CHARLOTTE
COLLEGE OF COMPUTING AND INFORMATICS

Deep Learning
Brain Tumor Classification

Group 1: Medical Image Classifiers
Anulitha Vardhan
Carsten Lingemann
Sundari Kala

Problem Description

- We are leveraging 3 different articles that each used a different pre-trained model (VGG16, ResNet50, InceptionV3) to classify Brain Tumor Images.
- We are also aiming to create a general medical image classification model that can be used to diagnose other medical conditions, such as leukemia, as well as chest x-rays detecting pneumonia and COVID-19.

Motivation

- Classification of medical images is important because it can lead to a correct or incorrect diagnosis and treatment plan.
- An untreated condition can continue to grow and cause irreversible damage.
- Medical images analysis can be a very useful research idea and can be a great career opportunity in the future.
- Building deep learning models for medical image analysis involves a diverse range of skills and knowledge, including computer science, mathematics, statistics, and biology.

Existing Approaches

Inception-V3

- Comparing accuracy of the Inception V3 performance with the VGG-16.
- 100 images in each category (normal/abnormal) are selected from the database of 130 subjects.
- Average accuracy of inception architecture is 95.1%, whereas VGG-16 is 92.8%
- Link to article

ResNet-50

- Used ResNet-50 to extract features of pre-preprocessed (normalized and resized) images, and fed them into a feature selection algorithm.
- Selected features used to train a SVM classifier. Learning rate, weight decay, and optimizer to improve performance.
- Used Grad-CAM to visualize the regions of the image that contributed the most towards the model's prediction.

VGG-16

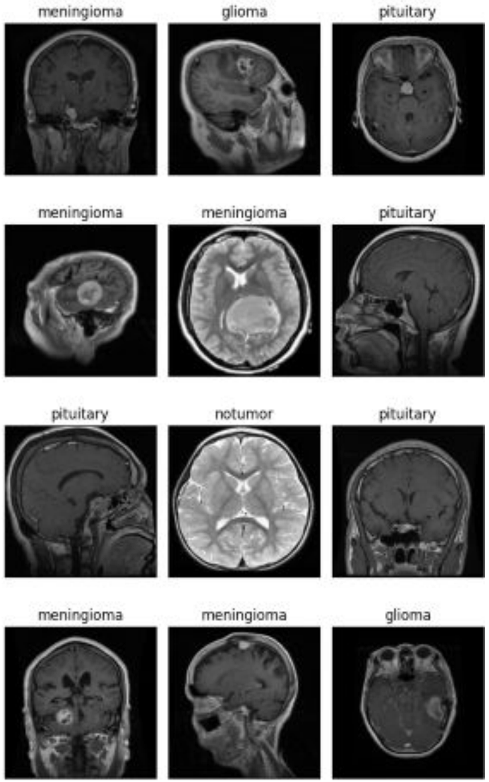
- Used 23-layer CNN for multiclass brain tumor classification (meningioma, glioma, pituitary) of 3064 images with 97.8% accuracy
- 23-layer CNN did not perform well for binary brain tumor classification (normal, abnormal) of only 152 images, so they used pretrained VGG-16 model with 100% accuracy
- Link to article

3 Different Data Sets

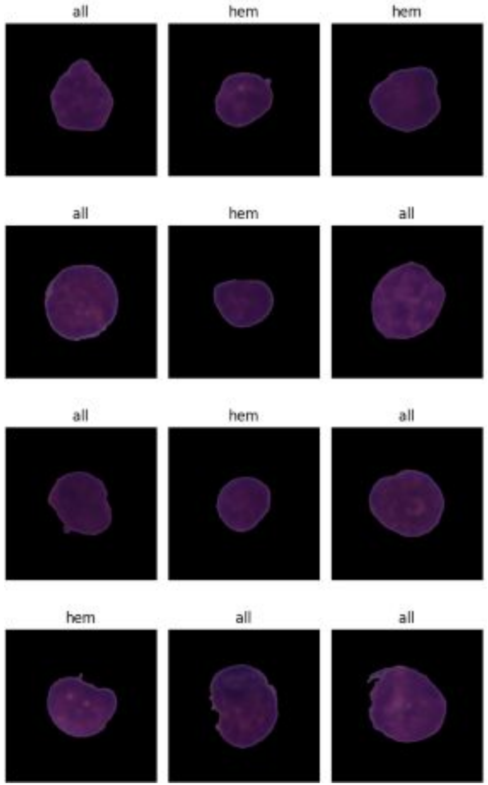
Data Set	Classes	Training	Validation	Test	Total
Brain Tumor	• Glioma Tumor • Meningioma Tumor • Pituitary Tumor • Normal	4570	1142	1311	7023
Leukemia	• ALL - Acute Lymphoblastic Leukemia • HEM - Normal	6090	1260	1263	8613
Chest X-Rays	• COVID-19 • Pneumonia-Bacterial • Pneumonia-Viral • Normal	4603	1841	2764	9208

