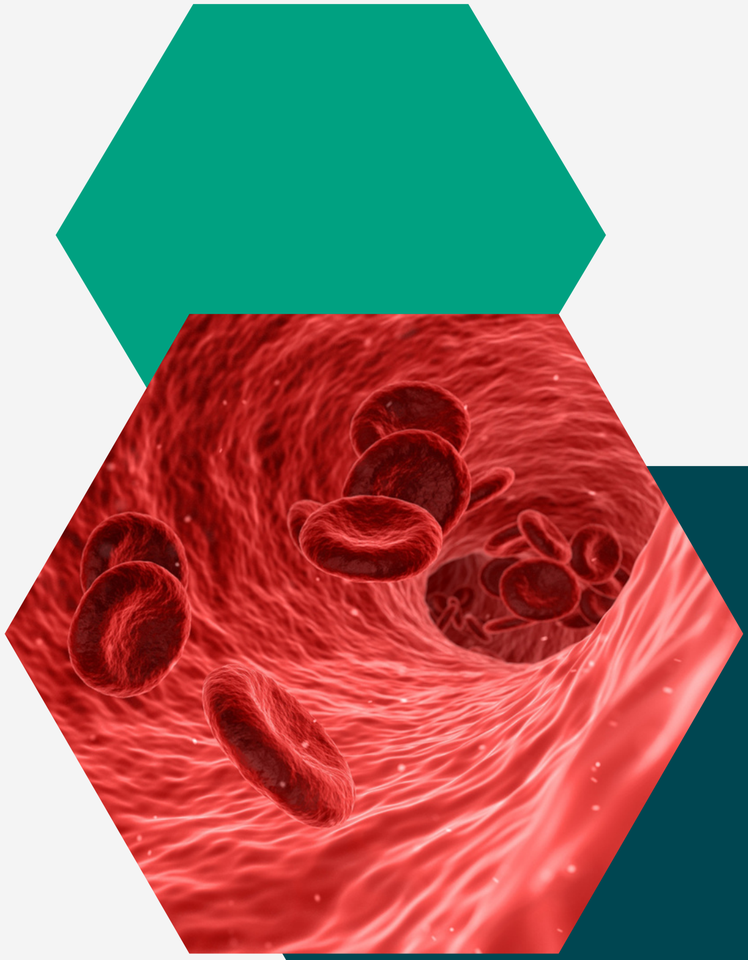
Methods and Apparatus for Detection of Types of Leukemia using Ensemble Learning

Contents

- Problem Definition
- Motivation
- Objective
- Literature Review
- Dataset
- Proposed System
- Results and Testing
- Event Participation and Paper Publication Details
- Conclusion
- References

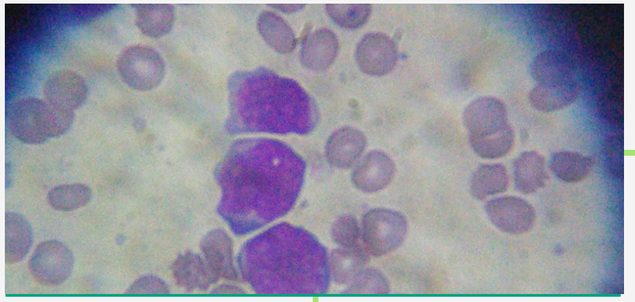
Problem Definition

- Leukemia, a group of blood cancers characterized by abnormal white blood cell production, requires timely and accurate detection of the specific types of leukemia for effective treatment. However, current diagnostic methods are subjective and prone to errors.
- To address this, a new approach utilizing pre-trained CNN models and ensemble learning has been proposed. By combining multiple models, this method aims to enhance accuracy and efficiency in classifying leukemia types. By leveraging the power of collective intelligence, this novel methodology can provide a reliable and robust system for improved leukemia classification.

