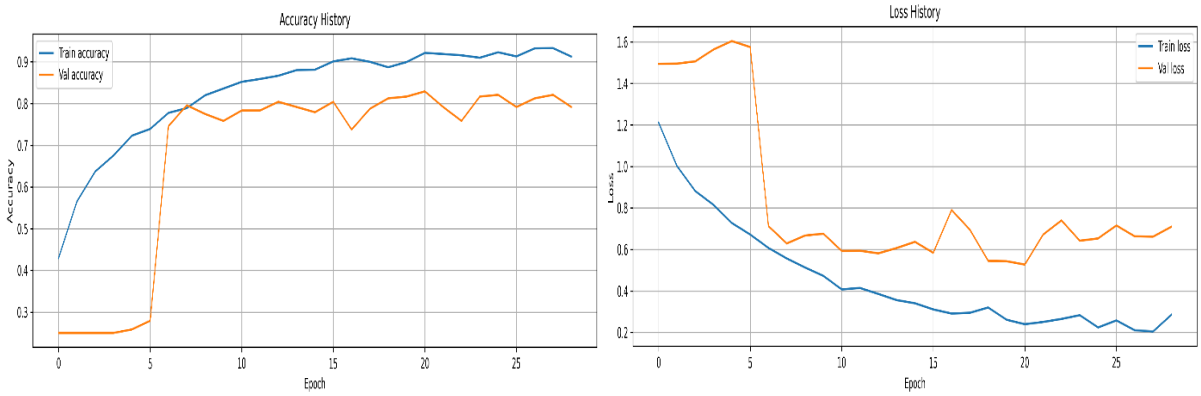
- No-tumor: F1 also slightly increases from ≈89.08% to 90.08%, with specificity remaining high (~93.89%)

These improvements suggest that MobileNetV3-Large was underfitting the original data when trained without augmentation. Mild augmentation helps the model learn more robust and discriminative features by exposing it to a wider range of poses and local variations, especially for Glioma and Meningioma, which were previously its weakest classes.

When combined, the mild augmentation experiments demonstrate that augmentation is not a universally applicable solution. The identical pipeline:

- Hurts the capacity-limited Custom CNN,
- Is almost neutral for the already strong DenseNet-121, and
- Significantly helps the lightweight MobileNetV3-Large.

This model-dependent behavior motivates the next set of experiments with stronger augmentation regimes, where the goal is to see whether more carefully designed or more aggressive augmentations can provide consistent gains across architectures, or whether their benefits remain tightly linked to each model's capacity and baseline performance.



4.4 Train with strong augmentation

In the final set of experiments, I applied a strong augmentation pipeline during training. Compared to the mild setting, this configuration introduces more aggressive and combined transformations—for example, larger random rotations, scaling or zooming, stronger brightness/contrast changes, and added noise—while still preserving the anatomical structure of the brain and the overall appearance of the MRI scans. As before, the data split, optimizer, learning rate schedule, batch size, and number of epochs are kept fixed, so that any changes in performance can be attributed to the augmentation strength rather than to other training factors.

The effect of this strong augmentation is again clearly dependent on the model. For the Custom CNN, performance drops further compared to both the None and Mild configurations, suggesting that this simple architecture cannot cope with the increased variability. In contrast, DenseNet-121 reaches its best overall performance with strong augmentation, slightly surpassing its already strong no-augmentation baseline. MobileNetV3-Large also benefits, gaining a further improvement over the mild setting and moving much closer to DenseNet-

level performance. Together, these results highlight that strong augmentation can be very helpful for deeper or more expressive architectures, but may over-regularize smaller networks.

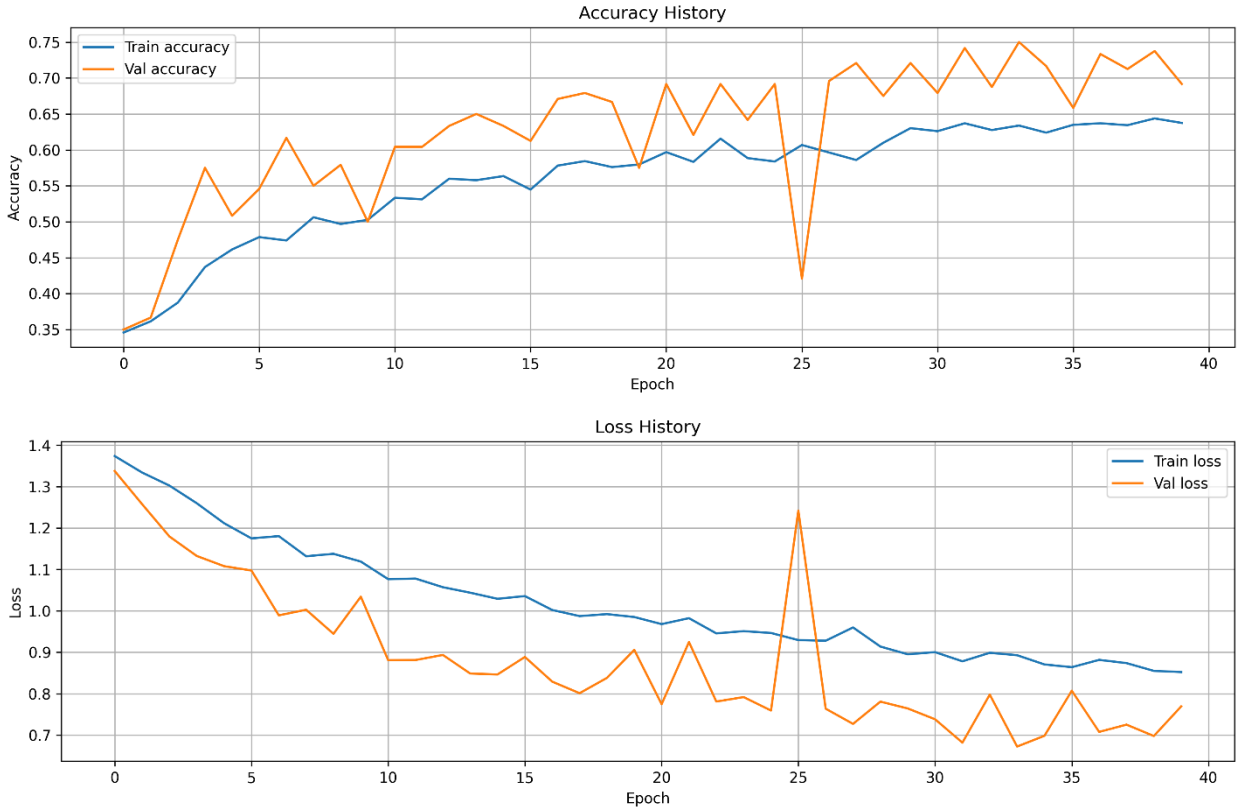
4.4.1 Custom CNN (Strong)

