Brain MRI Tumor Detection using Deep Learning

DenseNet121, Grad-CAM Visualization

May 2025 - June 2025

Woojin Park

Reason for Choosing this Topic (1/2)

1. Softmax + Multi-class Classification

Built a multi-class classification model using the Softmax function to improve accuracy on complex medical images with varied symptoms.

2. AI-Assisted Diagnosis

Aimed to support medical decisions through data-driven analysis and visualization, even without expert-level medical knowledge.

3. Lightweight and Accurate CNN

Improved accuracy while reducing model size by optimizing the baseline CNN architecture.

Reason for Choosing this Topic (2/2)

4. Why MRI?

MRI is widely used in clinical practice and provides reliable diagnostic information.

5. Why Brain Tumor?

Brain Tumors are a major pediatric health issue and have gained attention post-COVID, making them significant from a public health standpoint.

Key Concepts for Project Implementation:
A Brief Review of the DenseNet Paper (1/2)

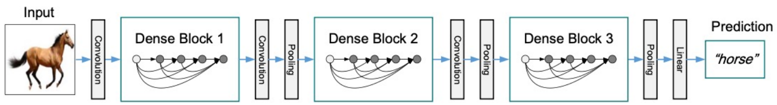


Figure 2: A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature-map sizes via convolution and pooling.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Table 1: DenseNet architectures for ImageNet. The growth rate for all the networks is $k = 32$. Note that each "conv" layer shown in the table corresponds the sequence BN-ReLU-Conv.

DenseNet Architecture

The DenseNet architecture show that the neural network is structured by grouping a **1x1 convolution** and a **3x3 convolution** into a **single bottleneck convolution unit**.

The number following DenseNet indicates the total number of layers in the network.

For example, in the DenseNet121 model, there is 1 initial convolution layer, 58 bottleneck convolution units (each consisting of 2 layers, totaling 116), 3 transition layers, and 1 classification layer, making up a total of 121 layers.

(Batch Normalization, ReLU, and Pooling layers are not included in this count.)

Key Concepts for Project Implementation:
A Brief Review of the DenseNet Paper (2/2)

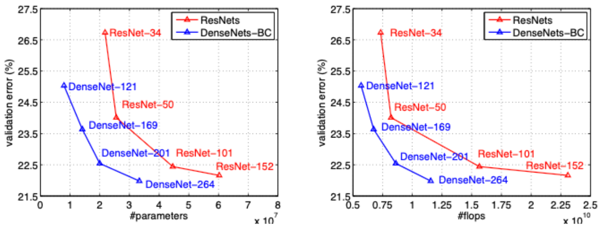


Figure 3: Comparison of the DenseNets and ResNets top-1 error rates (single-crop testing) on the ImageNet validation dataset as a function of learned parameters (*left*) and FLOPs during test-time (*right*).

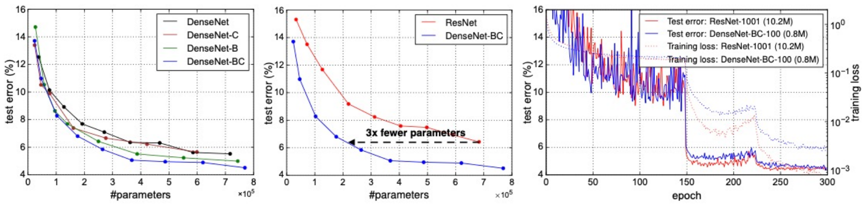


Figure 4: *Left:* Comparison of the parameter efficiency on C10+ between DenseNet variations. *Middle:* Comparison of the parameter efficiency between DenseNet-BC and (pre-activation) ResNets. DenseNet-BC requires about 1/3 of the parameters as ResNet to achieve comparable accuracy. *Right:* Training and testing curves of the 1001-layer pre-activation ResNet [12] with more than 10M parameters and a 100-layer DenseNet with only 0.8M parameters.

