UNIVERSITY VISVESVARAYA COLLEGE OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE

Application of Deep Learning in Medical Image Classification and Segmentation

TEAM MEMBERS
A Shriya (19GANSE001)
Harishankara H V (19GANSE017)
Kushal B (19GANSE024)
Mohammed Danish Rabbani (19GANSE028)

GUIDE
Dr. Thriveni J
Professor,
Dept of CSE, UVCE

Date-11-07-2023

Acknowledgements

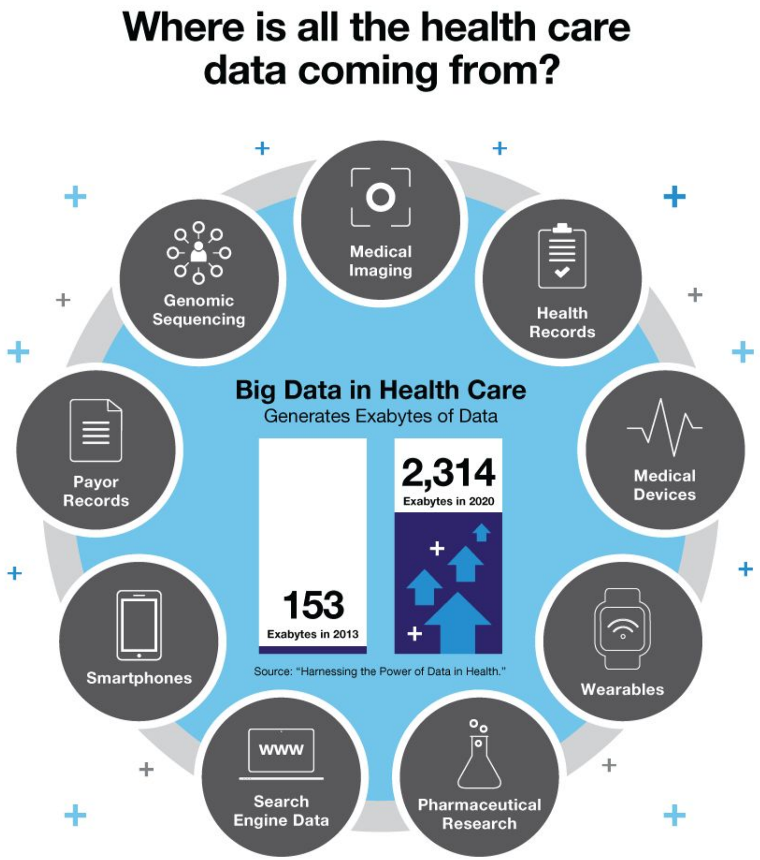
We would like to express our heartfelt gratitude to the following faculty:

- Dr Jayakara S M, Vice Chancellor, Bangalore University
- Dr H N Ramesh, Principal, UVCE
- Dr Vimala H S, Chairperson, Department of CSE
- Dr Thriveni J, Professor, Department of CSE

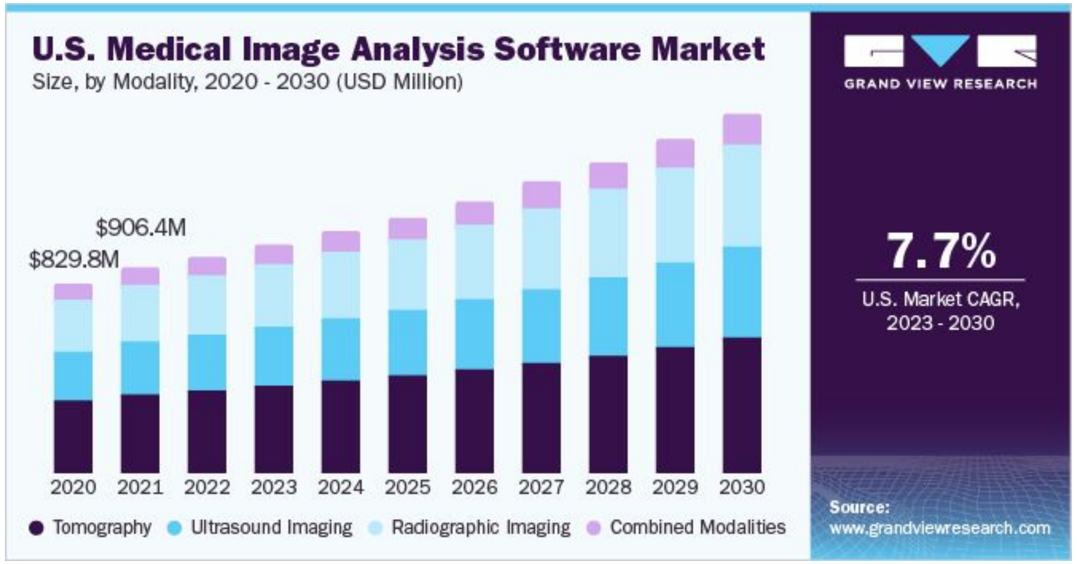
Contents

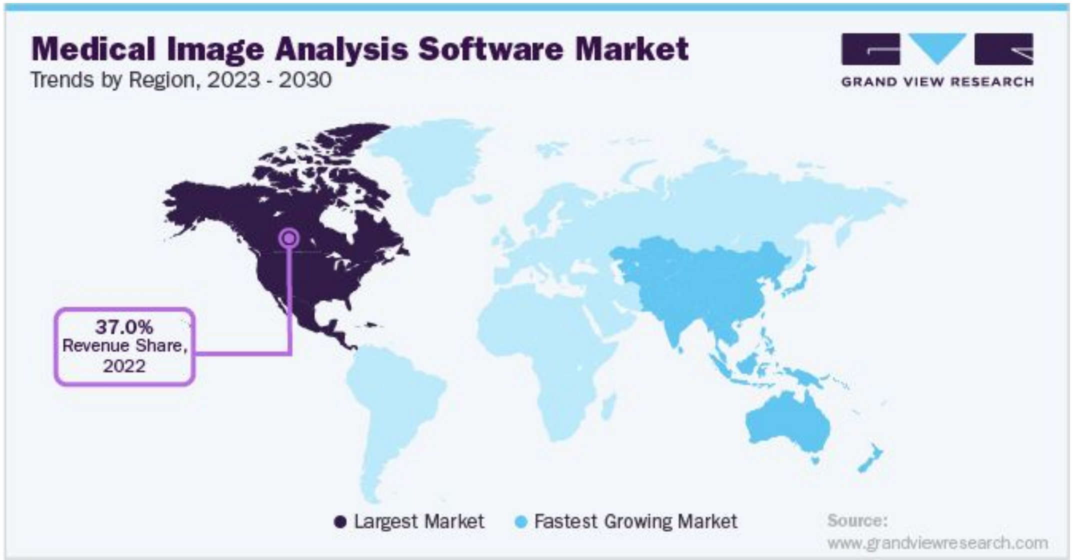
UNIVERSITY VISVESVARAYA
COLLEGE OF ENGINEERING

Background


Where is all the health care data coming from?
Medical Imaging
Health Records
Medical Devices
Wearables
Pharmaceutical Research
Search Engine Data
Smartphones
Payor Records
Genomic Sequencing
Big Data in Health Care
Generates Exabytes of Data
153
Exabytes in 2013
2,314
Exabytes in 2020
Source: "Harnessing the Power of Data in Health."

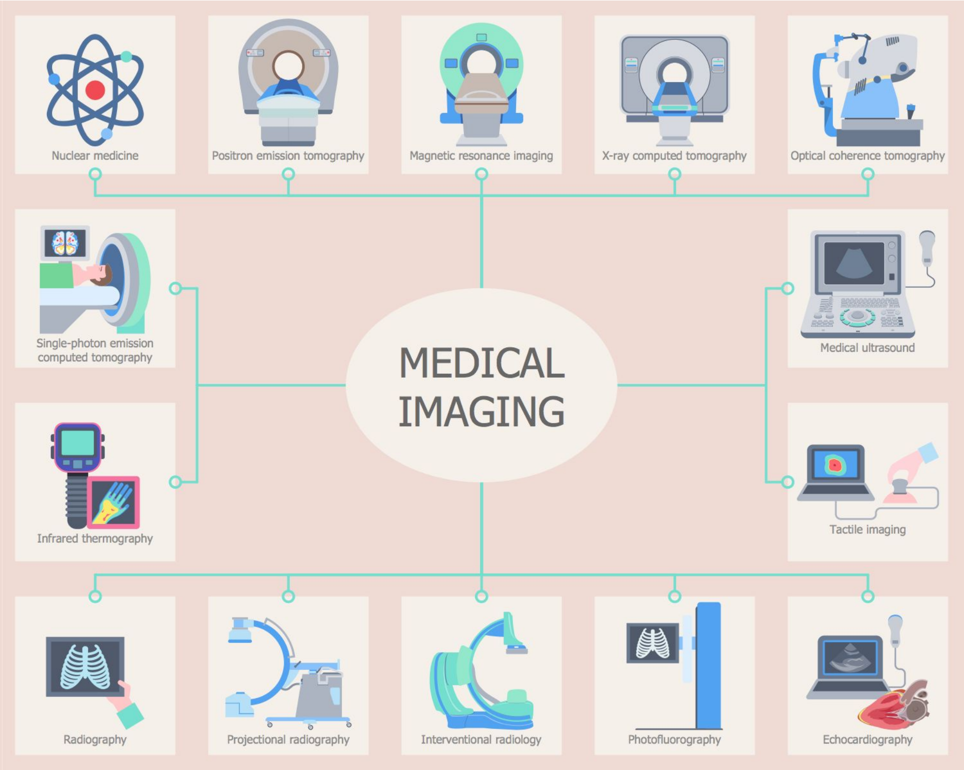
Background

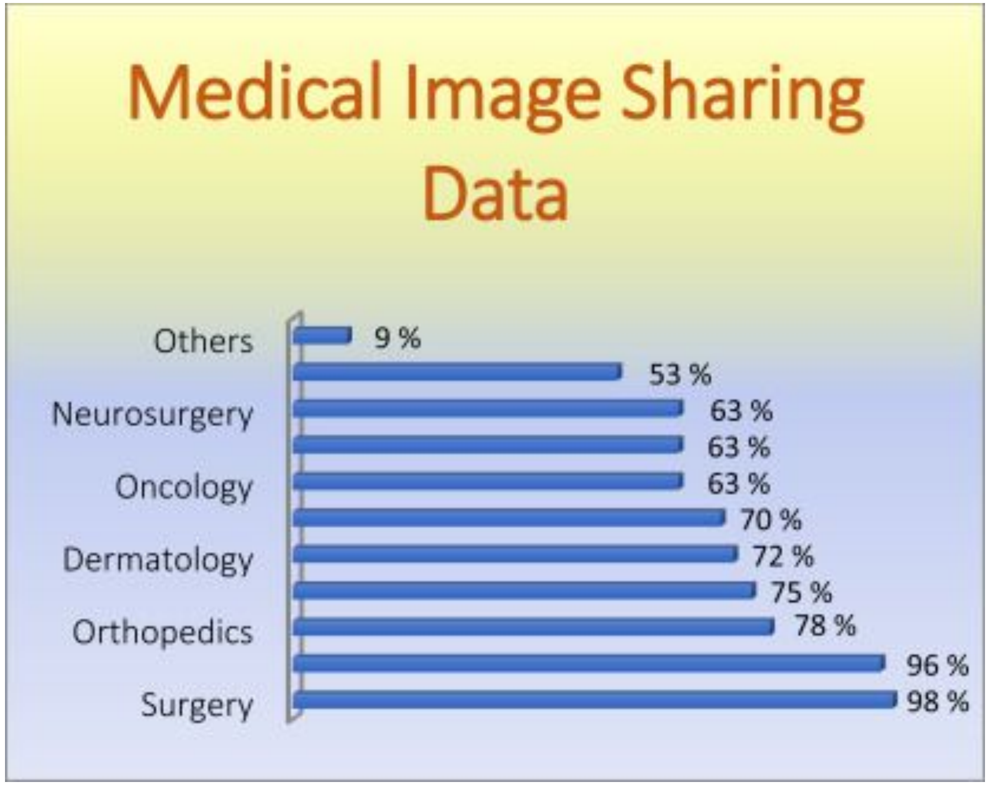




Source: https://www.grandviewresearch.com/industry-analysis/medical-image-analysis-software-market