Example Images

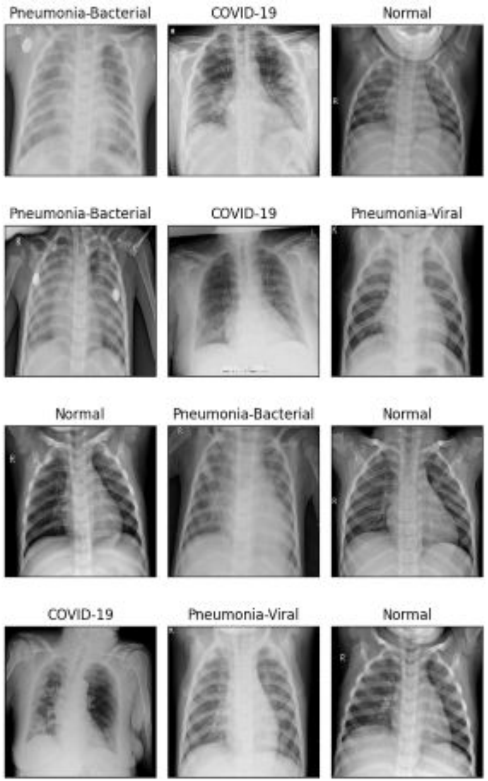
Brain Tumor



Leukemia



Chest X-Ray

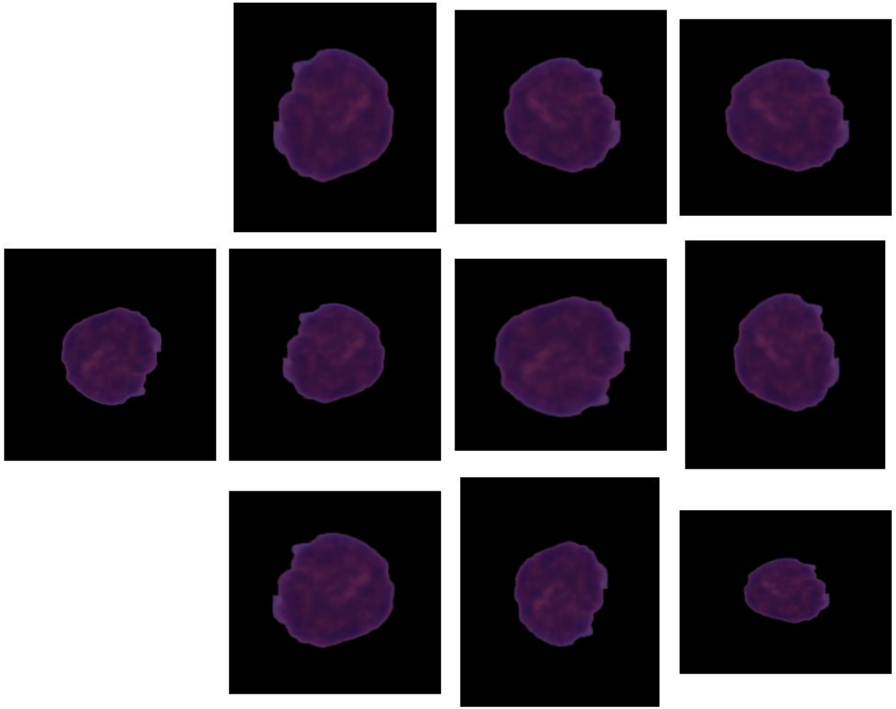


VGG-16 Image Augmentation

<u>Transformations Applied:</u>

- RandomZoom(0.3)
- RandomFlip(“horizontal”)
- RandomWidth(0.2)
- RandomHeight(0.2)
- RandomFlip(“vertical”)

Leukemia Augmentation Example



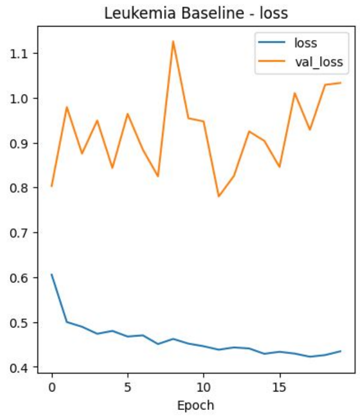
VGG-16 Experiments / Observations

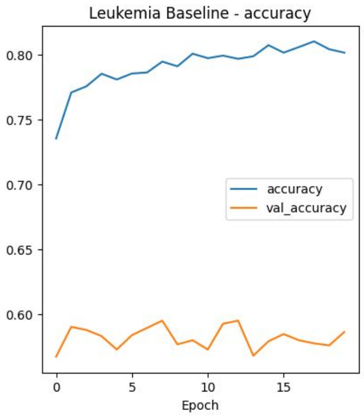
```
Model: "bt_baseline"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_4 (InputLayer)        [(None, 224, 224, 3)]     0

 sequential_1 (Sequential)   (None, None, None, 3)     0

 resizing_1 (Resizing)       (None, 224, 224, 3)       0

 tf.__operators__.getitem_1   (None, 224, 224, 3)      0
 (SlicingOpLambda)

 tf.nn.bias_add_1 (TFOpLambd  (None, 224, 224, 3)      0
 a)

 vgg16 (Functional)          (None, 7, 7, 512)         14714688

 global_average_pooling2d_1   (None, 512)              0
 (GlobalAveragePooling2D)

 dense_4 (Dense)             (None, 1024)              525312

 dense_5 (Dense)             (None, 1024)              1049600

 dense_6 (Dense)             (None, 512)               524800

 dropout_1 (Dropout)         (None, 512)               0

 dense_7 (Dense)             (None, 4)                 2052

=================================================================
Total params: 16,816,452
Trainable params: 2,101,764
Non-trainable params: 14,714,688
```