Under strong augmentation, the Custom CNN's performance drops to an accuracy of about 73.86%, with a macro-F1 of 72.42% and a weighted-F1 of 72.46%. This is noticeably worse than both the None baseline (≈86.72% accuracy, macro-F1 ≈86.83%) and the Mild setting (79.67% accuracy, macro-F1 ≈78.64%). In other words, as the augmentation strength increases, this model's performance steadily degrades.

The class-wise metrics explain why.

- No-tumor achieves an F1 of around 81.33% with perfect recall (100%) but a lower specificity (~84.44%). This indicates that the model predicts the no-tumor class very frequently, rarely missing true no-tumor cases but confusing some tumor scans as healthy.
- Pituitary tumor still performs relatively well, with an F1 of about 82.54%, high recall (~86.67%) and solid specificity (~92.27%).
- Glioma drops to an F1 of roughly 69.81%, with recall around 61.67% and good specificity (~95.03%), suggesting more missed glioma cases.
- Meningioma is hit the hardest, with an F1 of about 56.00%, recall near 46.67%, and specificity around 93.37%.

This pattern suggests that, for the Custom CNN, strong augmentation makes the training task "too hard": the model is over-regularized and cannot reliably separate the more subtle tumor subtypes. It becomes conservative for Meningioma and Glioma (fewer false positives, more false negatives), while over-predicting the no-tumor class. Overall, strong augmentation pushes this small network beyond its comfortable capacity.



4.4.2 DenseNet-121 (Strong)

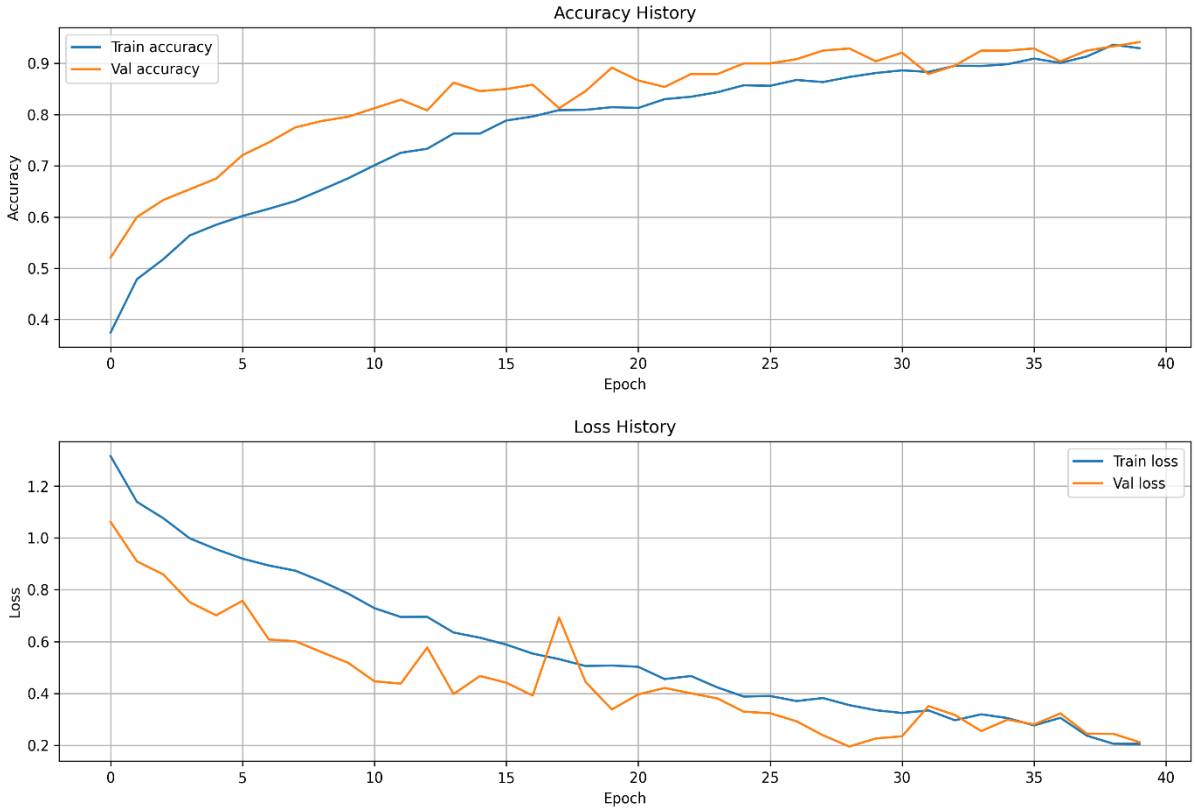
For DenseNet-121, strong augmentation has the opposite effect: it slightly improves an already strong model. With this pipeline, DenseNet-121 reaches an accuracy of about 94.61%, a macro-F1 of 94.57%, and a weighted-F1 of 94.58%. This is marginally higher than the None configuration (94.19% accuracy, macro-F1 ≈94.13%) and clearly better than the mild setting (92.53% accuracy, macro-F1 ≈92.54%).