Background


MEDICAL IMAGING
Nuclear medicine
Positron emission tomography
Magnetic resonance imaging
X-ray computed tomography
Optical coherence tomography
Single-photon emission computed tomography
Medical ultrasound
Infrared thermography
Tactile imaging
Radiography
Projectional radiography
Interventional radiology
Photofluorography
Echocardiography


Medical Image Sharing Data
Others
Neurosurgery
Oncology
Dermatology
Orthopedics
Surgery
9 %
53 %
63 %
63 %
63 %
70 %
72 %
75 %
78 %
96 %
98 %

Problem Statement

- **Background:** Medical image data account for the vast majority of medical data.
- **Problem:** How to utilise big medical data for best clinical practices to improve diagnosing techniques?
- **Solution:** Using intelligent imaging and deep learning in the field of big data analysis, combining the latest research progress and the work in the field, especially the classification and segmentation of medical images.
- **Outcome:** Early diagnosis of diseases

Project Focus

- **Medical Image Segmentation**
  - Brain Tumor Segmentation
- **Medical Image Classification**
  - Acute Lymphoblastic Leukemia
  - Brain Cancer
  - Breast Cancer
  - Cervical Cancer
  - Kidney Cancer
  - Lung Cancer
  - Colon Cancer
  - Lymphoma
  - Oral Cancer

Methodology

Transfer learning for Classification Models

- Transfer learning is a machine learning technique where knowledge gained from solving one problem is applied to a different but related problem.
- In transfer learning, a pre-trained model that has been trained on a large dataset and has learned useful patterns and representations is utilized as a starting point for solving a new task.

Methodology

Ensembling

- Machine learning technique where multiple models are combined to make predictions
- Using strengths of different models and reducing individual model errors for improved accuracy
- Reduces the risk of overfitting, enhances stability, and leverages the "wisdom of the crowd" principle

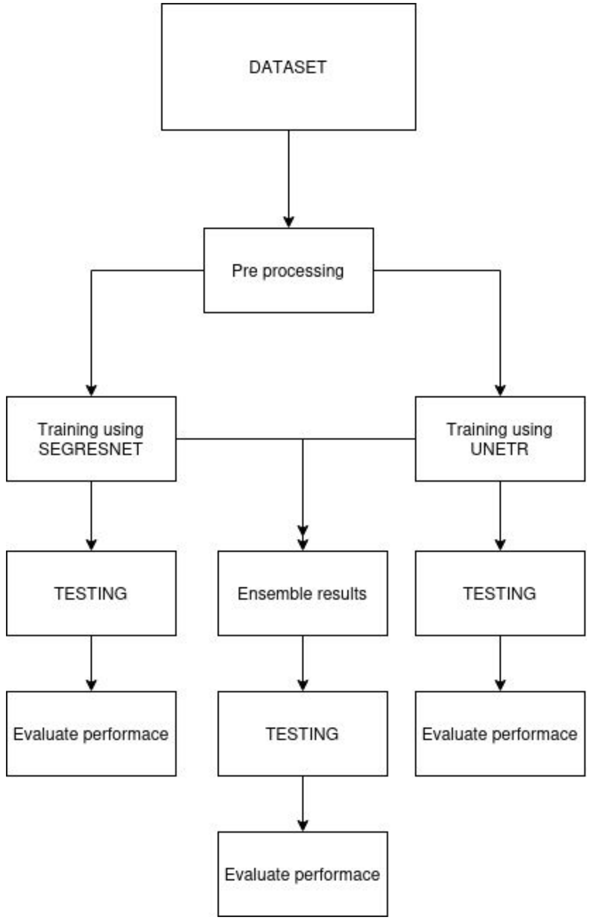
System Requirements

- **Hardware Requirements**
  - Intel i7 or higher
  - 16 GB RAM or higher
  - 12 GB CPU
- **Software Requirements**
  - Python 3.0 or higher
  - Jupyter
  - VSCode
  - Python Libraries: Pandas, Torchvision, Matplotlib, Numpy, Sklearn, Keras, TensorFlow, PyTorch, Monai

System Design

System Design for

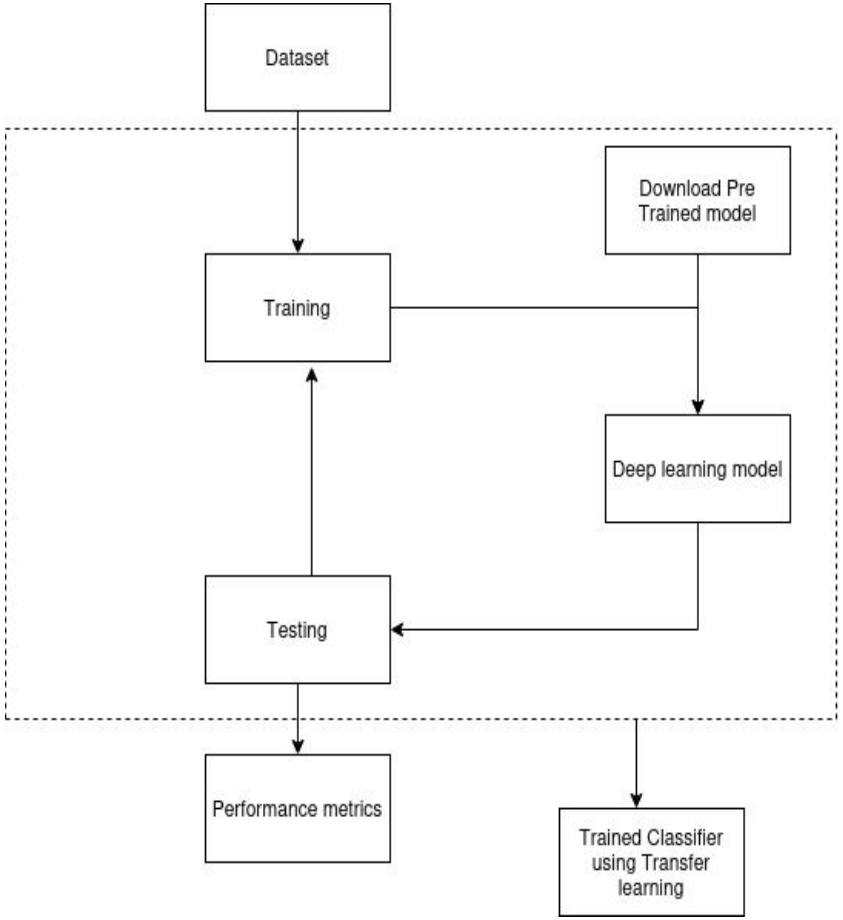
Brain Tumor Segmentation



System Design

System Design for

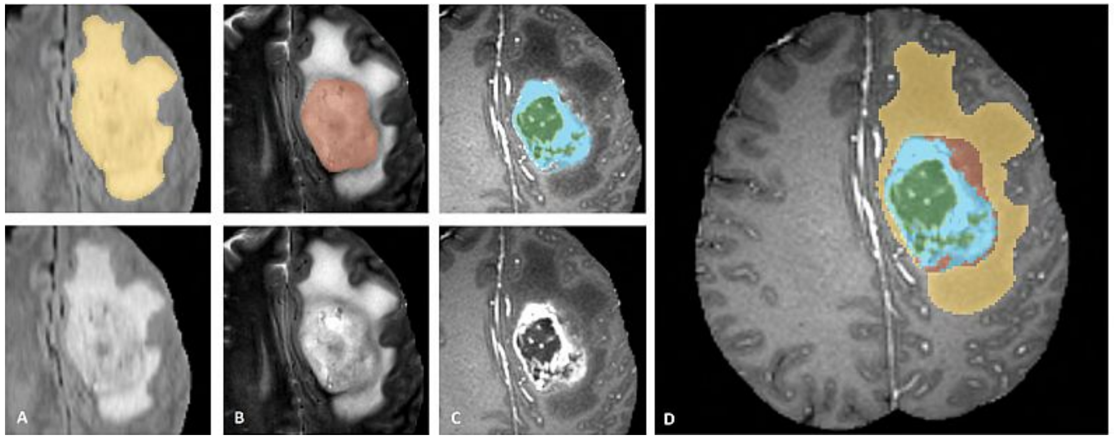
Classification



System Architecture

System Architecture for Brain Tumor Segmentation

- Dataset:
  - Size: 750 4D volumes (484 Training + 266 Testing) MRI images
  - Types: FLAIR, T1w, T1gd, T2w
  - Sources: BRATS 2016 and 2017



Glioma Sub-Regions

System Architecture

System Architecture for Brain Tumor Segmentation

Each MRI has 3 label images (Ground truth). The labels are:-

- Label 1 is the peritumoral edema
- Label 2 is the GD-enhancing tumor
- Label 3 is the necrotic and non-enhancing tumor core

These labels are converted to brats classes which are TC (Tumor core), WT (Whole tumor) and ET (Enhancing tumor).

System Architecture for Brain Tumor Segmentation

Label 2 and label 3 are merged to construct the tumor core.

Label 1, label 2 and label 3 are merged to construct Whole tumor.

Label 2 is enhancing tumors.

Each Image is cropped to a specific dimension.