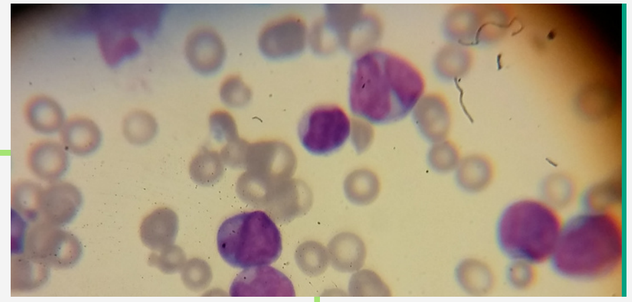


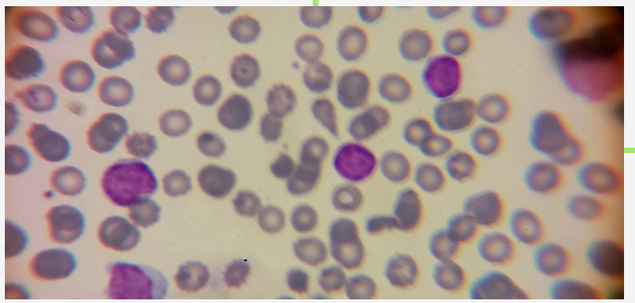
Types of Leukemia

ALL

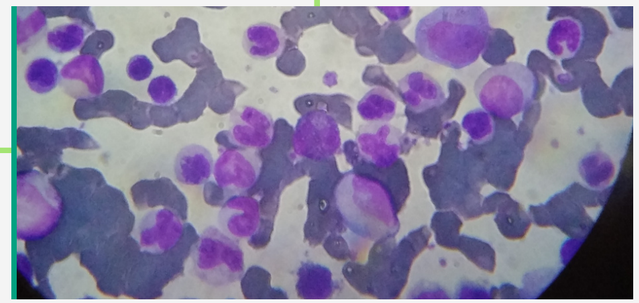


AML





CLL



CML

Motivation

- Blood cancer is an uprising issue and medical procedure is too sensitive and time-consuming to detect any blast cell.
- Traditional systems take time to predict the results.
- Detecting leukemia through optical blood smear testing under the supervision of a specialist is a critical and time-consuming stage.
- Manual testing includes blood tests, spinal fluid tests, bone marrow tests, imaging tests, etc.

Objectives

Detection of all 4 types of Leukemia using ensemble learning and pre-trained convolutional neural network (CNN) algorithms.

To develop a system that detects the presence of abnormal blood cells in a human body which indicates the sign of cancer cells using ensemble learning.

To check the performance analysis of each model.

To analyze different techniques and find the most appropriate technique for effective detection of disease.

Literature Review

Paper Title	Author Name	Conference/ Journal Name	Findings from Paper	Year of Publication
Executing Spark BigDL for Leukemia Detection from Microscopic Images using Transfer Learning	Aftab M O, Javed Awan M, Khalid S, Javed R, Shabir H.	1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)	Images from the dataset are resized and converted to grayscale as a part of image preprocessing. Using transfer learning and the BigDL library with the Apache Spark framework, you can achieve 97.33% training accuracy and 94.78% testing accuracy	2021
Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification	Luis H S Vogado, Rodrigo M S Veras, Flavio H D Araujo, Romuere R V Silva, Kelson R T Aires	Engineering Applications of Artificial Intelligence	Pre-trained CNN model such as AlexNet, CaffeNet, Vgg-f, and AlexNet + CaffeNet + Vgg-f for image processing and Support Vector Machine for classification and achieved an accuracy of 99.29%, 99.76%, 99.58%, and 98.02% respectively	2018
Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network.	Ahmed, Yigit, Isik, Alpkoçak	Diagnostics	Data augmentation is performed to balance the number of images. CNN models give superior precision than traditional machine learning methods.	2019