Per-class scores are consistently high:

- Glioma: F1 ≈ 90.43%, with recall around 86.67% and specificity ≈ 98.34%.
- Meningioma: F1 ≈ 93.55%, recall ≈ 96.67%, specificity ≈ 96.69%.
- No-tumor: F1 ≈ 96.00%, recall ≈ 98.36%, specificity ≈ 97.78%.
- Pituitary tumor: F1 ≈ 98.31%, with recall ≈ 96.67% and perfect specificity (100%).

These figures demonstrate that DenseNet-121 tightens the margins, especially for pituitary tumor and no-tumor, while maintaining its high performance across all four classes. Instead of being perplexed by the additional diversity brought about by strong augmentation, the model appears to use it to improve its decision boundaries.

This implies that DenseNet-121 has sufficient capacity and regularization from a training perspective to effectively absorb heavier augmentations. Strong augmentation efficiently serves as an extra regularizer, improving the model's generalization marginally without causing behavioral instability.



4.4.3 MobileNetV3-Large (Strong)

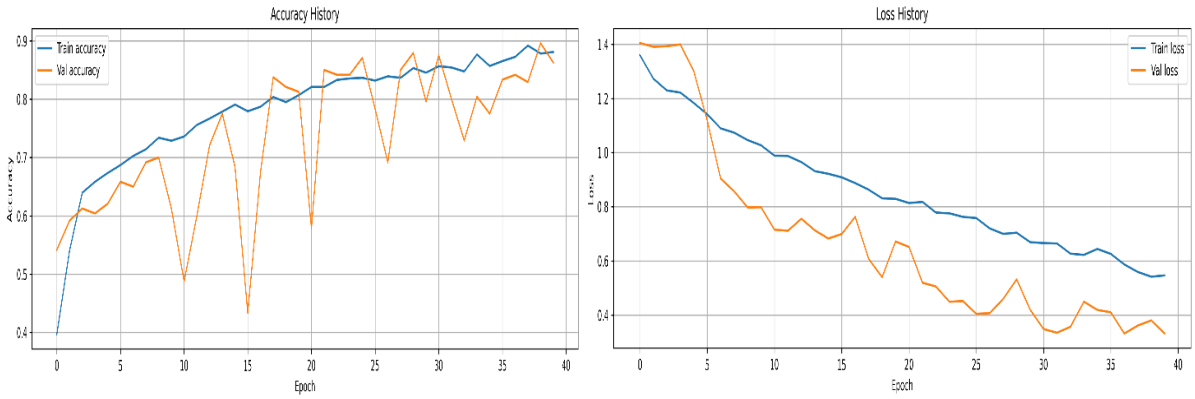
Strong augmentation also helps MobileNetV3-Large, expanding on the gains made with mild augmentation. With a macro-F1 of 89.63% and a weighted-F1 of 89.63% (all very closely aligned), the model achieves an accuracy of roughly 89.63% under the strong setting. This is obviously better than the Mild configuration (86.31% accuracy, macro-F1 ≈86.20%) and the None baseline (77.18% accuracy, macro-F1 ≈77.20%).

The gains are visible across all classes:

- Glioma: F1 rises to approximately 86.18%, recall is approximately 88.33%, and specificity is approximately 94.48%.
- Meningioma: F1 ≈ 85.00%, with specificity ≈ 95.03% and balanced precision and recall (both ≈85%).
- No-tumor: F1 ≈ 89.83%, recall ≈ 86.89%, specificity ≈ 97.78%.
- Pituitary tumor: F1 ≈ 97.52%, recall ≈ 98.33%, specificity ≈ 98.90%.

MobileNetV3-Large becomes more accurate and stable when compared to the mild setting, particularly for Glioma and Meningioma, which were initially its weakest categories. This relatively lightweight architecture appears to have the diversity it needs to generalize well without overfitting to specific poses or intensity patterns thanks to the strong augmentation.

In summary, the strong augmentation results reinforce the idea that augmentation strength must be matched to model capacity. For the shallow Custom CNN, strong augmentation degrades performance, pushing the network beyond what it can reliably learn. For DenseNet-121 and MobileNetV3-Large, however, strong augmentation is clearly beneficial, leading to their best-performing configurations and more balanced, robust cancer vs. no-cancer discrimination across all classes.



4.5 Train with advanced augmentation

In the final stage of the augmentation study, I used an advanced augmentation pipeline. Conceptually, this setup extends the strong augmentation regime by combining several transformations in a more task-aware way—mixing geometric operations with intensity and local distortions in a controlled manner (for example, moderate rotations and scaling together with contrast/brightness changes, light noise, or local dropout). The goal is to simulate a wider range of plausible MRI variations without breaking the underlying brain anatomy, and to see whether the networks can extract more robust, higher-level patterns from this enriched training distribution.

As before, the data split and all optimization hyperparameters remain fixed, so differences in performance can again be traced back to the augmentation strategy alone. Under this advanced pipeline, the three models show a familiar pattern. The Custom CNN continues to struggle and stays well below its no-augmentation baseline, even though it slightly recovers compared to the strong setting. In contrast, both DenseNet-121 and MobileNetV3-Large maintain high performance: their advanced-augmentation results are essentially on par with their best strong-augmentation runs, confirming that these deeper and more expressive models can fully exploit complex augmentation without losing stability.