System Architecture

System Architecture for Brain Tumor Segmentation

- Data Augmentation:
  - Transform the data by randomly adjusting the intensity and producing rotations and flips of the MRI image.
  - We apply a random (per channel) intensity shift (−0.1..0.1 of image std) and scale (0.9..1.1) on input image channels. We also apply a random axis mirror flip (for all 3 axes) with a probability 0.5.
- Modeling:
  - 2 architectures: SegResNet and UNETR

System Architecture

System Architecture for Classification

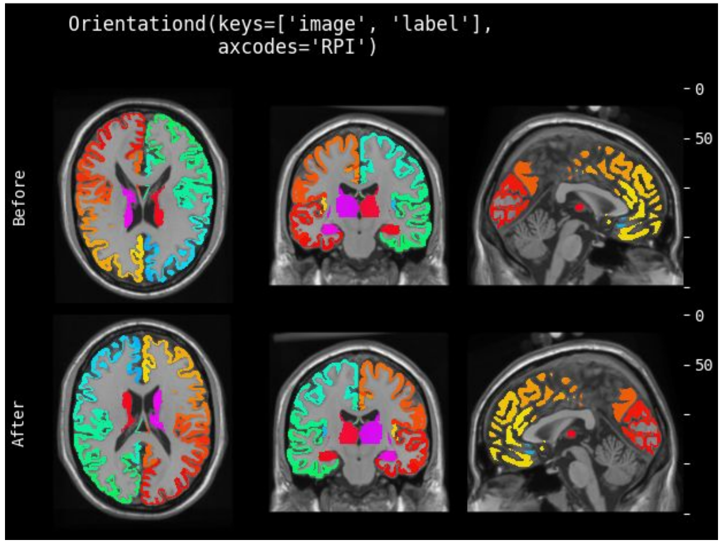
- Dataset:
  - Multi-cancer dataset from Kaggle
  - 5000 images in each sub-class
  - Image size 512x512
- Pre-processing:
  - Data Normalization
  - Image pre-processing
- Modeling:
  - CNN - ResNet, MobileNet, VGGNet, DenseNet

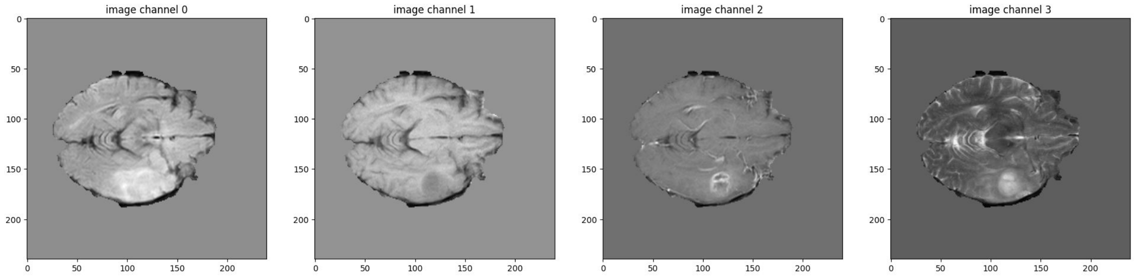
Cancer	Classes	Images
Acute Lymphoblastic Leukemia	4	20000
Brain Cancer	3	15000
Breast Cancer	2	10000
Cervical Cancer	5	25000
Kidney Cancer	2	10000
Lung and Colon Cancer	5	25000
Lymphoma	3	15000
Oral Cancer	2	10000

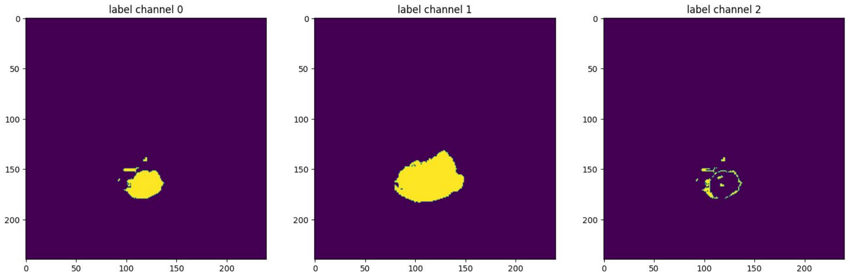
Implementation

Dataset

- Source: Medical Segmentation Decathlon (MSD)
- 80% training, 20% validation
- 484 mp-MRI images from patients diagnosed with either glioblastoma or lower-grade glioma



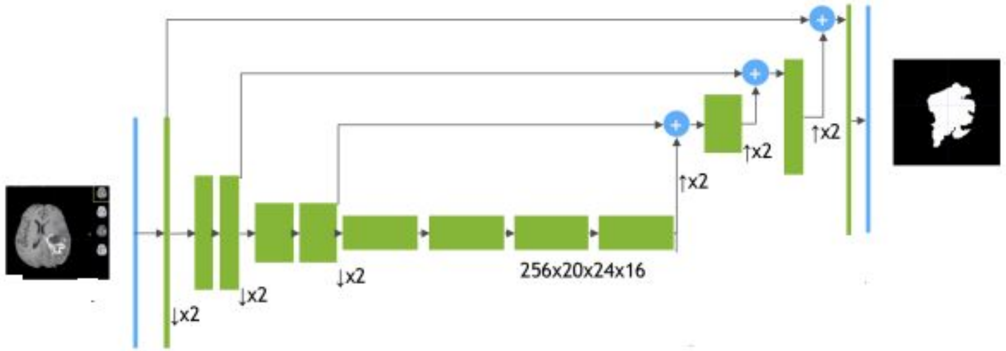




Implementation

SEGRESNET Architecture

- Based on ResNet, a CNN model
- SEGRESNET uses skip connections
- ResNet has powerful feature extraction capabilities
- Encoder-decoder based architecture



Schematic visualization of the network architecture

Architecture of each block

Implementation

Encoder

- The encoder part uses ResNet blocks, where each block consists of two convolutions with normalization and ReLU, followed by additive identity skip connection.
- For normalization, we use Group Normalization (GN), which shows better than BatchNorm performance when batch size is small (bath size of 1 in our case).
- We follow a common CNN approach to progressively downsize image dimensions by 2 and simultaneously increase feature size by 2. For downsizing we use strided convolutions.

Implementation

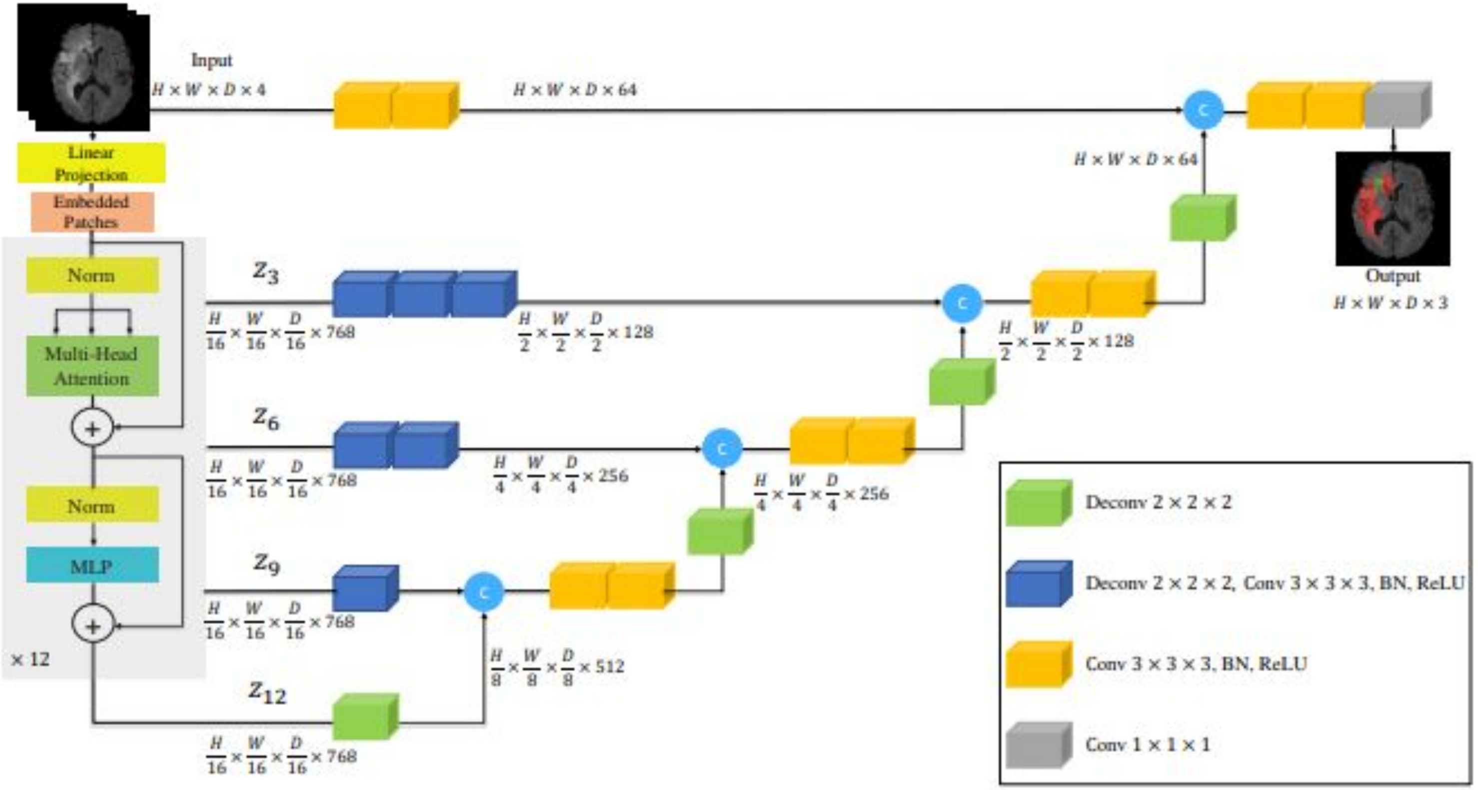
Decoder

- The decoder structure is similar to the encoder one, but with a single block per each spatial level.
- Each decoder level begins with upsizing: reducing the number of features by a factor of 2 (using 1x1x1 convolutions) and doubling the spatial dimension (using 3D bilinear upsampling), followed by an addition of encoder output of the equivalent spatial level.
- The end of the decoder has the same spatial size as the original image, and the number of features equal to the initial input feature size, followed by 1x1x1 convolution into 3 channels and a sigmoid function.

Implementation

UNETR Architecture

UNETR is a 3D medical image segmentation model that uses a transformer encoder and a CNN-based decoder to predict the segmentation mask.



Implementation

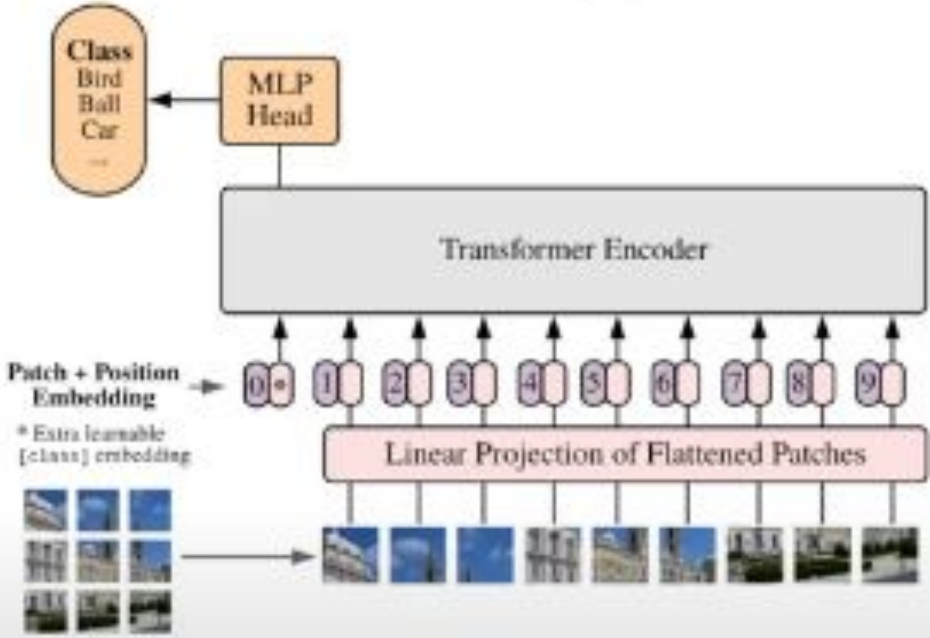
Key Features of UNETR?

1. It uses Vision Transformer as encoder to learn global contextual representations.

2. It uses a CNN Decoder to upsample the global representations and generate the final segmentation mask.

What is Vision Transformer?

Vision Transformer (ViT) is an architecture that is used for image recognition. It is based on the Transformer architecture, which was originally developed for natural language processing. ViTs have been shown to achieve state-of-the-art results on a variety of image recognition tasks, including ImageNet classification.