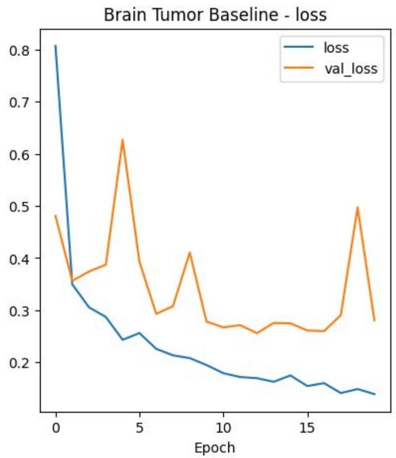
- Adding Batch Normalization layer between each dense function improved the Chest X-Ray test accuracy by ~1%, but decreased the test accuracy on the Brain Tumor and Leukemia data sets
- Using Binary Cross-entropy for the Leukemia data set, along with the sigmoid activation function on the output layer, and RMSProp optimizer improved the test accuracy by ~7%
- Increasing epochs from 25 to 50 led to overfitting, and reduced the test accuracy on all 3 data sets

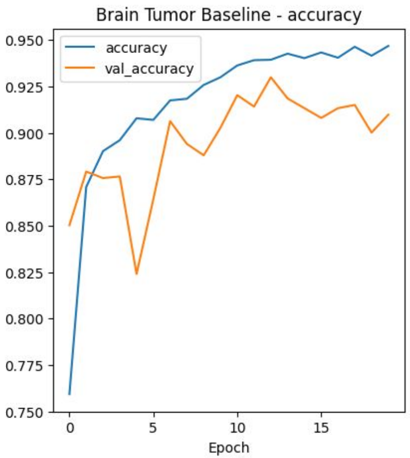
VGG-16 Results

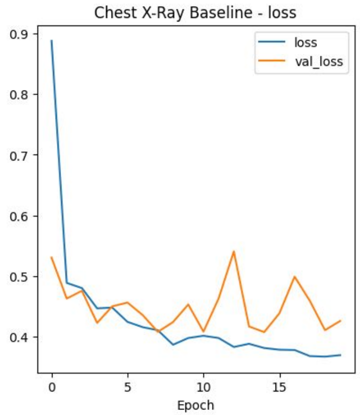
Test Accuracy	
Brain Tumor	91.99%
Leukemia	69.20%
Chest X-Ray	84.19%