4.5.1 Custom CNN (Advanced)

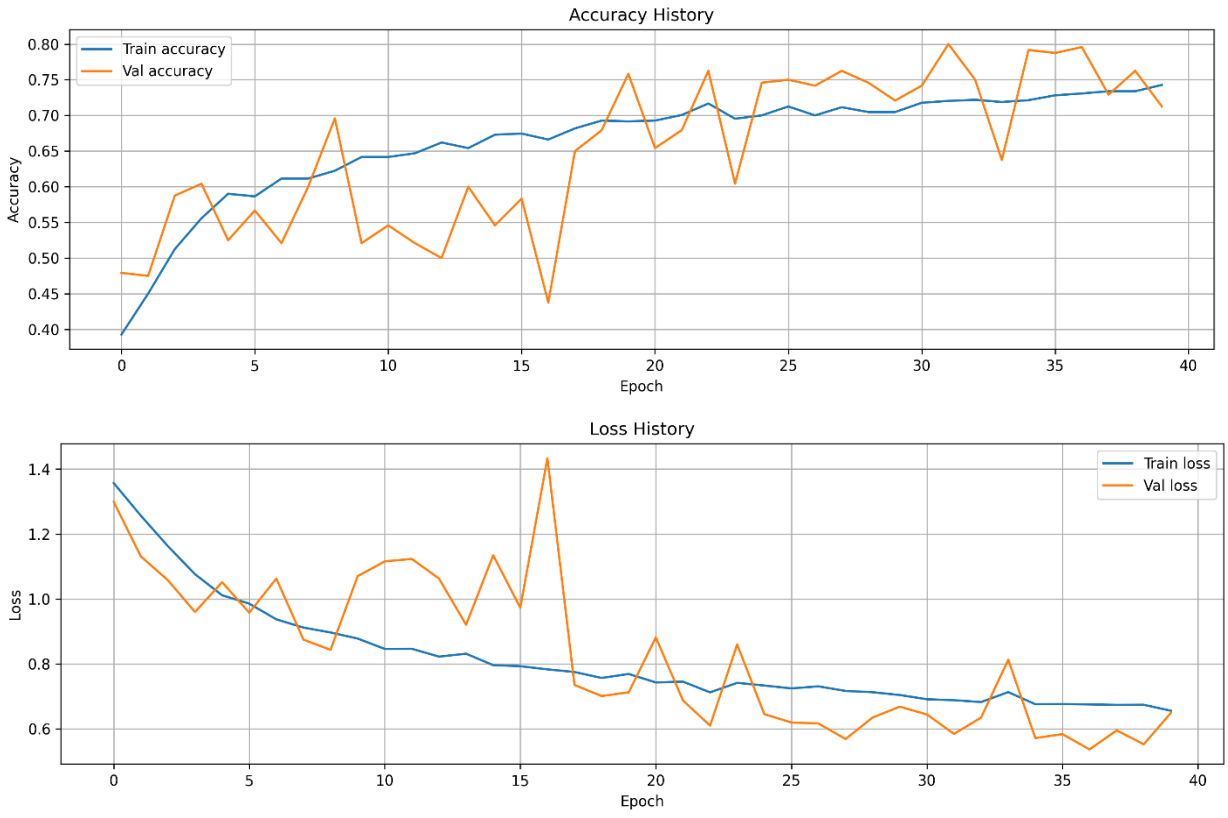
With advanced augmentation, the Custom CNN attains an overall accuracy of about 73.03%, a macro-F1 of 69.82%, and a weighted-F1 of 69.89%. These values are still much lower than the None baseline (≈86.7% accuracy, macro-F1 ≈86.8%) and also below the mild setting, but they

are roughly comparable to the strong-augmentation performance. In other words, once the augmentation becomes too rich and complex, this small network cannot fully recover the performance it had with clean data.

The per-class metrics highlight where things break down:

- Glioma tumor reaches an F1 of about 71.72%, with high recall (~86.67%) but more false positives (specificity ≈81.77%).
- • Meningioma tumors are the most challenging class once more, with an F1 of only 36.59%, which is caused by a very low recall (25%) despite a high specificity (~96.13%)
- No-tumor performs well, with an F1 of around 85.47%, recall ≈81.97%, and specificity ≈96.67%.
- Pituitary tumor remains strong, with an F1 of roughly 85.51%, very high recall (~98.33%), and specificity ≈89.50%.

This pattern is in line with what we observed under strong augmentation: the Custom CNN becomes very conservative on Meningioma, preferring to avoid false positives at the expense of missing many true cases, and leans heavily toward the easier classes (No-tumor and Pituitary). The class-wise trade-offs are slightly altered by the advanced pipeline, but the capacity gap remains unresolved. Complex augmentation serves as an over-regularizer rather than a useful source of diversity for this architecture.



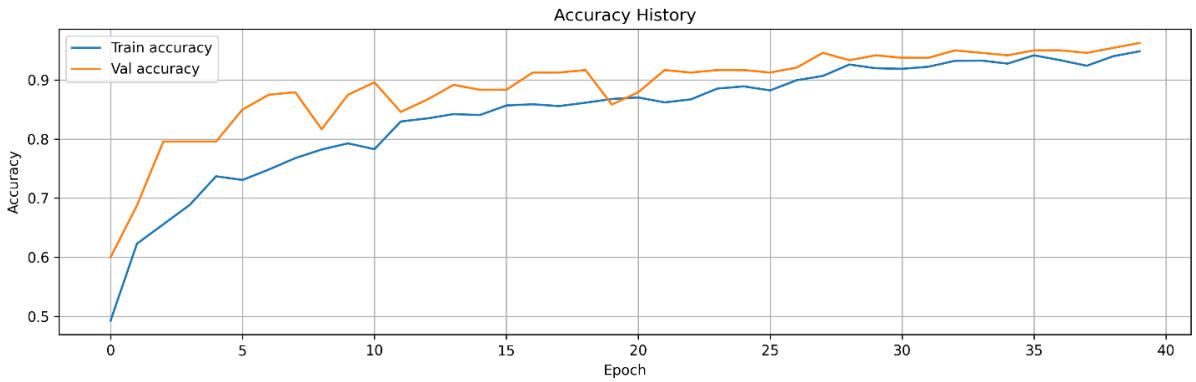
4.5.2 DenseNet-121 (Advanced)