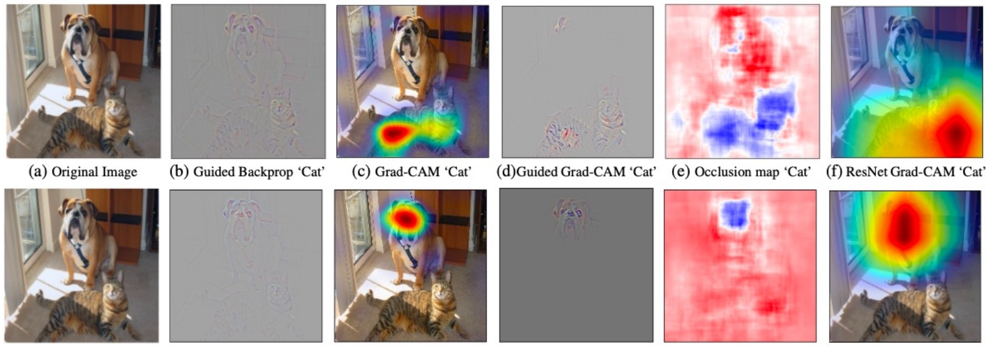
DenseNet vs ResNet

Through a comparison between DenseNet and ResNet, it is observed that **DenseNet** achieves a **lower validation error** with **fewer parameters** and **lower FLOPs** (Floating Point Operations) compared to ResNet. Therefore, it can be concluded that the DenseNet model outperforms the ResNet model in terms of **efficiency** and **performance**.

To achieve **higher accuracy** with **fewer parameters** and **reduced computational cost**, **DenseNet121** was chosen over ResNet50 as the CNN model for this project.

Key Concepts for Project Implementation:
A Brief Review of the Grad-CAM Paper

Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization 3



(a) Original Image (b) Guided Backprop 'Cat' (c) Grad-CAM 'Cat' (d)Guided Grad-CAM 'Cat' (e) Occlusion map 'Cat' (f) ResNet Grad-CAM 'Cat'

(g) Original Image (h) Guided Backprop 'Dog' (i) Grad-CAM 'Dog' (j)Guided Grad-CAM 'Dog' (k) Occlusion map 'Dog' (l)ResNet Grad-CAM 'Dog'

Fig. 1: (a) Original image with a cat and a dog. (b-f) Support for the cat category according to various visualizations for VGG-16 and ResNet. (b) Guided Backpropagation [53]: highlights all contributing features. (c, f) Grad-CAM (Ours): localizes class-discriminative regions, (d) Combining (b) and (c) gives Guided Grad-CAM, which gives high-resolution class-discriminative visualizations. Interestingly, the localizations achieved by our Grad-CAM technique, (c) are very similar to results from occlusion sensitivity (e), while being orders of magnitude cheaper to compute. (f, l) are Grad-CAM visualizations for ResNet-18 layer. Note that in (c, f, i, l), red regions corresponds to high score for class, while in (e, k), blue corresponds to evidence for the class. Figure best viewed in color.

Grad-CAM Visualization and Heatmap

Through Grad-CAM and heatmap visualization, we can **easily identify** which parts the **model learns** and **responds** to, enabling the implementation of explainable AI (XAI).

Research Hypothesis

- If the baseline CNN or ResNet50 model is used as-is for training, the accuracy is unlikely to significantly exceed that of the **DenseNet121** model. Even with pretrained weights applied, training on new medical images requires a **more suitable** architecture and neural network tuning.

- By **fine-tuning** the DenseNet121 architecture and **utilizing techniques** such as reduceLROnPlateau, callbacks, data augmentation, and **hyperparameter tuning**, the model's **accuracy can be improved**. Throughout this process, various techniques will be applied, and accuracy will be monitored accordingly.

- By applying **Grad-CAM** to the trained DenseNet121 model, it is possible to **visualize** which **regions** of the brain MRI the model focused on during classification. This **heatmap-based visualization** enables the **implementation of explainable AI** (XAI) in the brain tumor detection process.