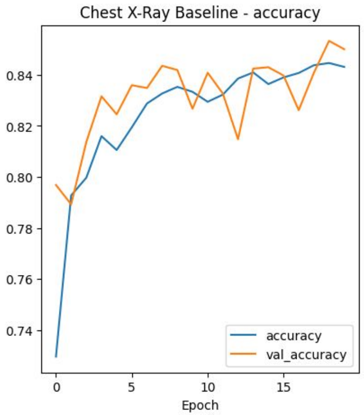




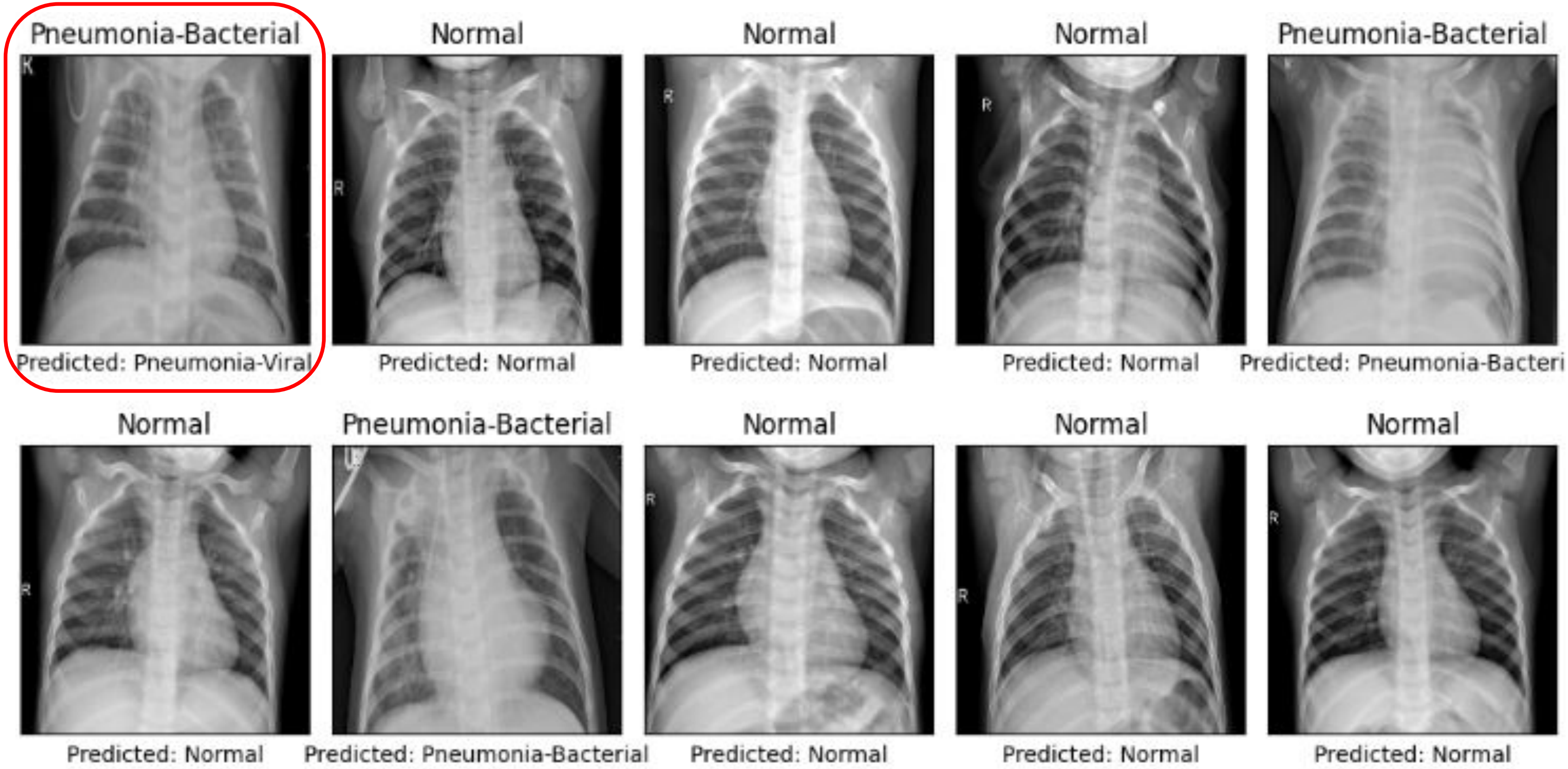








VGG-16 Predictions

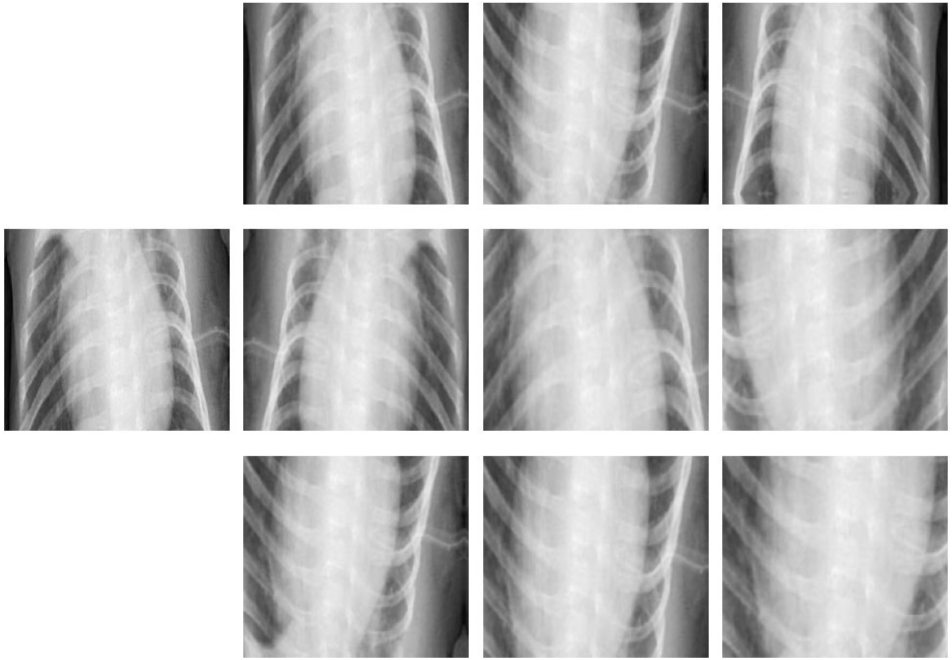


Inception-V3 Image Augmentation

Transformations Applied:

- RandomZoom(0.2)
- RandomFlip(“horizontal”)
- RandomWidth(0.1)
- RandomHeight(0.1)
- RandomFlip(“vertical”)
- RandomCrop(height=180, width=180)