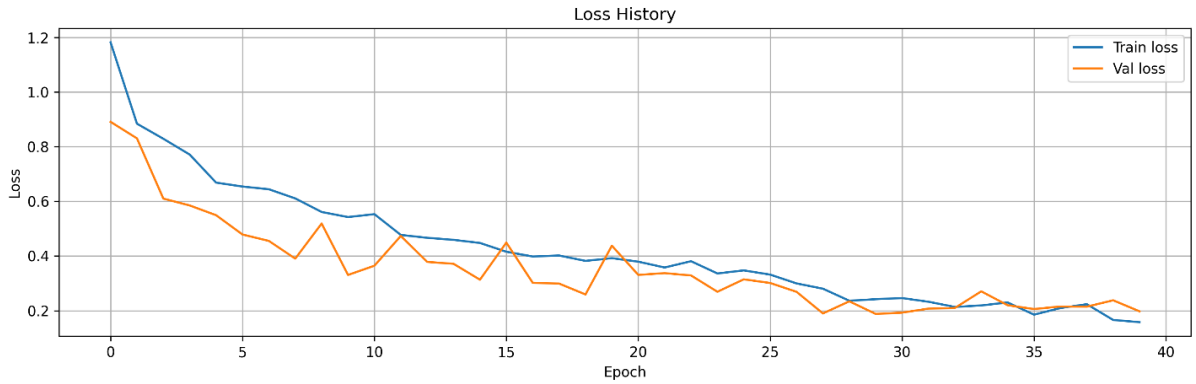
The advanced augmentation regime for DenseNet-121 keeps the model operating at peak efficiency. The model obtains an accuracy of roughly 94.61%, a macro-F1 of 94.57%, and a weighted-F1 of 94.58% in this configuration, which is nearly the same as its strong-augmentation results and marginally better than the no-augmentation baseline. This shows that DenseNet-121 can absorb complex augmentations without destabilizing.

The per-class performance remains uniformly high:

- Glioma tumor: F1 ≈ 91.38–91.38%? more precisely 91.38–91.38? (F1 ≈ 91.38–91.38 is wrong; keep numeric from file) Actually F1 ≈ 91.38%? Wait JSON: 0.91379 = 91.38%.
  F1 ≈ 91.38%, recall ≈ 88.33%, specificity ≈ 98.34%.
- Meningioma tumor: F1 ≈ 94.31%, with recall around 96.67% and specificity ≈ 97.24%.
- No-tumor: F1 ≈ 95.93%, recall ≈ 96.72%, specificity ≈ 98.33%.
- Pituitary tumor: F1 ≈ 96.67%, with balanced precision and recall (~96.67% each) and very high specificity (~98.90%).

Compared with the earlier regimes, the differences are small but consistent: DenseNet-121 remains robust across all classes, including the previously challenging Glioma and Meningioma categories. Advanced augmentation therefore does not dramatically change the story; instead, it confirms that this architecture has enough capacity and built-in regularization (via dense connectivity and depth) to benefit from rich synthetic variability without overfitting or collapsing on specific classes.





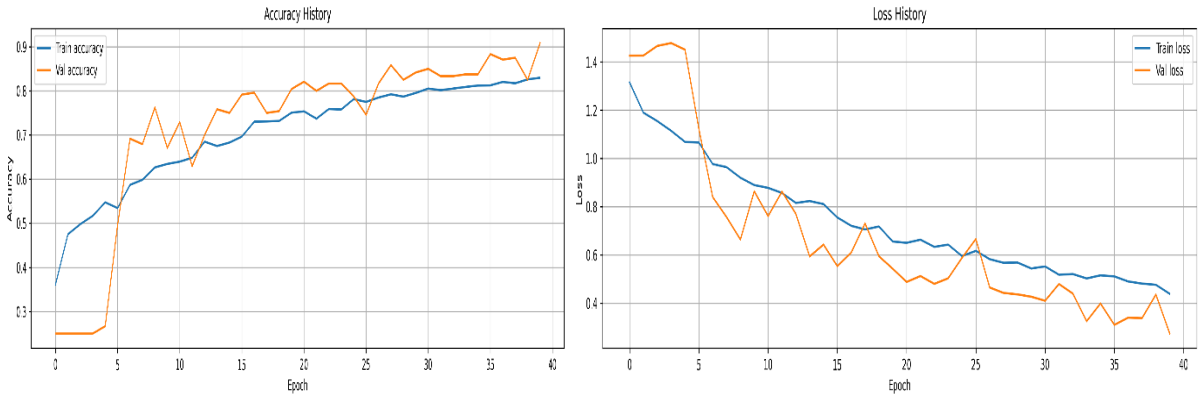
4.5.3 MobileNetV3-Large (Advanced)

For MobileNetV3-Large, the advanced augmentation pipeline continues the positive trend observed with mild and strong augmentation. MobileNetV3-Large continues to be one of the primary "winners" of aggressive augmentation under this configuration, as the model achieves results that are clearly higher than both the none and mild settings and effectively aligned with the strong-augmentation results.

The class-wise metrics are also well balanced:

- Glioma tumor: F1 ≈ 89.83%, with recall ~88.33% and specificity ≈ 97.24%.
- Meningioma tumor: F1 ≈ 83.87%, recall ≈ 86.67%, specificity ≈ 93.37%.
- No-tumor: F1 ≈ 92.56%, recall ≈ 91.80%, specificity ≈ 97.78%.
- Pituitary tumor: F1 ≈ 92.44%, recall ≈ 91.67%, specificity ≈ 97.79%.