Paper Title	Author Name	Conference/ Journal Name	Findings from Paper	Year of Publication
IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning.	Nighat Bibi, Misba Sikandar, Ikram Ud Din, Ahmad Almogren, Sikandar Ali	Journal of Healthcare Engineering	Preprocessing steps such as rotation, flipping, and shifting of an image are performed on the dataset. DenseNet-121 and ResNet-34 models are used for classification and attained an accuracy of 99.91% and 99.56% respectively.	2020
Deep Transfer Learning in Diagnosing Leukemia in Blood Cells.	Mohamed Loey, Mukdad Naman, Hala Zayed	Computers	For 1st model, they used AlexNet for image processing and machine learning algorithms for classification where the accuracy of SVM is 99.93%, and for the second model, they used AlexNet for image processing as well as classification where accuracy is 100%	2020
Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images: A Systematic Review.	Mustafa G., Farkhondeh A., Azamossadat H., Davood B., Hassan A. and Arash R.	Scientific Programming	Image segmentation can be performed using various algorithms such as thresholding, k-means clustering, boundary-based segmentation, region-based segmentation, etc. Transfer learning gives better results than machine learning algorithms.	2021

Paper Title	Author Name	Conference/ Journal Name	Findings from Paper	Year of Publication
Leukemia Disease Detection and Classification Using Machine Learning Approaches: A Review	A Ratley, J Minj, P Patre.	First International Conference on Power, Control and Computing Technologies (ICPC2T)	For noise removal, a wiener filter and median filter are used. Machine learning algorithms such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Linear Discriminant Analysis(LDA) are used.	2020
Classification of Leukemia and Leukemoid Using VGG-16 Convolutional Neural Network Architecture	G Sriram, T R Ganesh Babu, R Praveena, J V Anand	MCB Molecular and Cellular Biomechanics	VGG-16 model is used for both feature extraction and classification.	2022
Automated Detection of Acute Lymphoblastic Leukemia From Microscopic Images Based on Human Visual Perception	Bodzas A, Kodytek P, Zidek J	Frontiers in Bioengineering and Biotechnology	For Image Segmentation, Leukocyte Localization and Region Extraction are performed explicitly containing thresholding and cytoplasm extraction to make the image noise-free. Feature extraction includes the extraction of features like edge, texture, shape, colors, etc.	2020
Ensemble of Convolutional Neural Networks to diagnose Acute Lymphoblastic Leukemia from microscopic images	Chayan M, Md Kamrul H, Mohiuddin A, Md Abdul A, Md Tasnim J, Aishwariya D, Md	Informatics in Medicine Unlocked	Data augmentation technique is used. A combination of ensemble learning and transfer learning is used, where the models used are DenseNet, Xception, MobileNet, and InceptionResNet	2021

Dataset

Raabin Leukemia Dataset

- All samples were taken from patients who were referred to the medical laboratory, which works with Takht-e Tavous Laboratory in Tehran, Iran.
- It should be noted that the imaging was done using a Zeiss microscope and an LG J3 smartphone camera.
- All images are in the size 3120 x 4160
- Total Images: 1800, out of this 400 were taken for training and testing in an 80:20 ratio
- Link: https://raabindata.com/

Proposed System

Project Plan

- Project Duration: 6 months
- Project Phases and Milestones
  - Phase 1: Research and Data Collection:
    - Milestone 1: Literature review and understanding of leukemia types and diagnostic methods.
    - Milestone 2: Collection and preprocessing of diverse diagnostic data, including blood cell morphology, and clinical features.