Implementation

UNIVERSITY VISVESVARAYA
COLLEGE OF ENGINEERING

Training

$$\mathbf{L}_{dice} = \frac{2 * \sum p_{true} * p_{pred}}{\sum p_{true}^2 + \sum p_{pred}^2 + \epsilon}$$

Loss function

$$\alpha = \alpha_0 * \left(1 - \frac{e}{N_e}\right)^{0.9}$$

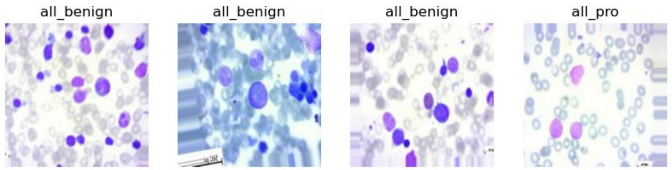
Optimizer function

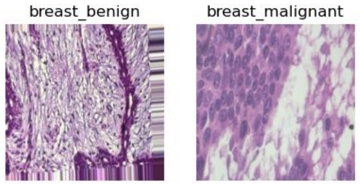
Ensemble

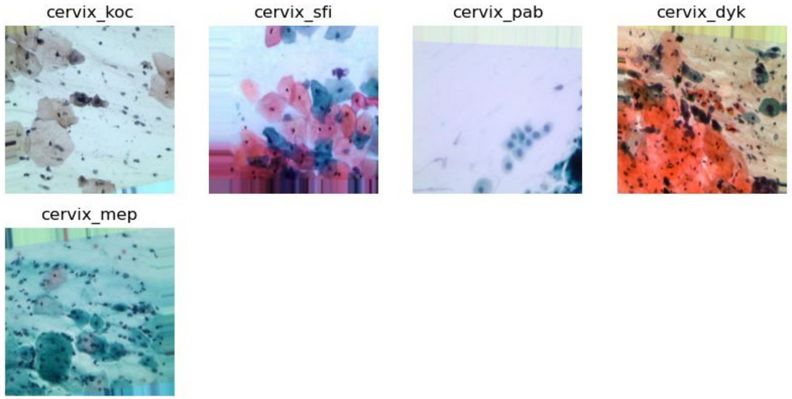
```
    def __call__(self, img: Union[Sequence[NdarrayOrTensor], NdarrayOrTensor]
        img_ = self.get_stacked_torch(img)
        if self.weights is not None:
            self.weights = self.weights.to(img_.device)
            shape = tuple(self.weights.shape)
            for _ in range(img_.ndimension() - self.weights.ndimension()):
                shape += (1,)
            weights = self.weights.reshape(*shape)

            img_ = img_ * weights / weights.mean(dim=0, keepdim=True)

        out_pt = torch.mean(img_, dim=0)
        return self.post_convert(out_pt, img)
```

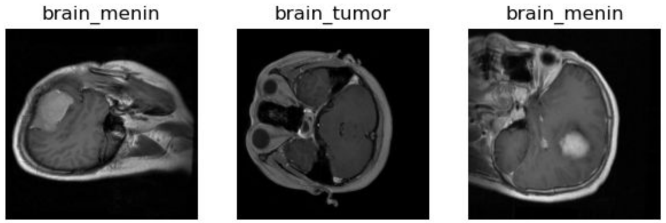
Implementation

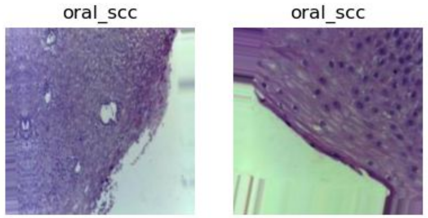
Implementation of Classification Models

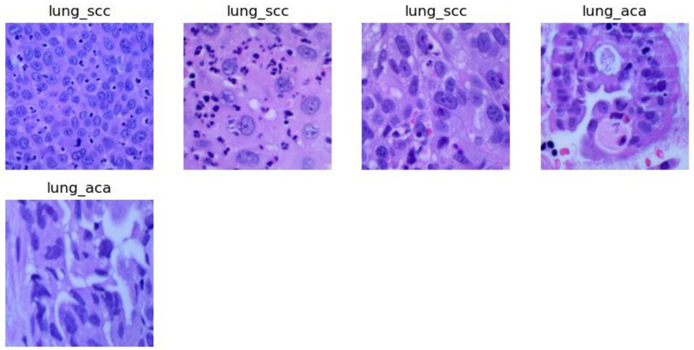


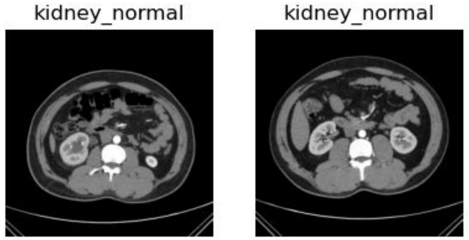


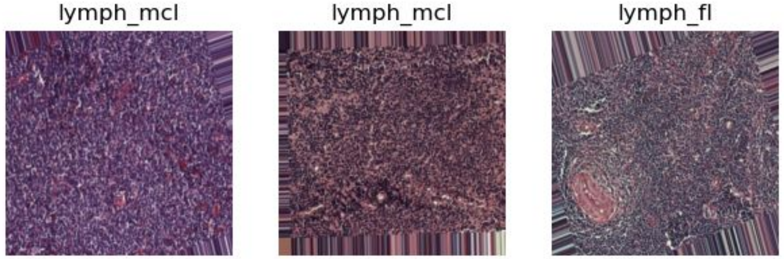












Neural Networks

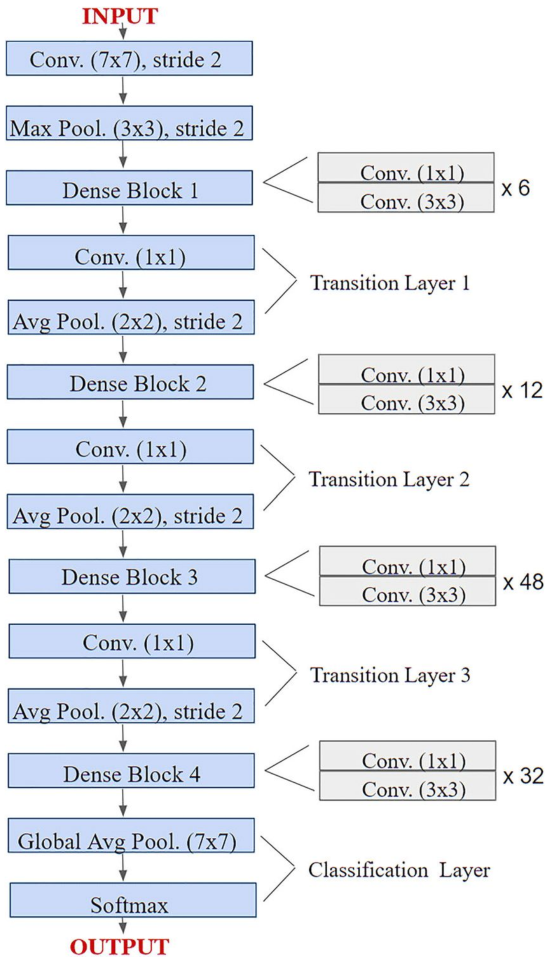
- What is a neuron?
- What is a neural network?
- What is a Convolutional Neural Network?

Implementation

UNIVERSITY VISVESVARAYA
COLLEGE OF ENGINEERING

DenseNet-201

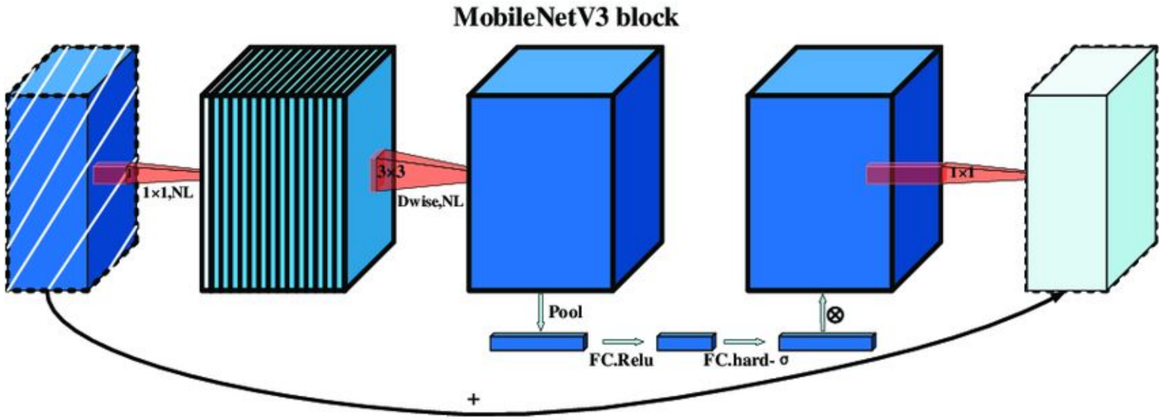
- Dense convolutional network
- Basic building block - dense block
- 201 layers
- Addresses vanishing gradients
- Strengthens feature propagation



Implementation

MobileNetv3Small

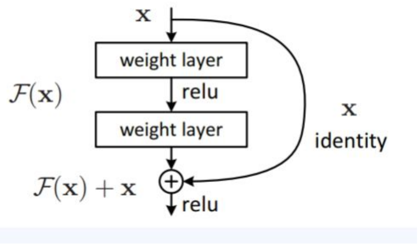
- Lightweight MobileNetv3
- Balance between model size and accuracy
- Depth-wise separable convolutions
- Inverted residuals
- Efficient use of activations



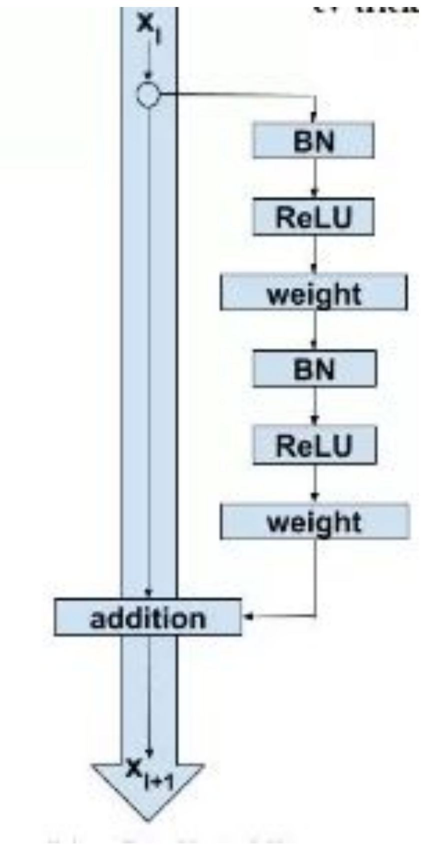
Implementation

ResNet50v2

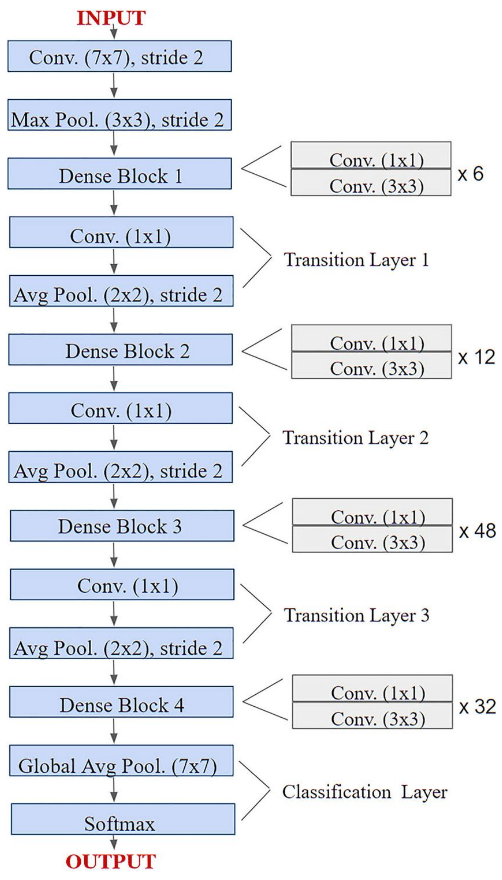
- Residual network architecture
- 50 layers
- Building blocks - residual blocks
- Residual connections
- Identity mapping
- Bottleneck structure
- Pre-activation
- Deeper architecture