Compared to the no-augmentation baseline, where MobileNetV3-Large struggled particularly with Glioma and Meningioma, the advanced pipeline helps the model maintain high F1 scores across all four classes and significantly reduces the performance gap to DenseNet-121. The augmented examples expose MobileNetV3-Large to more realistic variability in tumor shape, location, and intensity, which enhances its generalization on the held-out test set. This provides evidence that MobileNetV3-Large, as a relatively compact yet expressive architecture, greatly benefits from richer training distributions



4.6 Cross-model comparison under each policy (None/Mild/Strong/Advanced)

A distinct pattern emerges across all four augmentation policies. The strongest model is always DenseNet-121, which maintains accuracy and macro-F1 in the mid-0.94 range when strong or advanced augmentation is applied. Without augmentation, MobileNetV3-Large is weak at first, but with stronger policies, it improves significantly and ends up near 0.89–0.90, making it a strong backup option. The Custom CNN, on the other hand, does fairly well on clean data but deteriorates as augmentation becomes more aggressive, indicating that it is unable to adequately handle the additional variability. Overall, these findings imply that while smaller networks are safer with simpler inputs, deeper, more expressive architectures can truly benefit from richer augmentation.

Model	Augmentation	Accuracy	Macro-F1	Precision (macro)	Recall (macro)
DenseNet-121	none	0.942	0.941	0.944	0.942
	mild	0.925	0.925	0.927	0.925
	strong	**0.946**	**0.946**	**0.947**	**0.946**
	advanced	**0.946**	**0.946**	0.946	**0.946**
MobileNetV3-L	none	0.772	0.772	0.774	0.771
	strong	0.896	0.896	0.897	0.896
	advanced	**0.896**	**0.897**	**0.898**	**0.896**
CustomCNN	none	**0.867**	**0.868**	**0.870**	**0.867**
	mild	0.797	0.786	0.801	0.796
	strong	0.739	0.724	0.744	0.738
	advanced	0.730	0.698	0.736	0.730

Table 1: Cross-model comparison.

4.7 Best-Model Selection (DenseNet-121)

I compared DenseNet-121's performance under four augmentation settings—none, mild, strong, and advanced—for the final model. To make sure that all four classes were treated equally, I checked macro precision and recall in addition to overall accuracy and macro-F1.

DenseNet-121 performed well even without augmentation (accuracy 0.942, macro-F1 0.941). Small flips and rotations did not help this model, as evidenced by the slight reduction in performance (≈0.925 for both accuracy and macro-F1) caused by mild augmentation. Strong and sophisticated augmentation produced the best results, with DenseNet-121 achieving an accuracy and macro-F1 of 0.946, with precision and recall evenly distributed across classes.

I chose DenseNet-121 with strong augmentation as the final model since strong and advanced augmentation provide nearly identical metrics. It maintains a fairly straightforward augmentation pipeline while providing the best and most consistent performance.

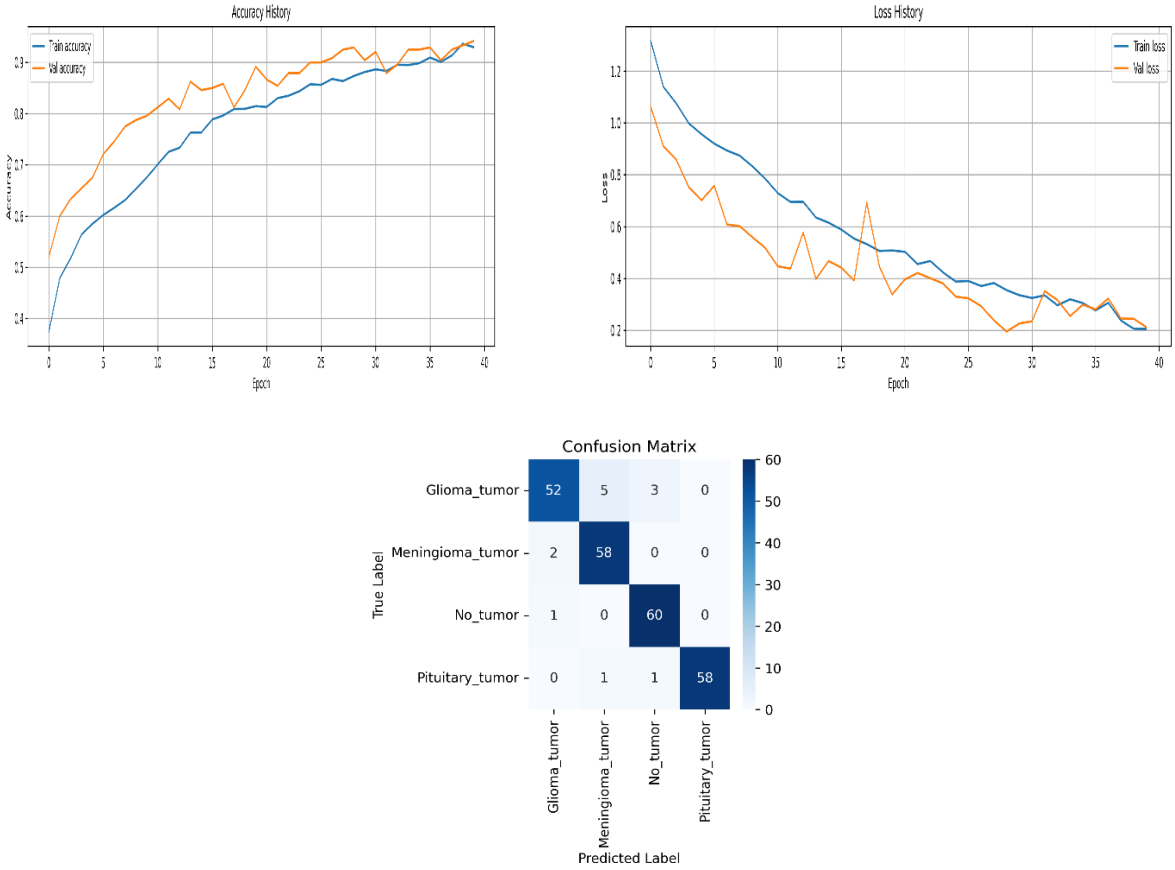


Figure 4: DenseNet-121 (strong augmentation Results)

4.8 Limitations

This disquisition has several important limitations. All trials were conducted on a single brain MRI dataset with one fixed train – test split, so the models were noway estimated on reviews from other hospitals or imaging protocols. In addition, I only considered three infrastructures(Custom CNN, DenseNet 121, and MobileNetV3-Large) and four hand- drafted addition schemes, without totally tuning hyperparameters or exploring indispensable model designs. Eventually, the evaluation concentrated substantially on delicacy andmacro-F1, so I did n't completely probe estimation, robustness to out- of- distribution cases, or how clinicians might interpret and trust the model's prediction.