Project Experimental Process

1. **Download Dataset** from Kaggle – 7,023 Brain Tumor MRI Images
2. **Train ResNet50 and DenseNet121**
3. **Compare ResNet50 and DenseNet121**
4. **Experimental Analysis**
5. **Visualization: Grad-CAM**

Train ResNet50 and DenseNet121

```
==========================================
Total params: 23595908 (90.01 MB)
Trainable params: 23542788 (89.81 MB)
Non-trainable params: 53120 (207.50 KB)
```

```python
# 경로 설정
BASE_DIR = 'dataset/classification'
TRAIN_DIR = os.path.join(BASE_DIR, 'train')
VAL_DIR = os.path.join(BASE_DIR, 'val')
IMG_SIZE = 224
BATCH_SIZE = 32
NUM_CLASSES = 4
EPOCHS = 15

# 데이터 로딩, 증강
train_datagen = ImageDataGenerator(
    rescale=1./255,
    horizontal_flip=True,
    rotation_range=20,
    zoom_range=0.2
)
```

ResNet50

```
==========================================
Total params: 7041604 (26.86 MB)
Trainable params: 6957956 (26.54 MB)
Non-trainable params: 83648 (326.75 KB)
```

```python
# 경로 설정
BASE_DIR = 'dataset/classification'
TRAIN_DIR = os.path.join(BASE_DIR, 'train')
VAL_DIR = os.path.join(BASE_DIR, 'val')
IMG_SIZE = 224
BATCH_SIZE = 32
NUM_CLASSES = 4
EPOCHS = 20

# 데이터 로딩, 증강
train_datagen = ImageDataGenerator(
    rescale=1./255,
    horizontal_flip=True,
    rotation_range=25,
    zoom_range=0.2,
    brightness_range=[0.8, 1.2],
)
```

DenseNet121

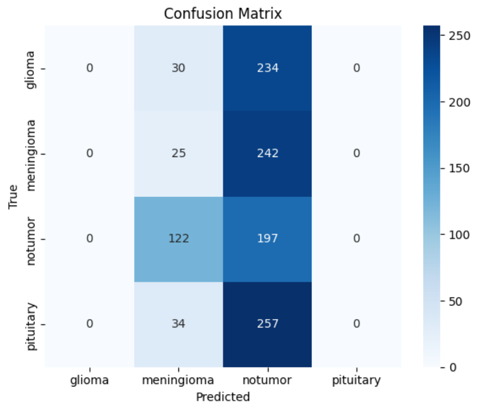
Compared to the training of ResNet50 and DenseNet121, all conditions were kept the same except for the **number of epochs** and the **brightness range parameter** in data augmentation.

Compare ResNet50 and DenseNet121

ResNet50

Classification Report:

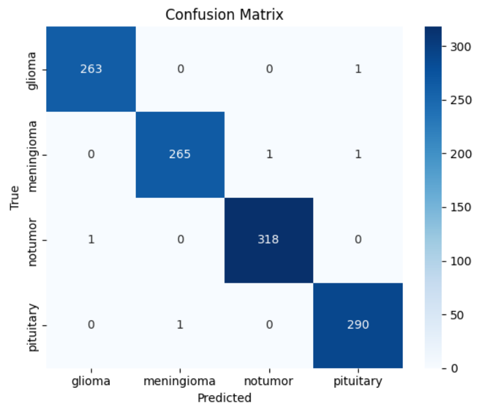
	precision	recall	f1-score	support
glioma	0.00	0.00	0.00	264
meningioma	0.12	0.09	0.10	267
notumor	0.21	0.62	0.32	319
pituitary	0.00	0.00	0.00	291
accuracy			0.19	1141
macro avg	0.08	0.18	0.11	1141
weighted avg	0.09	0.19	0.11	1141



DenseNet121

Classification Report:

	precision	recall	f1-score	support
glioma	1.00	1.00	1.00	264
meningioma	1.00	0.99	0.99	267
notumor	1.00	1.00	1.00	319
pituitary	0.99	1.00	0.99	291
accuracy			1.00	1141
macro avg	1.00	1.00	1.00	1141
weighted avg	1.00	1.00	1.00	1141

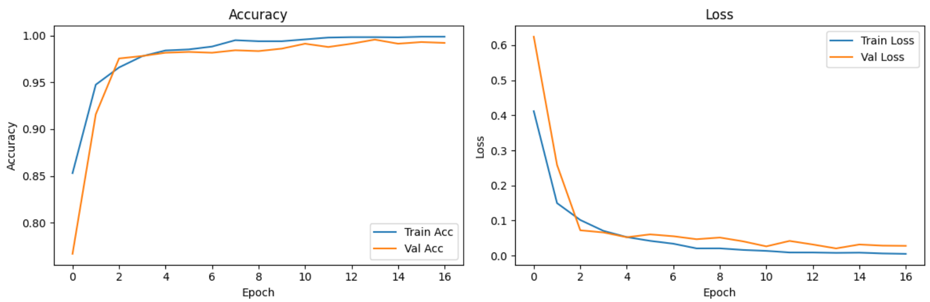


ResNet50 and DenseNet121

While ResNet50 failed to predict the **glioma** and **pituitary** classes entirely and misclassified most samples as **no tumor**, DenseNet121 successfully classified all four given classes with high accuracy.