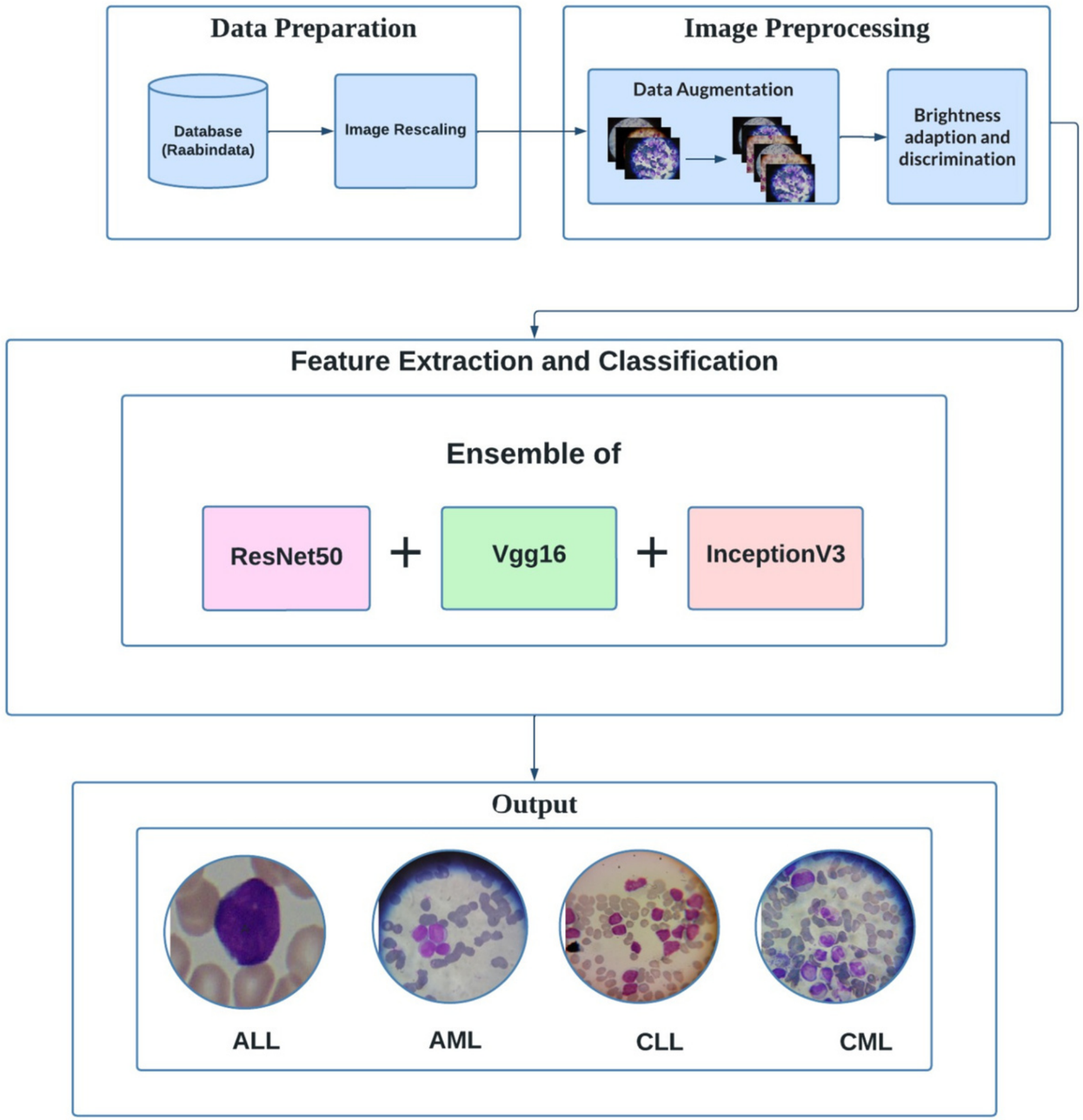
  - Phase 2: Ensemble Learning Model Development
    - Milestone 3: Implementation of various pretrained CNN models suitable for leukemia-type detection and then ensemble them.
    - Milestone 4: Training, compiling and evaluating the models using performance metrics.

- Phase 3: System Integration and Evaluation
  - Milestone 5: Development of a user-friendly interface for the automated leukemia-type detection system.
  - Milestone 6: Integration of the ensemble learning models into the system and testing its functionality.
  - Milestone 7: Evaluation of the system's accuracy, reliability, and performance

- Phase 4: Documentation
  - Milestone 8: Documentation of the project, including detailed technical specifications, algorithms used, and user manual.

- Project Deliverables
  - Comprehensive literature review on Leukemia types and diagnostic methods.
  - Preprocessed and integrated dataset for leukemia type detection.
  - Ensemble learning models implemented and trained.
  - The automated leukemia-type detection system with a user-friendly interface.
  - Documentation including technical specifications, algorithms used, and user manual.
  - Evaluation report showcasing the accuracy, reliability, and performance of the system.