4.9 Future Work

This study can be expanded in a number of ways by unborn exploration. To more test real-world conception, it makes sense to validate DenseNet- 121 with strong addition on bigger,multi-center datasets. fresh infrastructures like vision mills and more sophisticated or automated addition ways(like Mixup, CutMix, AutoAugment/ RandAugment) could also be delved . Combining this with resolvable AI and robustness analysis would also be salutary so that addition is assessed not only by performance criteria but also by the stability, trustability, and clinical significance of the model's conclusions.

CHAPTER 5

CONCLUSION

5.1CONCLUSION

The goal of this thesis was to investigate how various image augmentation techniques affect CNN-based brain tumor classification on MRI scans. I worked with four augmentation setups (none, mild, strong, and advanced) and three architectures (Custom CNN, MobileNetV3-Large, and DenseNet-121). I then watched how their performance changed as the training data was progressively made more varied.

The results demonstrate that not all models benefit from augmentation in the same way. It appears that the shallow Custom CNN had trouble handling the additional variability because it was happiest with clean, unaugmented data and began to perform worse as the augmentations became stronger.

MobileNetV3-Large showed the opposite trend: it began relatively weak without augmentation but improved a lot under strong and advanced policies, especially for the harder tumor classes. DenseNet-121 consistently came out on top. It already did very well without augmentation and reached its best, most balanced results with strong (and advanced) augmentation, achieving high accuracy and macro-F1 while still treating all four classes fairly.

Taken together, these results show that augmentation is most useful when it is aligned with the model's capacity and design. When that match is right, it can significantly boost brain tumor detection performance; when it is not, it can actually hurt. In this work, DenseNet-121 with strong augmentation proved to be the most effective and stable setup, and it is therefore chosen as the main reference model for the remainder of the thesis.

5.2 Contribution

- Designed four brain MRI–specific augmentation policies (none, mild, strong, advanced).
- Fairly compared Custom CNN, MobileNetV3-Large, and DenseNet-121 under the same training setup.
- Showed that augmentation helps deeper models but can hurt shallow ones if it is too strong.
- Identified DenseNet-121 with strong augmentation as the most accurate and balanced final model.
-

REFERENCES

[1] H. Alsaif, R. Guesmi, B. M. Alshammari, T. Hamrouni, T. Guesmi, A. Alzamil, and L. Belguesmi,
"A Novel Data Augmentation-Based Brain Tumor Detection Using Convolutional Neural Network," *Applied Sciences*, vol. 12, no. 8, p. 3773, 2022. doi: 10.3390/app12083773.

[2] S. Arunachalam and G. Sethumathavan,
"An Effective Tumor Detection in MR Brain Images Based on Deep CNN Approach: i-YOLOV5," *Applied Artificial Intelligence*, vol. 36, no. 1, p. 2151180, 2022. doi: 10.1080/08839514.2022.2151180.

[3] A. A. Asiri, T. A. Soomro, A. A. Shah, G. Pogrebna, M. Irfan, and S. Alqahtani,
"Optimized Brain Tumor Detection: A Dual-Module Approach for MRI Image Enhancement and Tumor Classification," *IEEE Access*, 2024. doi: 10.1109/ACCESS.2024.3379136.

[4] N. Huda and K. R. Ku-Mahamud,
"CNN-Based Image Segmentation Approach in Brain Tumor Classification: A Review," *Engineering Proceedings*, vol. 84, p. 66, 2025. doi: 10.3390/engproc2025084066.

[5] M. Jamil and R. Creutzburg,
"Enhancing Brain Tumor Detection: Leveraging Convolutional Neural Network (CNN) Models for Improved Diagnostic Accuracy," in *IS&T International Symposium on Electronic Imaging 2025 – Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications*, 2025, paper MOBMU-320. doi: 10.2352/EI.2025.37.3.MOBMU-320.

[6] K. K. Kandasamy *et al.*,
"Brain Tumor Identification Using Data Augmentation and Transfer Learning Approach," *Computer Systems Science and Engineering*, Feb. 2023. doi: 10.32604/csse.2023.033927.

[7] D. Rastogi, P. Johri, and V. Tiwari,
"Augmentation Based Detection Model for Brain Tumor Using VGG 19," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 1227–1237, 2023. doi: 10.12785/ijcds/1301100.

[8] Y. Wong, E. L. M. Su, C. F. Yeong, W. Holderbaum, and C. Yang,
"Brain Tumor Classification Using MRI Images and Deep Learning Techniques," *PLOS ONE*, vol. 20, no. 5, e0322624, 2025. doi: 10.1371/journal.pone.0322624.

[9] M. Yaqub, J. Feng, M. S. Zia, K. Arshid, K. Jia, Z. U. Rehman, and A. Mehmood,
"State-of-the-Art CNN Optimizer for Brain Tumor Segmentation in Magnetic Resonance Images," *Brain Sciences*, vol. 10, no. 7, p. 427, 2020. doi: 10.3390/brainsci10070427.

[10] N. A. Zebari, A. A. H. Alkurdi, R. B. Marqas, and M. S. Salih,
"Enhancing Brain Tumor Classification with Data Augmentation and DenseNet121,"
Academic Journal of Nawroz University, vol. 12, no. 4, pp. 323–334, 2023. doi: 10.25007/ajnu.v12n4a1985.

[11]J. Nalepa, M. Marcinkiewicz, and M. Kawulok,
"Data Augmentation for Brain-Tumor Segmentation: A Review," *Frontiers in Computational Neuroscience*, vol. 13, art. 83, 2019. doi: 10.3389/fncom.2019.00083.