Residual Blocks



RESNETv2

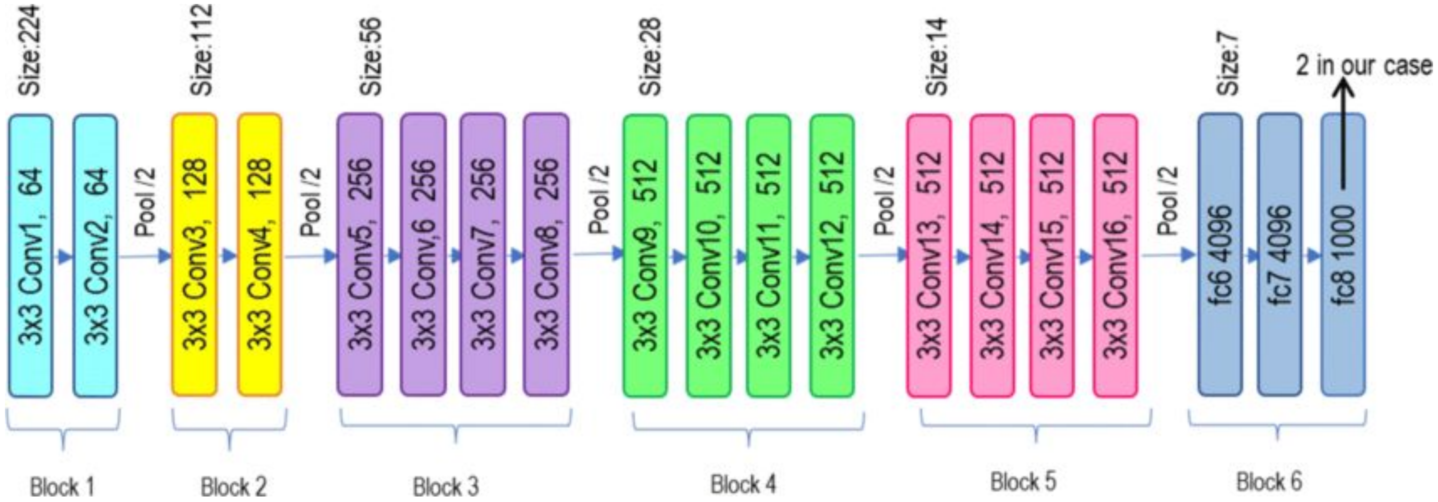


Architecture

Implementation

VGG19 Architecture

- Visual Geometry Group, 19 layers
- Convolutional layers, max pooling layers, fully connected layers, increasing depth
- Captures both local and global information from images



Implementation

Training

Optimiser - Adam Optimiser

- Adam + RMSprop
- Adaptive Learning Rates
- Momentum Optimization
- Bias Correction
- Weight Decay
- Computational Efficiency

Set learning rate = 0.01

Annealer - ReduceLROnPlateau

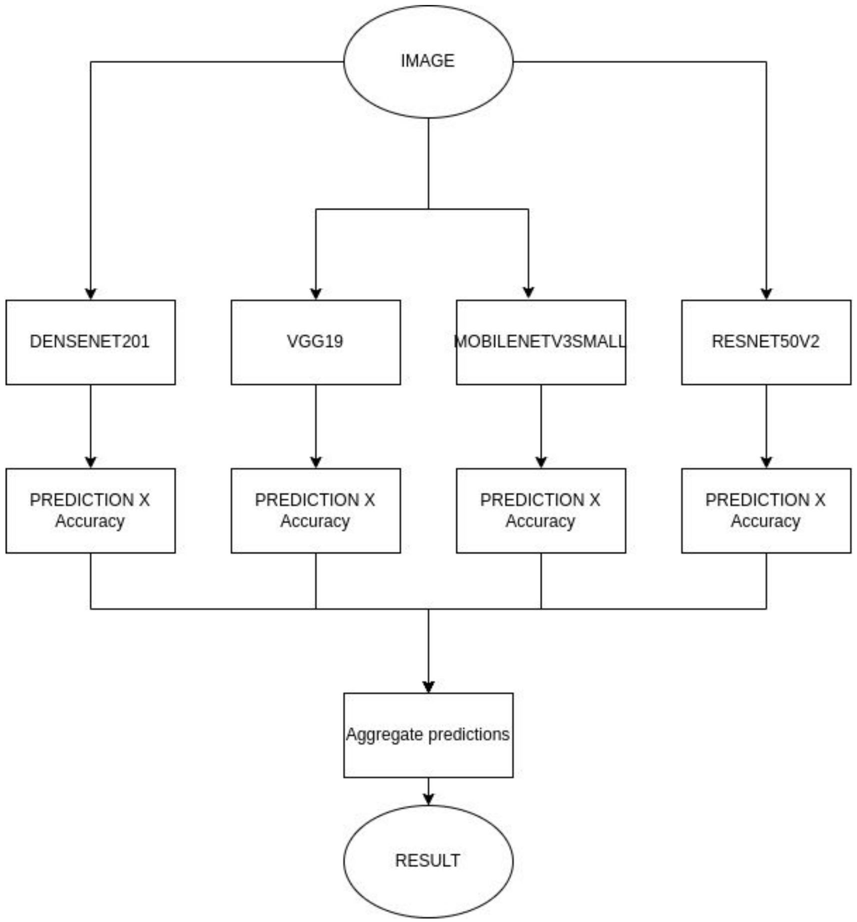
- ReduceLROnPlateau
- Monitor Metric
- Patience
- Factor
- Mode
- Threshold
- Cool-down

“ReduceLROnPlateau(monitor='val_accuracy', factor=0.5, patience=5, verbose=1, min_lr=1e-3)”

Implementation

Ensemble

- ML technique that involves combining multiple individual models, called base models or weak learners, to create a more robust and accurate prediction model.
- We have tried to ensemble the predictions of all the models based on the accuracy as a heuristic knowledge to weight the predictions of each model.



Performance Metrics

Metrics Used for Segmentation

- Dice score

Metrics Used for Classification

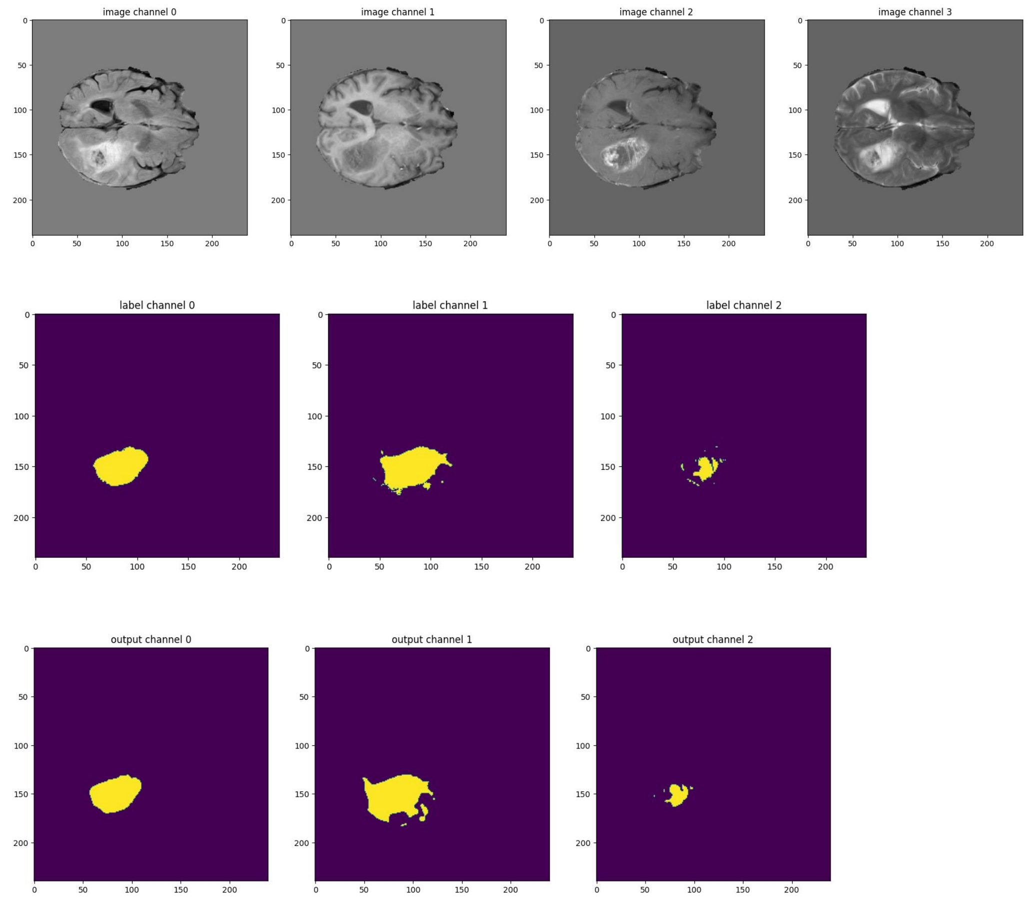
- Performance metrics
- Confusion matrix
- Accuracy
- Precision
- Recall or sensitivity
- F1 score

Performance Analysis

Result for Brain Tumor Segmentation - SEGRESNET

Dice Scores:

- Metric on original image spacing: 0.7625482082366943
- Metric for tumor Core: 0.8078
- Metric for Whole tumor: 0.9002
- Metric enhancing tumor: 0.5796
- Results achieved at 63 epochs