Chest X-Ray Augmentation Example



Inception-V3 Experiments / Observations

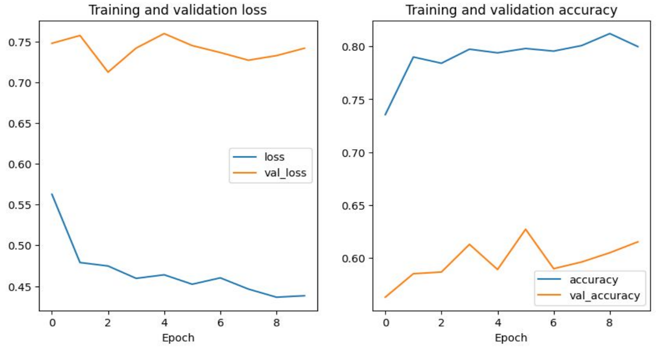
- Tried different optimizers, activation and loss functions.
- Added the Batch Normalization to improve the performance, however it didn't predicted the values accurately.
- The following activation, loss and optimizers performed well for my model.
  - Activation - softmax
  - Optimizer - Adam
  - Loss function - sparse categorical crossentropy

```
Model: "model_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_3 (InputLayer)        [(None, 224, 224, 3)]     0

 sequential_2 (Sequential)   (None, 180, 180, 3)       0

 resizing_1 (Resizing)       (None, 224, 224, 3)       0

 tf.math.truediv_1 (TFOpLamb  (None, 224, 224, 3)      0
 da)

 tf.math.subtract_1 (TFOpLam  (None, 224, 224, 3)      0
 bda)

 inception_v3 (Functional)   (None, 5, 5, 2048)        21802784

 max_pooling2d_5 (MaxPooling  (None, 2, 2, 2048)       0
 2D)

 flatten_1 (Flatten)         (None, 8192)              0

 dense_2 (Dense)             (None, 512)               4194816

 dropout_1 (Dropout)         (None, 512)               0

 dense_3 (Dense)             (None, 4)                 2052

=================================================================
Total params: 25,999,652
Trainable params: 4,196,868
Non-trainable params: 21,802,784
_________________________________________________________________
```