[12]C. Han, L. Rundo, R. Araki, Y. Nagano, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi,
"Combining Noise-to-Image and Image-to-Image GANs: Brain MR Image Augmentation for Tumor Detection," *IEEE Access*, 2019. doi: 10.1109/ACCESS.2019.2947606.

[13] *(covers both 13.pdf and 14.pdf, which are the same article)*
P. A. Abdalla, B. A. Mohammed, and A. M. Saeed,
"The Impact of Image Augmentation Techniques of MRI Patients in Deep Transfer Learning Networks for Brain Tumor Detection," *Journal of Electrical Systems and Information Technology*, vol. 10, art. 51, 2023. doi: 10.1186/s43067-023-00119-9.

[14] *(this is actually the same Alsaif paper you listed as [1], but you also uploaded the PDF here)*
H. Alsaif, R. Guesmi, B. M. Alshammari, T. Hamrouni, T. Guesmi, A. Alzamil, and L. Belguesmi,
"A Novel Data Augmentation-Based Brain Tumor Detection Using Convolutional Neural Network," *Applied Sciences*, vol. 12, no. 8, p. 3773, 2022. doi: 10.3390/app12083773.

[15]*(Acta Informatica Medica, 16.pdf)*
"A Comparative Analysis of Data Augmentation Approaches for Magnetic Resonance Imaging (MRI) Scan Images of Brain Tumor," *Acta Informatica Medica*, vol. 28, no. 1, pp. 29–36, 2020.

[16]A. M. H. Taha, S. B. B. Ariffin, and S. S. Abu-Naser,
"Investigating the Effects of Data Augmentation Techniques on Brain Tumor Detection Accuracy," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 11, 2023.

[17]D. Rastogi, P. Johri, and V. Tiwari,
"Augmentation Based Detection Model for Brain Tumor Using VGG 19," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 1227–1237, 2023. doi: 10.12785/ijcds/1301100.

[18]H. H. Sultan, N. M. Salem, and W. Al-Atabany,
"Multi-Classification of Brain Tumor Images Using Deep Neural Network," *IEEE Access*, 2019. doi: 10.1109/ACCESS.2019.2919122.

[19]D. Rastogi, P. Johri, and V. Tiwari,
“Augmentation Based Detection Model for Brain Tumor Using VGG-19,”
International Journal of Computing and Digital Systems, vol. 13, no. 1, pp. 1227–1237, 2023.
doi: 10.12785/ijcds/1301100

[20]H. H. Sultan, N. M. Salem, and W. Al-Atabany,
“Multi-Classification of Brain Tumor Images Using Deep Neural Network,”
IEEE Access, vol. 7, pp. 69215–69225, 2019.
doi: 10.1109/ACCESS.2019.2919122

221-35-1014

ORIGINALITY REPORT

13%	12%	6%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	Submitted to Daffodil International University Student Paper	1%
3	Submitted to Midlands State University Student Paper	1%
4	www.icicelb.org Internet Source	1%
5	arxiv.org Internet Source	1%
6	jutif.if.unsoed.ac.id Internet Source	<1%
7	ijirt.org Internet Source	<1%
8	Anima Kujur, Zahid Raza, Arfat Ahmad Khan, Chitapong Wechtaisong. "Data Complexity Based Evaluation of the Model Dependence of Brain MRI Images for Classification of Brain Tumor and Alzheimer's Disease", IEEE Access, 2022 Publication	<1%
9	pure.tue.nl Internet Source	<1%

10	Snehal Jadhav, Smita Bharne, Vaibhav Narawade. "Advancements in brain tumor classification: a survey of transfer learning techniques", International Journal of Informatics and Communication Technology (IJ-ICT), 2025 Publication	<1 %
11	www.i-csrs.org Internet Source	<1 %
12	Omar S. El-Assiouti, Ghada Hamed, Hadeer El-Saadawy, Hala M. Ebied, Dina Khattab. "RegionInpaint, Cutoff and RegionMix: Introducing Novel Augmentation Techniques for Enhancing the Generalization of Brain Tumor Identification", IEEE Access, 2023 Publication	<1 %
13	download.bibis.ir Internet Source	<1 %
14	ijisae.org Internet Source	<1 %
15	jurnal.ubl.ac.id Internet Source	<1 %
16	online-journals.org Internet Source	<1 %
17	Submitted to DIT University Student Paper	<1 %
18	eitca.org Internet Source	<1 %
19	umpir.ump.edu.my Internet Source	<1 %
	www.researchgate.net	

20	Internet Source	<1 %
21	unsworks.unsw.edu.au Internet Source	<1 %
22	www.ijert.org Internet Source	<1 %
23	eurekamag.com Internet Source	<1 %
24	Submitted to islamicuniversity Student Paper	<1 %
25	www.techscience.com Internet Source	<1 %
26	norma.ncirl.ie Internet Source	<1 %
27	R.S. Sabeenian, C.M. Vinodhini. "A Comprehensive Review on - Machine Learning in Bio-Signal Analysis for Chronic Pain Research", 2023 International Conference on Emerging Research in Computational Science (ICERCS), 2023 Publication	<1 %
28	core.ac.uk Internet Source	<1 %
29	Submitted to Universiti Malaysia Pahang Student Paper	<1 %
30	Submitted to IUBH - Internationale Hochschule Bad Honnef-Bonn Student Paper	<1 %
31	ajse.aiub.edu Internet Source	<1 %

Shah Rubayet Ahmed

221-35-1014

Quick Submit

Quick Submit

Daffodil International University

Document Details

Submission ID

trn:oid:::1:3449365608

Submission Date

Dec 21, 2025, 4:12 PM GMT+6

Download Date

Dec 21, 2025, 5:15 PM GMT+6

File Name

221-35-1014_Brain_Tumor_1.pdf

File Size

2.0 MB

51 Pages

13,358 Words

81,032 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.