Original Image

Ground Truth

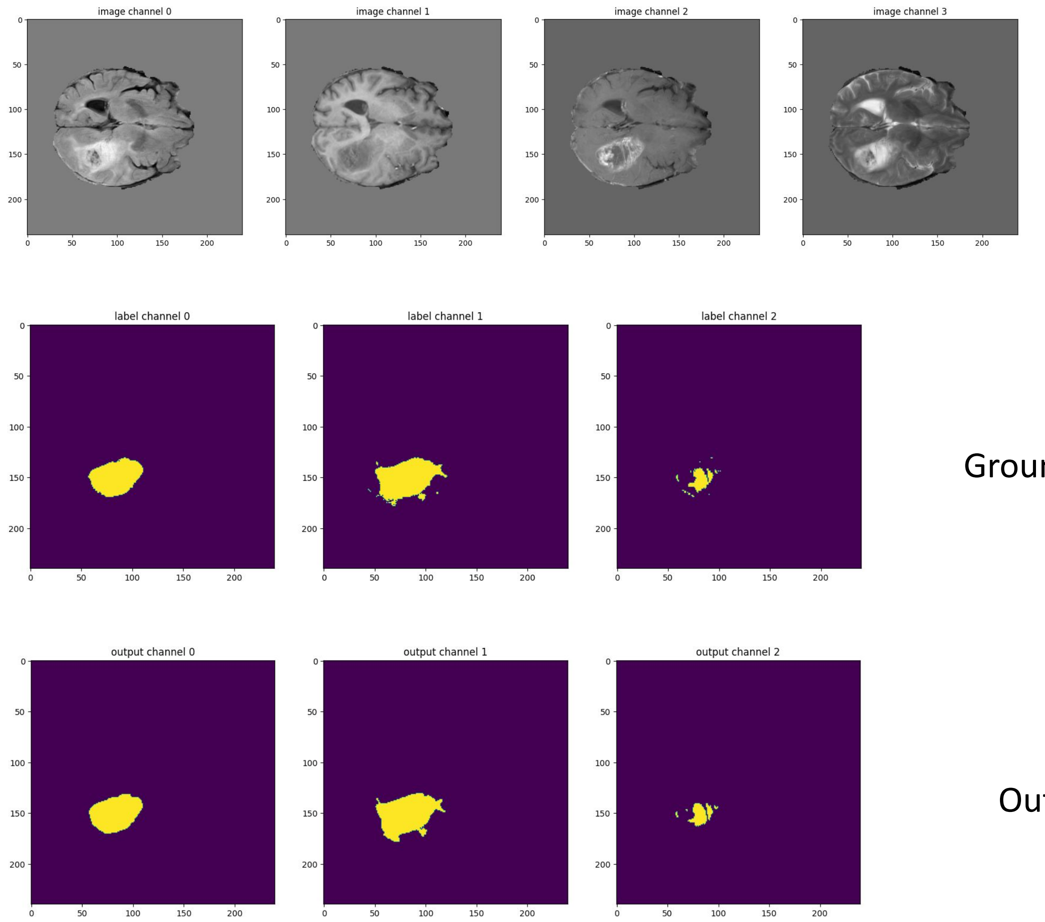
Output

Performance Analysis

Brain Tumor Segmentation - UNETR

Dice Scores:

- Metric on original image spacing: 0.737675666809082
- Metric tumor Core: 0.7786
- Metric Whole tumor: 0.8912
- Metric Enhancing tumor: 0.5432
- Results achieved at 47 epochs



Original Image

Ground Truth

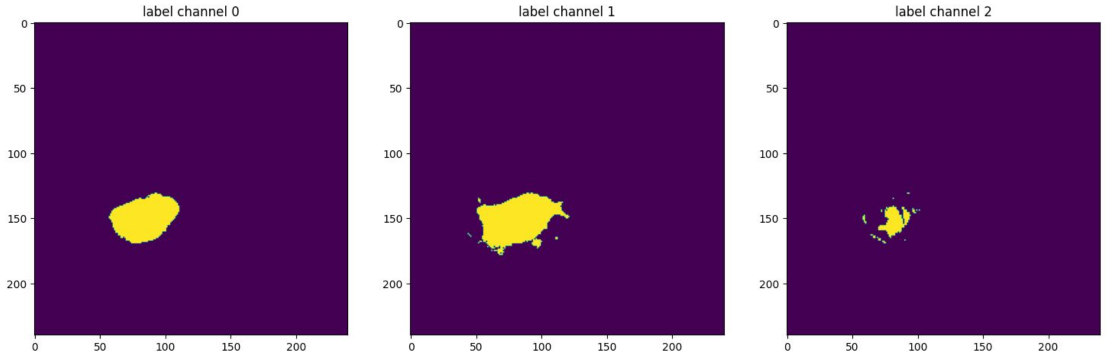
Output

Performance Analysis

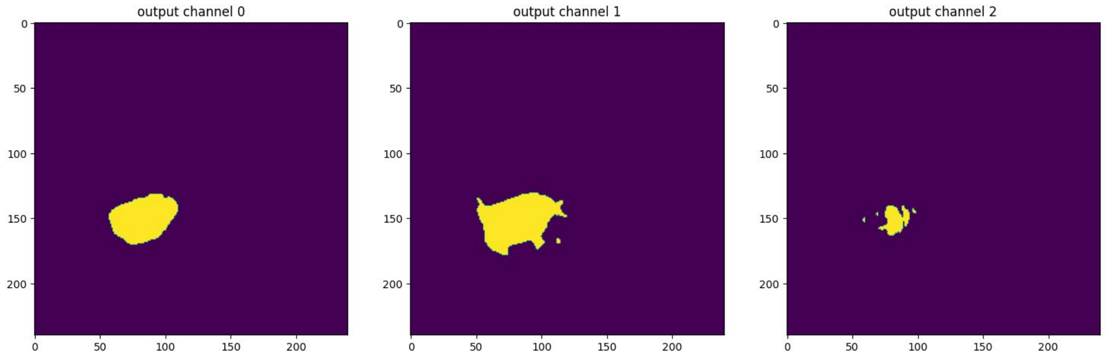
Brain Tumor Segmentation - Mean Ensemble

Dice Scores:

- Metric on original image spacing: 0.7645719647407532
- Metric Tumor Core: 0.8091
- Metric Whole tumor: 0.9032
- Metric enhancing tumor: 0.5814



Ground Truth

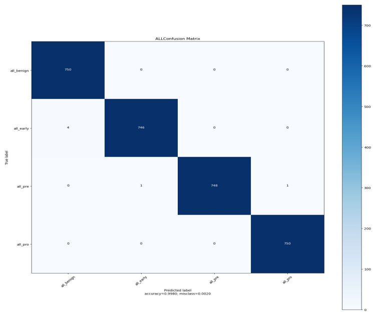


Model Output

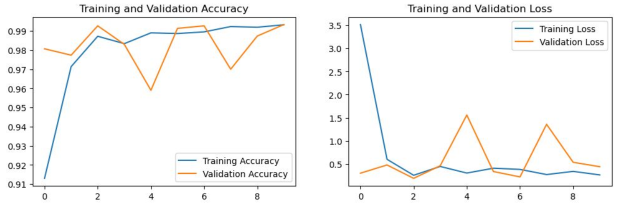
Performance Analysis

Acute Lymphoblastic Leukemia

Models	Recall	Precision	F1Score
DenseNet201	0.9903333333333333	0.9904421811723344	0.9903241375685572
MobileNetV3 Small	0.9916666666666667	0.991678857124651	0.9916619096443816
VGG19	0.9913333333333333	0.9913810676610213	0.99133643551295
ResNet50V2	0.7963333333333333	0.8631494378615919	0.7981504464349779
Ensemble	0.998	0.998006178550939	0.9979997690422237

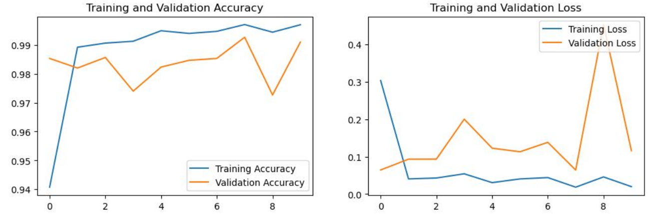


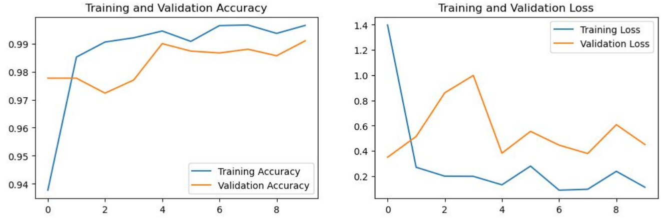
Ensemble



DenseNet201

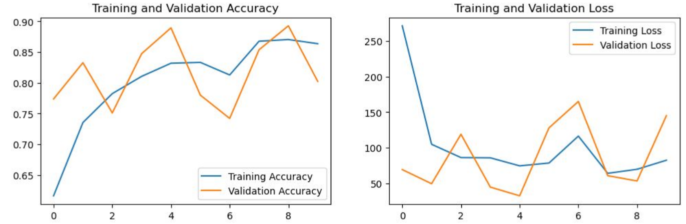
MobileNetv3Small





VGG19

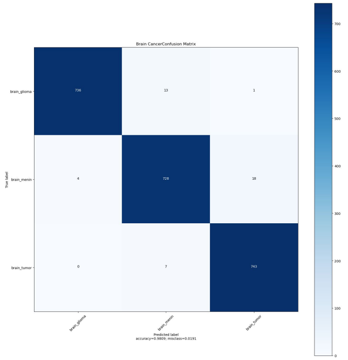
ResNet50V2



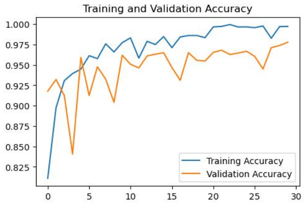
Performance Analysis

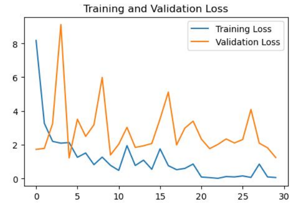
Brain Cancer

Models	Accuracy	Recall	Precision	F1Score
DenseNet201	0.9595555555555556	0.9595555555555556	0.959682605299268	0.9595922637062334
MobileNetV3Small	0.9697777777777777	0.9697777777777777	0.9699118252039262	0.9697824194666587
VGG19	0.9684444444444444	0.9684444444444444	0.9688725780364567	0.9683784986919409
ResNet50V2	0.8742222222222222	0.8742222222222222	0.8893087706148176	0.8757004946044663
Ensemble	0.9808888888888889	0.9808888888888889	0.980974081159352	0.9808954375713036



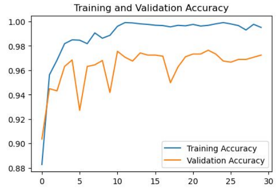
Ensemble

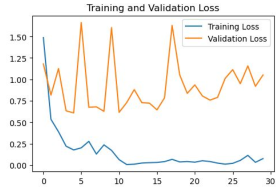


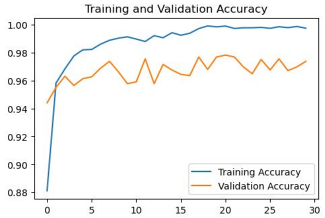


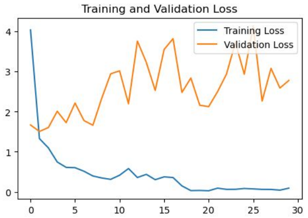
DenseNet201

MobileNetv3Small



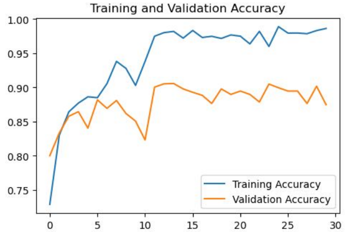


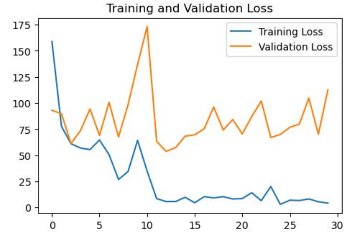




VGG19

ResNet50V2

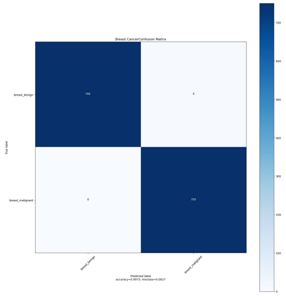




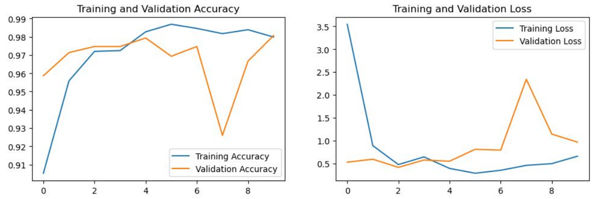
Performance Analysis

Breast Cancer

Models	Accuracy	Recall	Precision	F1Score
DenseNet201	0.98	0.98	0.9806699470907005	0.9799930286819941
MobileNetV3Small	0.9926666666666667	0.9926666666666667	0.9926675425200756	0.992666663407406
VGG19	0.9913333333333333	0.9913333333333333	0.9914040955230887	0.991333021322101
ResNet50V2	0.7926666666666666	0.7926666666666666	0.8479715662049592	0.784087653470172
Ensemble	0.9973333333333333	0.9973333333333333	0.9973474801061007	0.9973333143702354