Compared to ResNet50, the **DenseNet121** model has **fewer parameters** and a **smaller size**, while achieving **higher** classification **accuracy**.

Experimental Analysis



Accuracy – Loss Graph

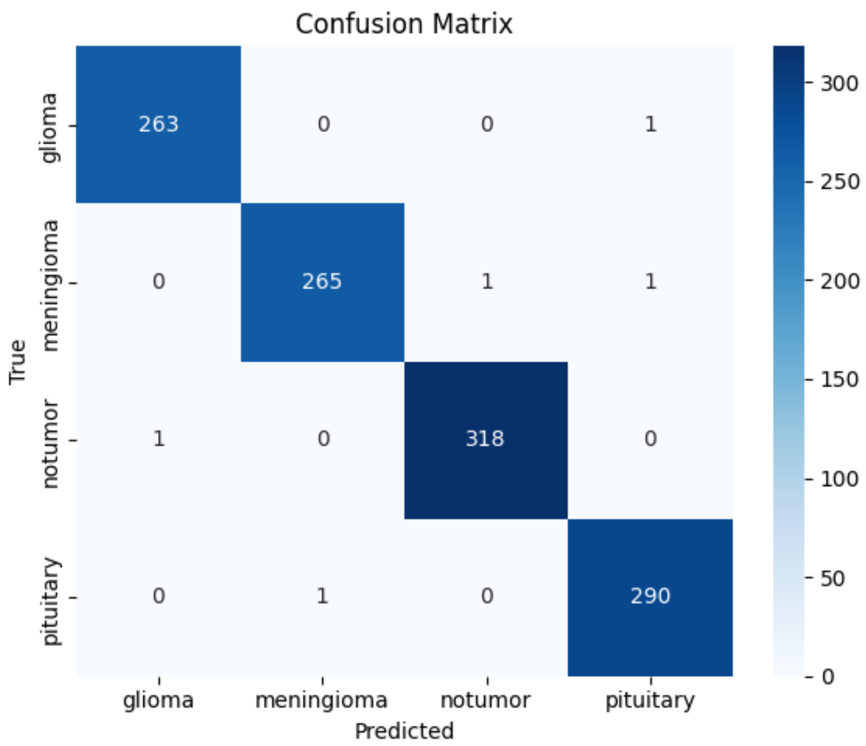
Accuracy increased for both training and validation, while loss decreased for both, indicating that overfitting did not occur.

```
Classification Report:

              precision    recall  f1-score   support

      glioma       1.00      1.00      1.00       264
  meningioma       1.00      0.99      0.99       267
     notumor       1.00      1.00      1.00       319
   pituitary       0.99      1.00      0.99       291

    accuracy                           1.00      1141
   macro avg       1.00      1.00      1.00      1141
weighted avg       1.00      1.00      1.00      1141
```

Classification Report

Attempted classification into four classes using **Softmax function** and achieved high accuracy.