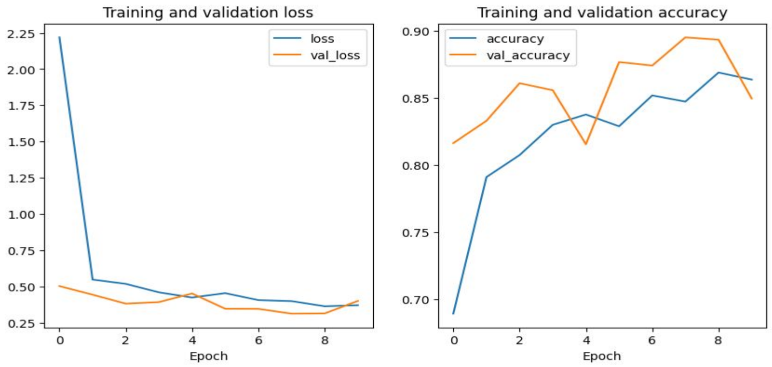
Inception-V3 Results

Test Accuracy	
Brain Tumor	86.12%
Leukemia	75.14%
Chest X-Ray	75.58%

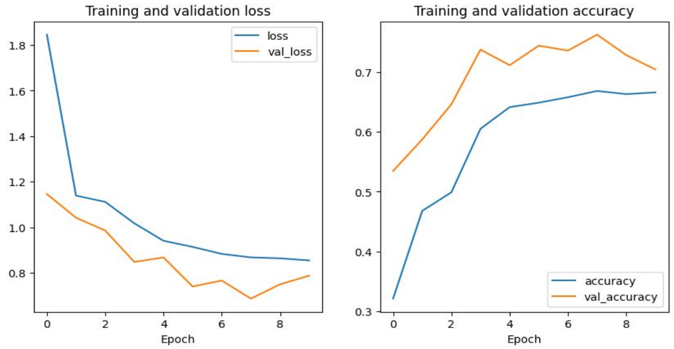
Leukemia



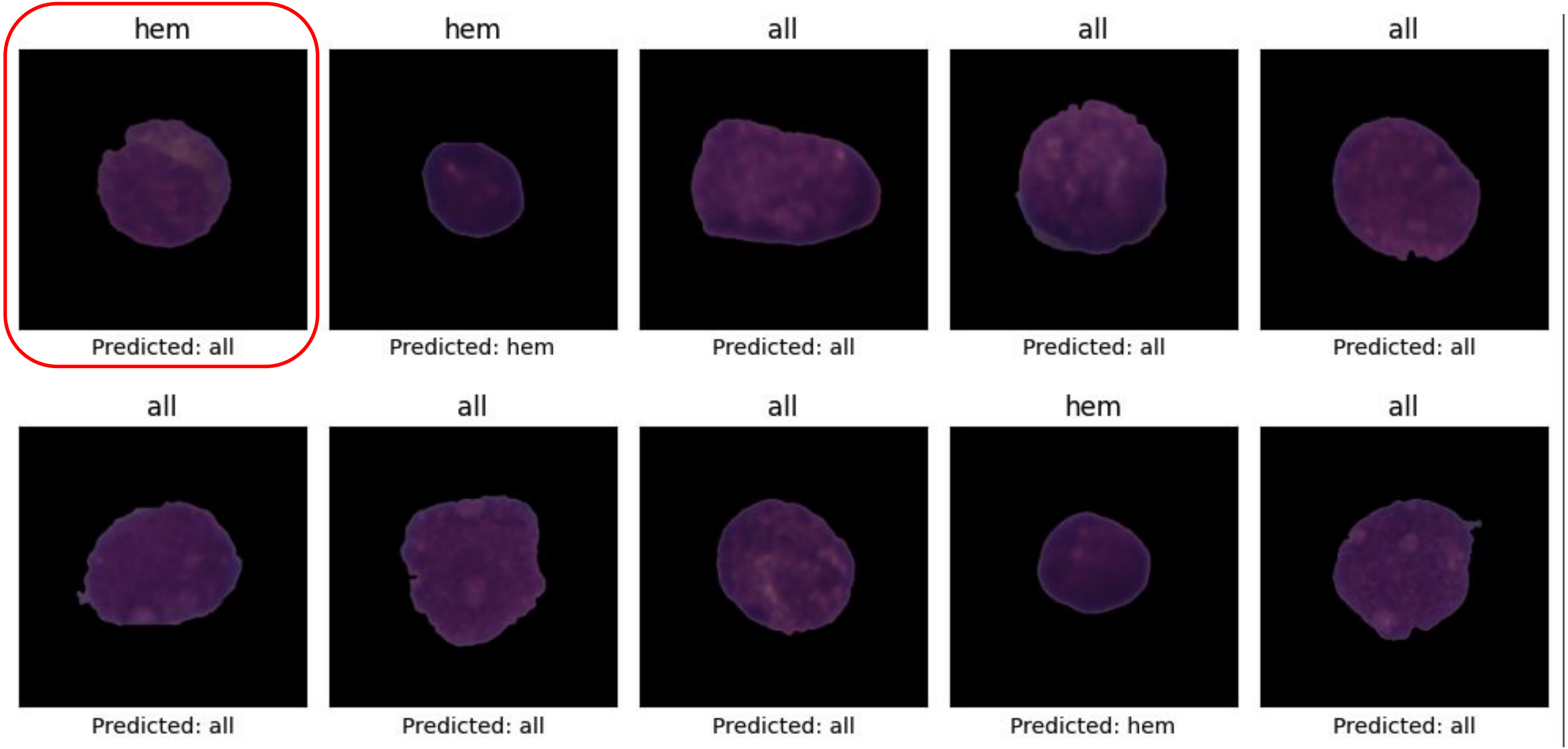
Brain Tumor



Chest X-Ray



Inception-V3 Predictions

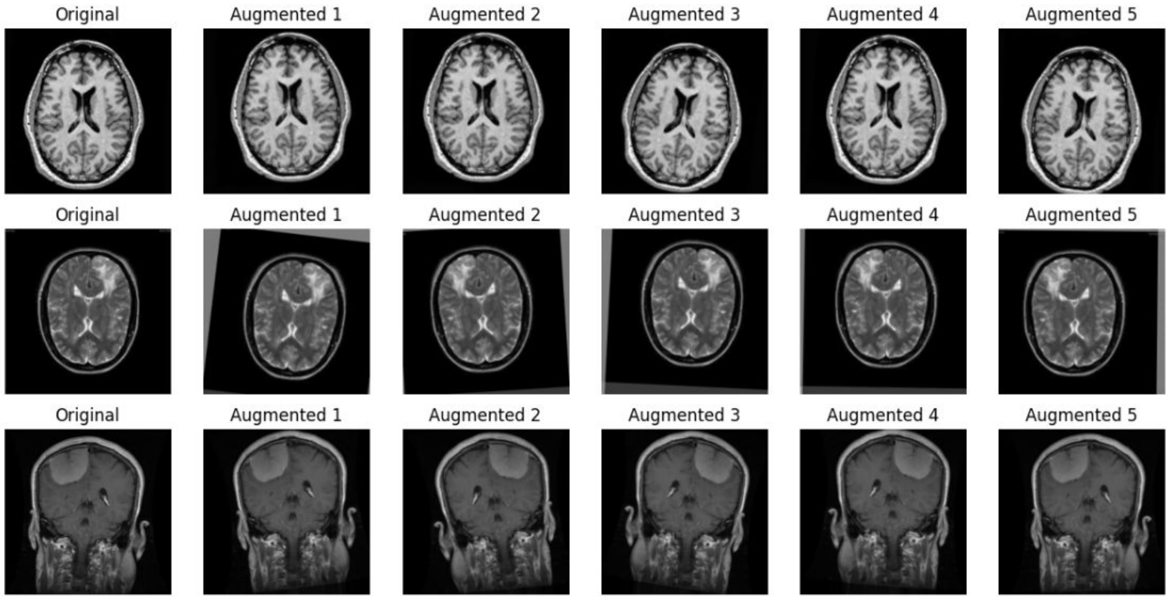


ResNet-50 Image Augmentation

Transformations Applied:

- rotation_range=10,
- width_shift_range=0.05,
- height_shift_range=0.05,
- horizontal_flip=True