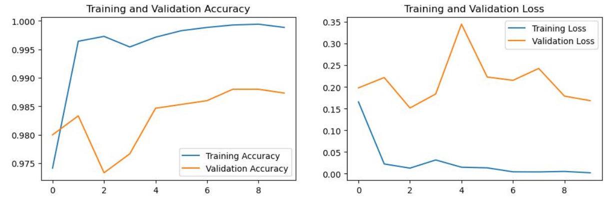


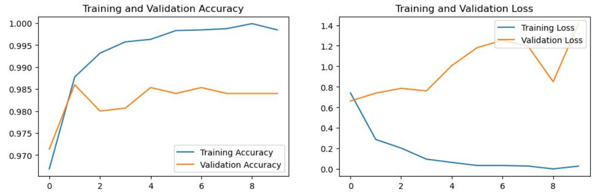
Ensemble



DenseNet201

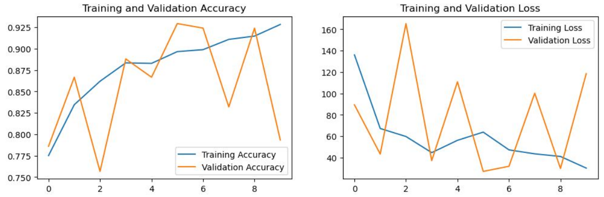
MobileNetv3Small





VGG19

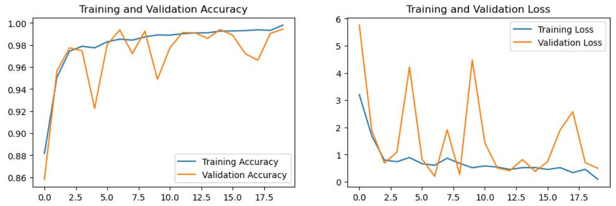
ResNet50V2



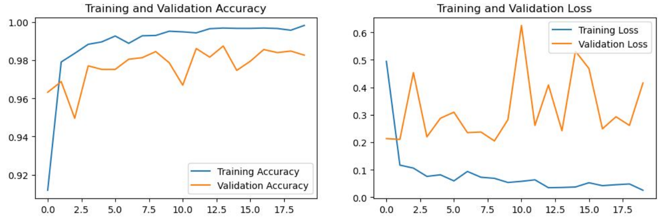
Performance Analysis

Cervical Cancer

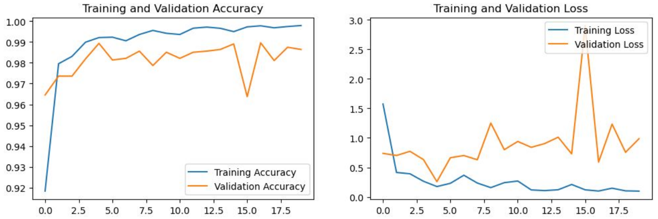
Models	Accuracy	Recall	Precision	F1Score
DenseNet201	0.9944	0.9944	0.994463029969946	0.9944038407241459
MobileNetV3Small	0.9826666666666667	0.9826666666666667	0.9829379948701118	0.9826295355747524
VGG19	0.9890666666666666	0.9890666666666666	0.9891694364407169	0.9890777752461215
ResNet50V2	0.8512	0.8512	0.881911965145657	0.847642216679388
Ensemble	0.9981333333333333	0.9981333333333333	0.9981358160750206	0.9981333307240217



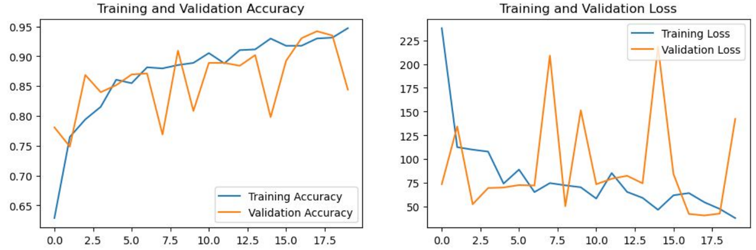
DenseNet201



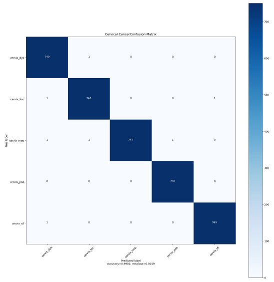
MobileNetv3Small



VGG19



ResNet50V2

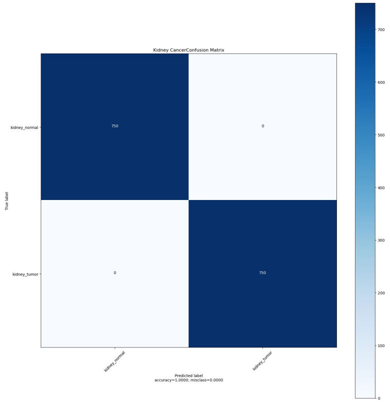


Ensemble

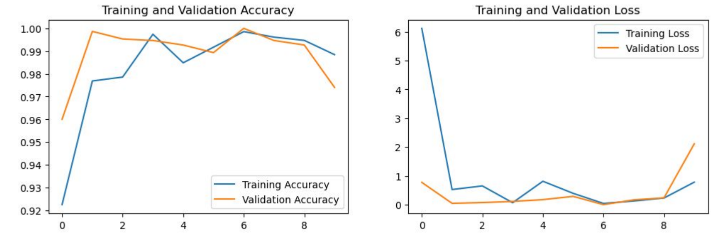
Performance Analysis

Kidney Cancer

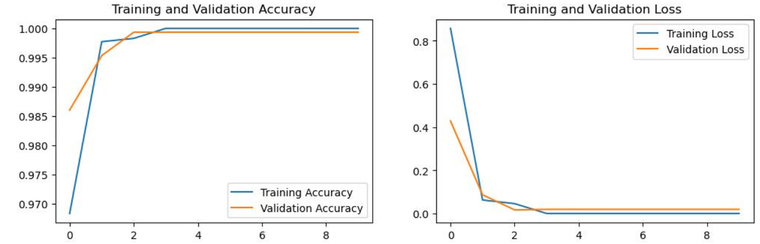
Models	Recall	Precision	F1Score
DenseNet201	0.9426666666666667	0.997179125528914	0.9691569568197395
MobileNetV3Small	1.0	1.0	1.0
VGG19	1.0	1.0	1.0
ResNet50V2	0.9933333333333333	1.0	0.9966555183946488
Ensemble	1.0	1.0	1.0



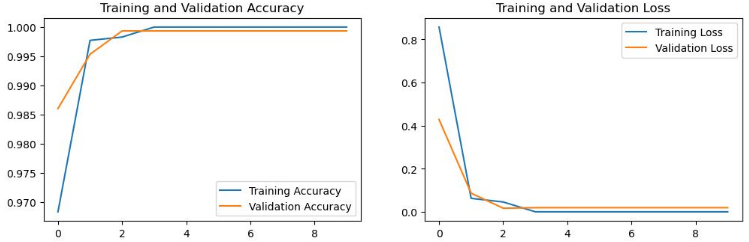
Ensemble



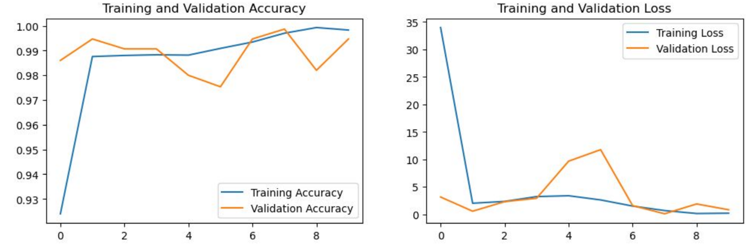
DenseNet201



MobileNetv3Small



VGG19

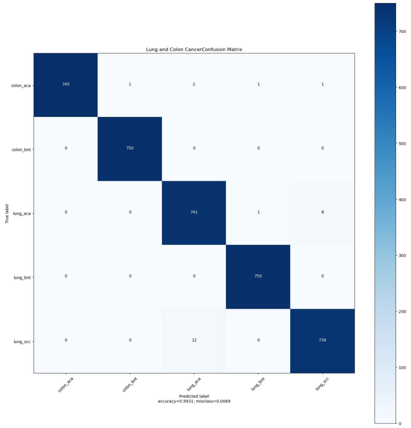


ResNet50V2

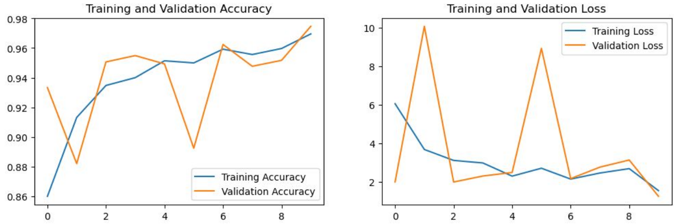
Performance Analysis

Lung and Colon Cancer

Models	Recall	Precision	F1Score
DenseNet201	0.9728	0.9729875598804594	0.972848129781705
MobileNetV3Small	0.9864	0.9864941910877201	0.9864197297322221
VGG19	0.9733333333333334	0.974243236543145	0.9732906860342894
ResNet50V2	0.868	0.8908159830262765	0.8611486751050
Ensemble	0.9930666666666667	0.9930835256960889	0.9930694579629511