Experimental Analysis

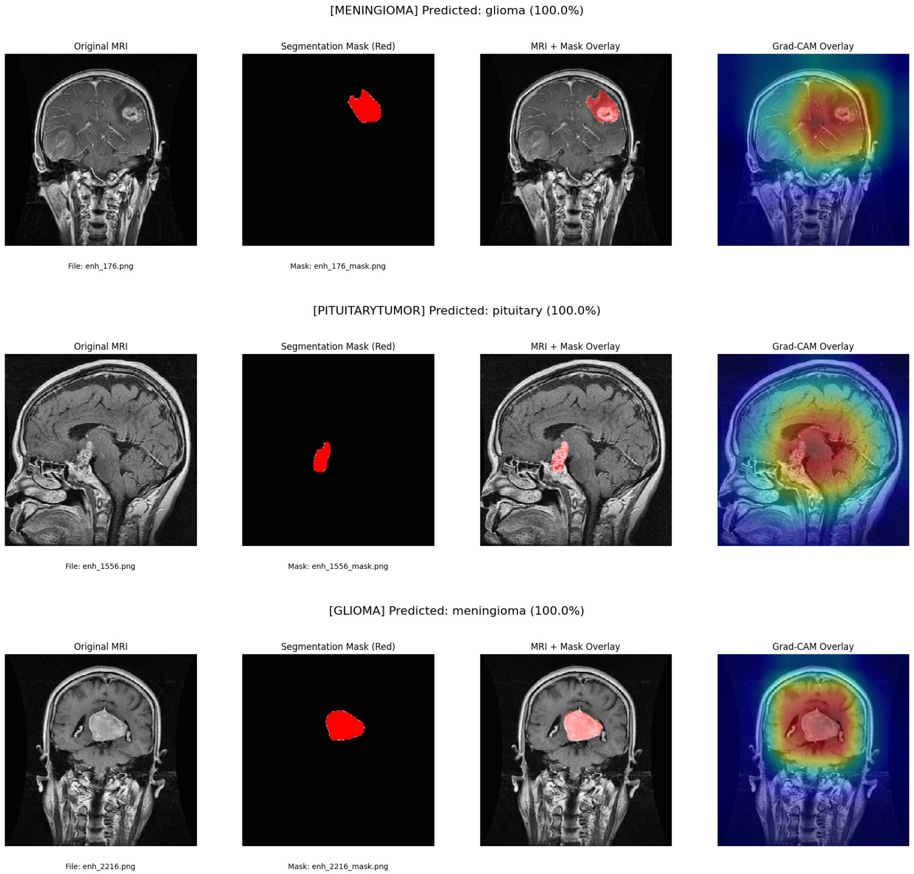


Confusion Matrix

The confusion matrix helps identify which cases were misclassified for each class.

Since most of the predicted classifications were accurate, it can be concluded that the analysis using the DenseNet121 model in this experiment was successful.

Visualization: Grad-CAM



Classification

Visualized the classification of three types of brain tumors individually.

Original MRI + Segmentation Mask

Since the dataset provides separate original images and masks, they were combined to help visualize the lesions.

Grad-CAM Overlay

Heatmap visualization was used to identify which regions the model focused on during training, providing interpretability.

Results and Discussion

Results:

- Among the ResNet50 and DenseNet121 models, the **fully fine-tuned DenseNet121** model achieved the best performance.
- Using **Grad-CAM and heatmap visualizations**, it was confirmed that the regions the model focused on during the classification of brain tumors aligned with the actual lesion areas. This demonstrates that even by simply examining the Grad-CAM overlay images, the model aligns with the goals of explainable AI.

Discussion:

- With **more diverse brain tumor datasets**, it would be possible to classify a wider range of tumor types and severity levels.
- On the current hardware, training the model took approximately one hour. It may be possible to improve efficiency by **finding a model** with fewer parameters and lower computational cost than DenseNet121, or by **designing a custom model inspired by DenseNet's architecture** specifically tailored for brain tumor classification.

Future Research Direction

- **Training a Classification Model Using Diverse Data**: It is possible to develop models capable of distinguishing brain tumors not only from brain MRI images, but also from video data, various physiological signals, and biometric inputs.

- **Direction for Utilizing the Model Architecture**: By leveraging the architecture of the current model, classification models can be extended to detect a variety of diseases beyond brain tumors.

- **Model Compression and Deployment**: By reducing the computational cost and size of the model (i.e., improving efficiency), the trained model can be deployed across various devices. This enables the development of software-based solutions, such as mobile applications for assisting brain tumor diagnosis.

- **Multimodal Extension (LLM + Medical Analysis)**: By integrating Large Language Models (LLMs), the system can provide explainable AI capabilities and be expanded into a comprehensive multimodal medical service.

Reference

RR Selvaraju et al., “**Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization**”, ICCV’17

Gao Huang et al., ”**Densely Connected Convolutional Networks**”, CVPR’17