Leukemia Augmentation Example



ResNet-50 Experiments / Observations

```
 global_average_pooling2d (Glob  (None, 2048)        0           ['conv5_block3_out[0][0]']
 alAveragePooling2D)

 dropout (Dropout)              (None, 2048)        0           ['global_average_pooling2d[0][0]'
                                                                 ]

 dense (Dense)                  (None, 4)           8196        ['dropout[0][0]']

==================================================================================================
Total params: 23,595,908
Trainable params: 23,542,788
Non-trainable params: 53,120
```

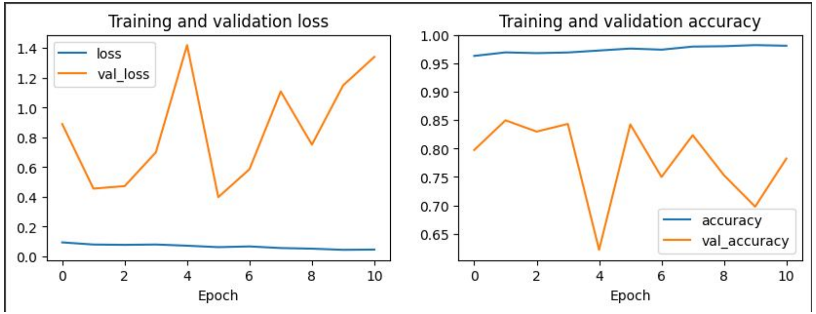
Complicated architecture reduced val-accuracy

- Training the entire Resnet [unfrozen] with early stopping, and LRonPlateau with only Global Average pooling layer, Dropout and Dense layers yielded 97% val_accuracy and 98% test accuracy.
- ResNet50 model was not optimal for classifying leukemia and chest MRI scans at only 50% and 30% accuracy respectively.
- MRI images performed poorly with complex architecture and failed to generalize for X-ray and Leukemia datasets which possibly require a more complex model.