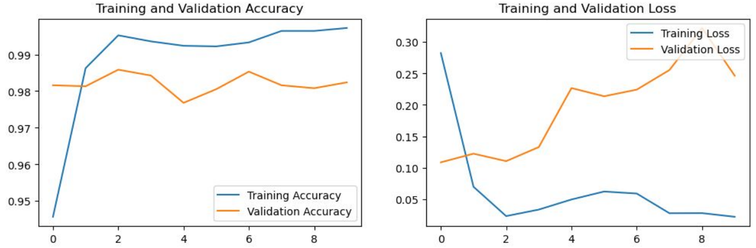


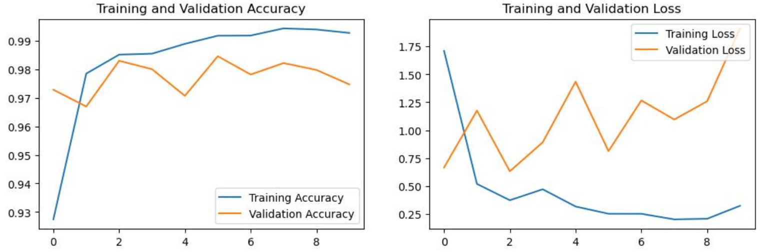
Ensemble



DenseNet201

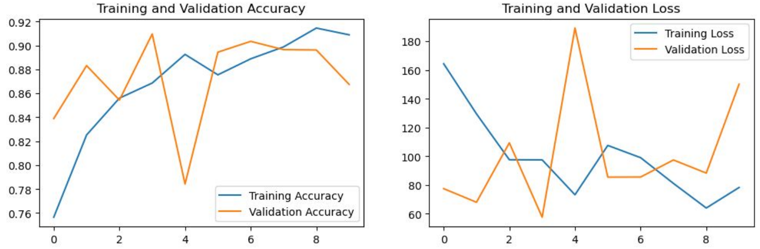
MobileNetv3Small





VGG19

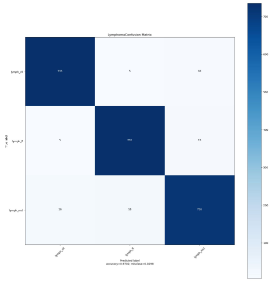
ResNet50V2



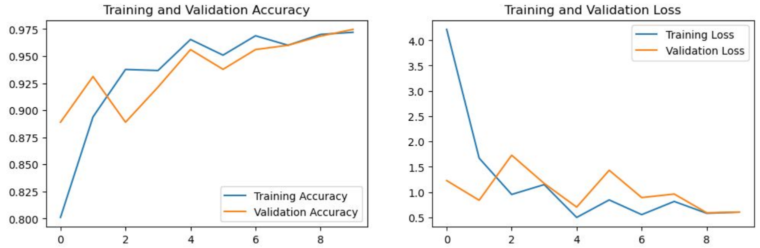
Performance Analysis

Lymphoma

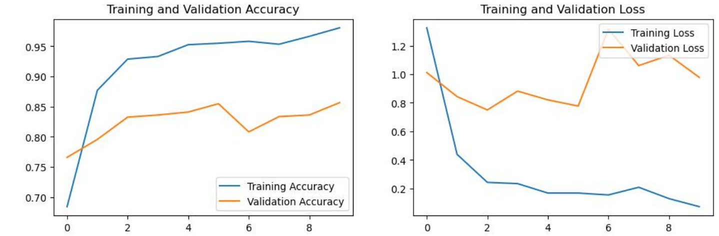
Models	Accuracy	Recall	Precision	F1Score
DenseNet201	0.9688888888888889	0.9688888888888889	0.9689020280935331	0.9688874465479973
MobileNetV3Small	0.8613333333333333	0.8613333333333333	0.863749712001629	0.8618748161763887
VGG19	0.8417777777777777	0.8417777777777777	0.8426514454459834	0.8415871408756553
ResNet50V2	0.6862222222222222	0.6862222222222222	0.778109181931267	0.6394392144701045
Ensemble	0.9702222222222222	0.9702222222222222	0.9702118355619148	0.9701907890979771



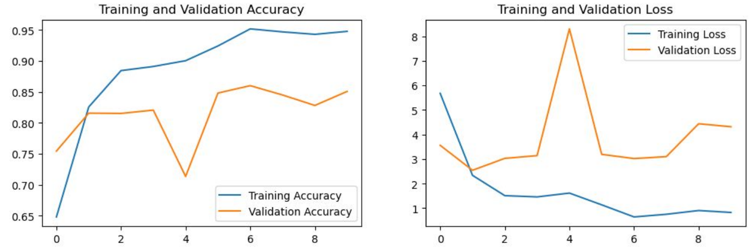
Ensemble



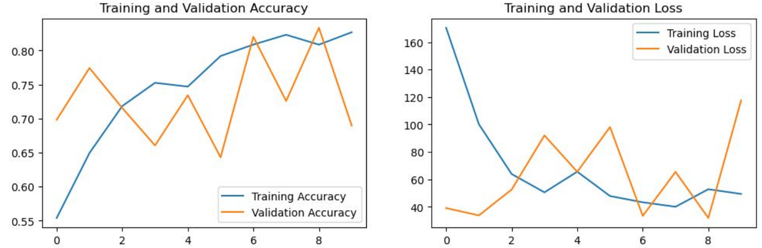
DenseNet201



MobileNetv3Small



VGG19



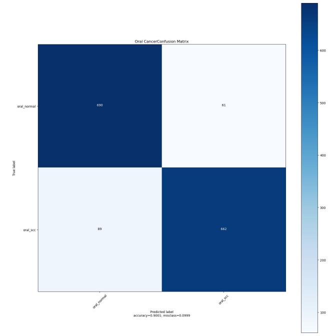
ResNet50V2

Performance Analysis

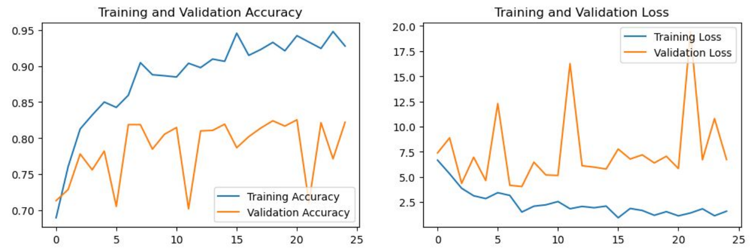
UNIVERSITY VISVESVARAYA COLLEGE OF ENGINEERING

Oral Cancer

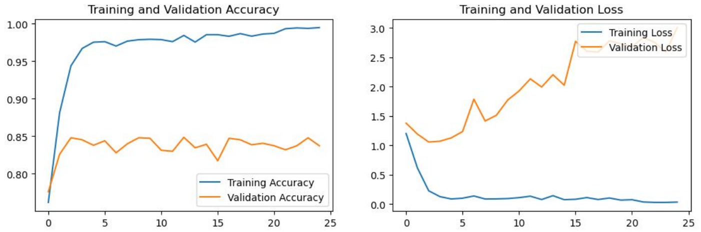
Models	Recall	Precision	F1Score
DenseNet201	0.8362183754993342	0.8406859442513094	0.835679673824873
MobileNetV3Small	0.8495339547270306	0.8495364337044347	0.8495336879432624
VGG19	0.8302263648468708	0.8302269503546099	0.8302262895927
ResNet50V2	0.8055925432756325	0.8071807301923951	0.8053409365280007
Ensemble	0.9001331557922769	0.9006901425205559	0.9000984382898342



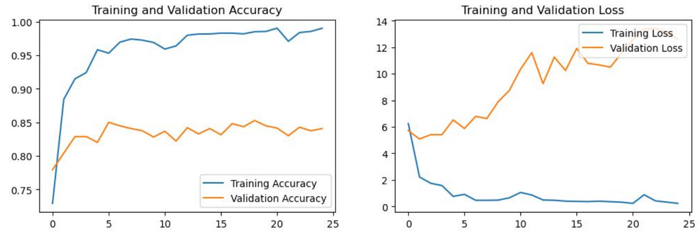
Ensemble



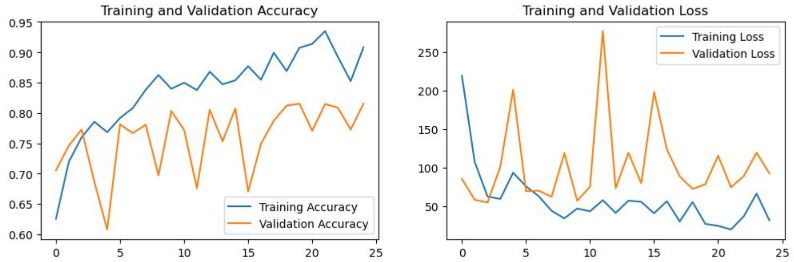
DenseNet201



MobileNetv3Small



VGG19



ResNet50V2

Conclusion and Future Work

- Managed to build 2 models and tried to ensemble it to segment brain tumors and achieved desirable results
- Ensemble-based approaches offer significant advantages in medical image classification and segmentation. They enhance the performance, robustness, and generalization capabilities of the system, enabling more accurate diagnoses and treatment decisions. Leveraging the power of ensembles can lead to improved healthcare outcomes and contribute to advancements in medical imaging technology.
- Conclusion: The detection of cancer using machine learning techniques offers several advantages that can significantly impact healthcare and improve patient outcomes. Traditional cancer diagnosis often requires extensive manual analysis and interpretation of medical images and biopsy samples, which can be time-consuming. Machine learning algorithms can automate and expedite the diagnostic process, analyzing images and data in a fraction of the time, thereby reducing waiting times for patients and enabling faster treatment initiation.

Conclusion and Future Work

- Scope for improvement:
  - Bigger dataset
  - Cross-validation ensemble
  - Swin-UNETR Architecture
  - Better validation data sourced through key radiologists
- Future Work:
  - More and potentially better models
  - Creating bigger ensemble

References

UNIVERSITY VISVESVARAYA
COLLEGE OF ENGINEERING

[1] Abadi, Martın et al., “Tensorflow: Large-scale machine learning on heterogeneous distributed systems” (2016).

[2] Abraham, Nabila and Naimul Mefraz Khan, “A novel focal tversky loss function with improved attention u-net for lesion segmentation” (2019).

[3] Anithadevi, D and K Perumal, “A hybrid approach based segmentation technique for brain tumor in MRI Images” (2016).

[4] Hatamizadeh A, Tang Y, “Unetr: Transformers for 3d medical image segmentation” (pp. 574-584), (2022).

[5] Myronenko A, “3D MRI brain tumor segmentation using autoencoder regularization” (2019).

[6] Gu, Shanqing, Manisha Pednekar and Robert Slater, “Improve Image Classi- fication Using Data Augmentation and Neural Networks” (2019).

[7] Salçin, Kerem et al, “Detection and classification of brain tumors from MRI images using faster R-CNN” (2019).

[8] Jonas Prellberg, Oliver Kramer, “Acute Lymphoblastic Leukemia Detection from Microscopic Images Using Weighted Ensemble of Convolutional Neural Networks” (2019).

[9] Marco Aurélio Granero, Cristhian Xavier Hernández, Marcos Eduardo Valle, “Quaternion-Valued Convolutional Neural Network Applied for Acute Lymphoblastic Leukemia Diagnosis” (2021).

[10] Md Zahangir Alom, Chris Yakopcic, Tarek M. Taha, Vijayan K. Asari, “Breast Cancer Classification from Histopathological Images with Inception Recurrent Residual Convolutional Neural Network” (2018).

[11] Aditya Golatkar, Deepak Anand, Amit Sethi, “Classification of Breast Cancer Histology using Deep Learning” (2017).

[12] Rishav Pramanik, Momojit Biswas, et al, “A fuzzy distance-based ensemble of deep models for cervical cancer detection” (2019).

[13] Nanditha B R, Geetha Kiran A Sanathkumar M P, “Oral Cancer Detection using Machine Learning and Deep Learning Techniques” (2021).

[14] Panwar, S.A., Munot, M.V., Gawande, S. and Deshpande, “A reliable and an efficient approach for diagnosis of brain tumor using transfer learning” (pp.283-294), (2021).

[15] Ullah, N., Khan, J.A., Khan, M.S., Khan, W., Hassan, I., Obayya, M., Negm, N. and Salama, A.S, “An Effective Approach to Detect and Identify Brain Tumors Using Transfer Learning. Applied Sciences” (2022).

[16] Dalia Alzu’bi,Malak Abdullah, et.al, “Kidney Tumor Detection and Classification Based on Deep Learning Approaches: A New Dataset in CT Scans” (2022).

[17] Md. Alamin Talukder, Md. Manowarul Islam, et.al, “Machine Learning-based Lung and Colon Cancer Detection using Deep Feature Extraction and Ensemble Learning” (2022).

[18] Satvik Garg, Somya Garg, “Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps” (2021).

[19] Maria Frucci, Daniel Riccio, et.al, “A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images” (2021).

[20] Zeyad Ghaleb Al-Mekhlafi, Ebrahim Mohammed Senan, et.al, “Diagnosis of Histopathological Images to Distinguish Types of Malignant Lymphomas Using Hybrid Techniques Based on Fusion Features” (2022).

THANK YOU