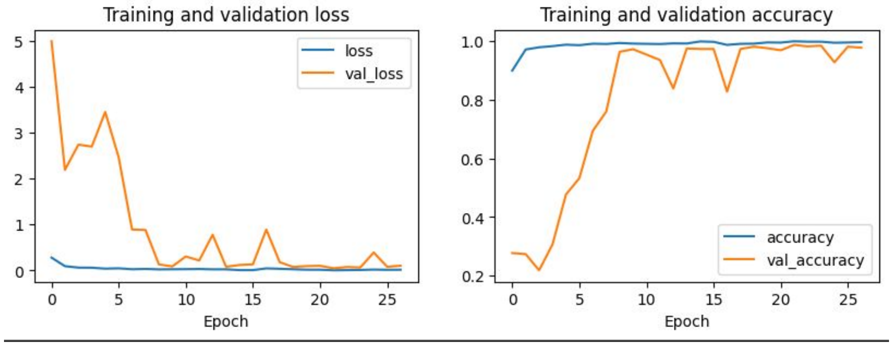
ResNet-50 Results

Test Accuracy	
Brain Tumor	98.93%
Leukemia	53.92%
Chest X-Ray	35.6%

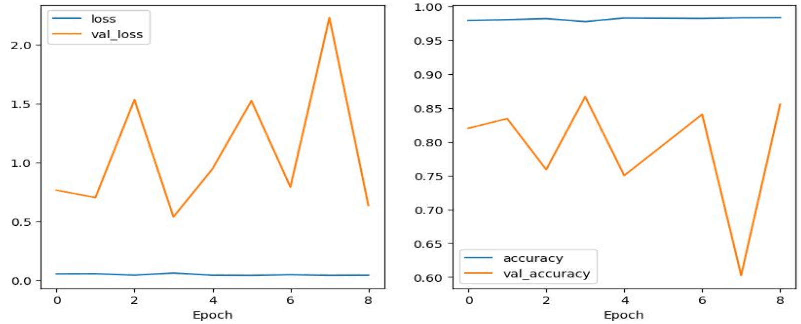
Leukemia



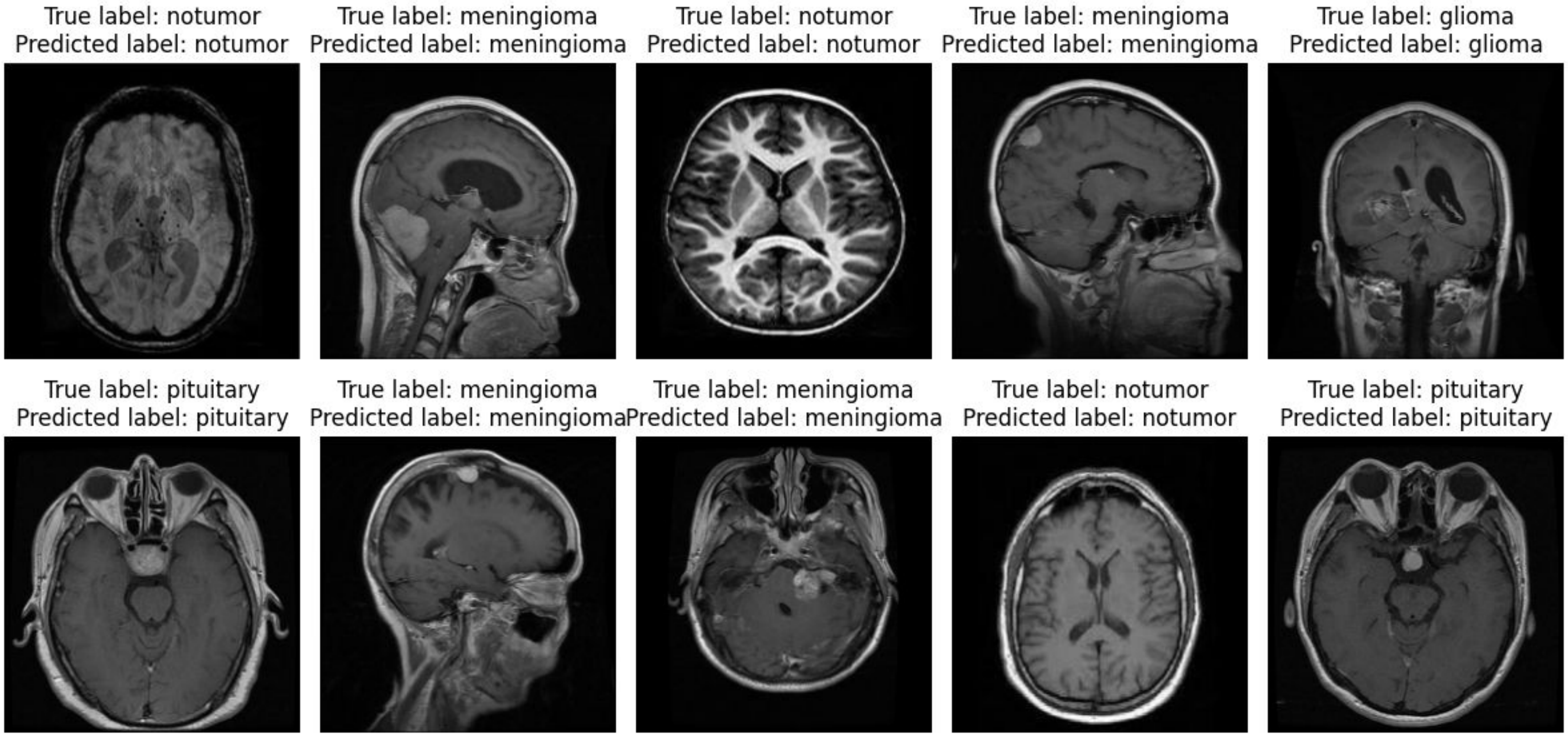
Brain Tumor



Chest X-ray



ResNet-50 Predictions



Challenges

- Existing research papers were using methods that are now deprecated (i.e. ImageDataGenerator), so it took some time to replicate their results using the newer methods (image_dataset_from_directory) and not all of the prior features were available in the newer methods
- Maxing out GPU consumption on GoogleColab
- Ran out of time to experiment with Albumentations library for image augmentation
- Figuring out which optimizer to use for each pre-trained model and data set
- ResNet50 performed poorly on complex architectures, & low information available on Brain MRI classification with 4 categories

Future Work

- Create a custom loss function to more heavily penalize false negatives, since the consequence of missing a condition is much worse than getting a false positive
- Experiment with Albumentations image augmentation library
- Customize the image augmentation pre-processing for each dataset (i.e. crop and apply brightness specifically for the leukemia data set)
- Experiment with combining all 3 pre-trained models
- Experiment with using un-frozen layers on other pre-trained models
- Could potentially leverage more advanced techniques (i.e. transformers)

Conclusion

- Could not generalize a model to work with all 3 data sets:
  - ResNet-50 performed the best on the Brain Tumor MRI data set
  - Inception-V3 performed the best on the Leukemia data set
  - VGG-16 performed the best on the Chest X-Ray data set
- Different pre-trained models could work better on different types of data sets
- Different architectures (including loss functions, optimizers, and activations) are necessary for binary classification vs. multi classification

References

- **Nickparvar, Msoud. (2021, September 24). *Brain Tumor MRI Dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset**
- **Khan, Md Saikat Islam et al. “Accurate brain tumor detection using deep convolutional neural network.” Computational and structural biotechnology journal vol. 20 4733-4745. 27 Aug. 2022, doi:10.1016/j.csbj.2022.08.039**
- **Sadad, Tariq, et al. "Brain tumor detection and multi‐classification using advanced deep learning techniques." Microscopy Research and Technique 84.6 (2021): 1296-1308.**
- **Baloni, Dev, and Shashi Kant Verma. "Detection of hydrocephalus using deep convolutional neural network in medical science." Multimedia Tools and Applications 81.12 (2022): 16171-16193.**
- **Elmorshedy, Abeer. (2022, December 10). *leukemia* [Data set]. Kaggle. https://www.kaggle.com/datasets/abeerelmorshedy/leukemia**
- **R Tamilarasi and S Gopinathan 2021 J. Phys.: Conf. Ser. 1964 072022**