System Architecture



VGG 16

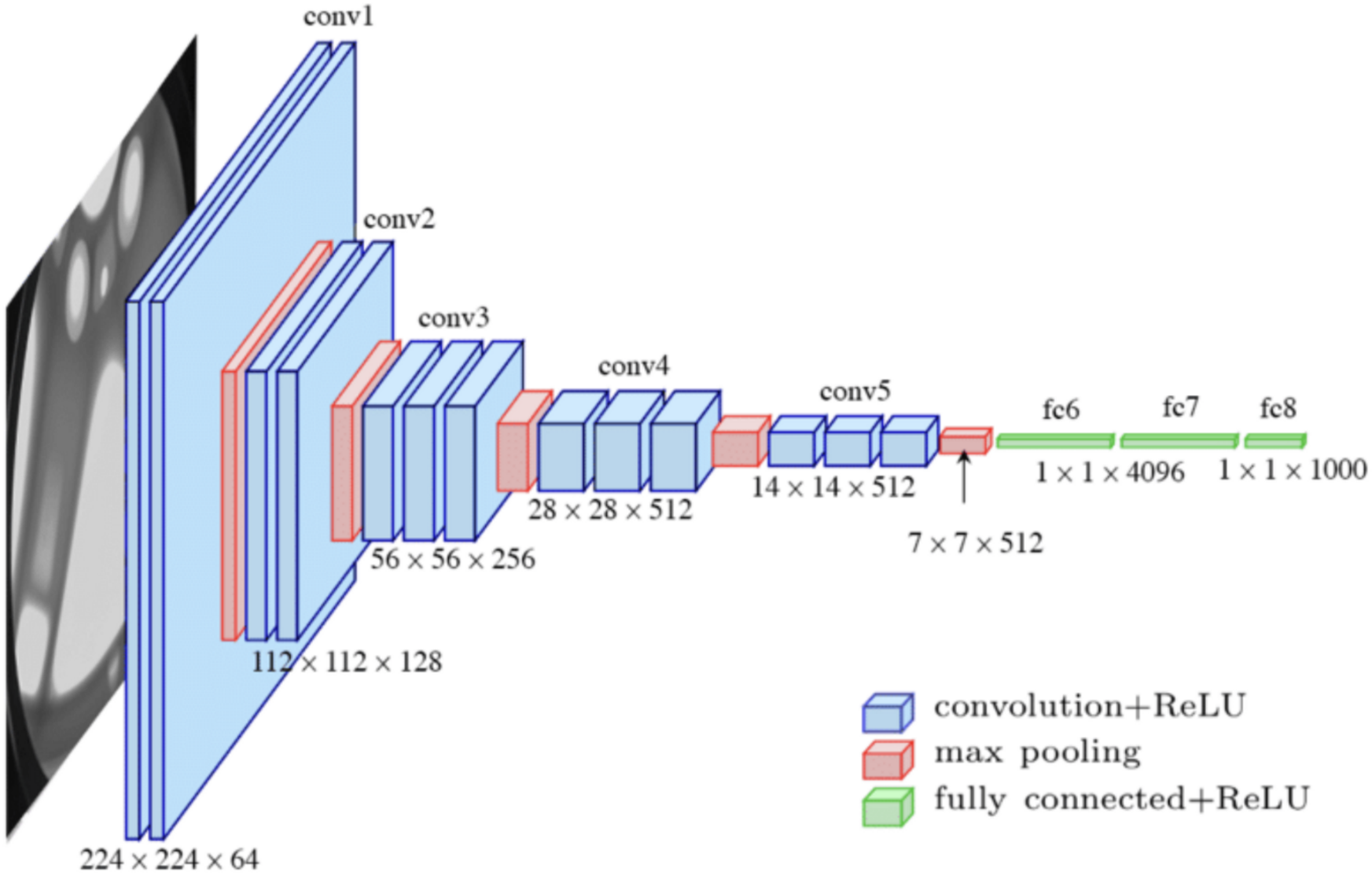
VGG-16 is a convolutional neural network that is 16 layers deep

---

Focuses on having convolution layers of 3x3 filter with stride 1 and utilized the same padding and max pool layer of 2x2 filter with stride 2

Finally, it has three FC (completely connected layers) and a softmax for output

VGG 16 Architecture



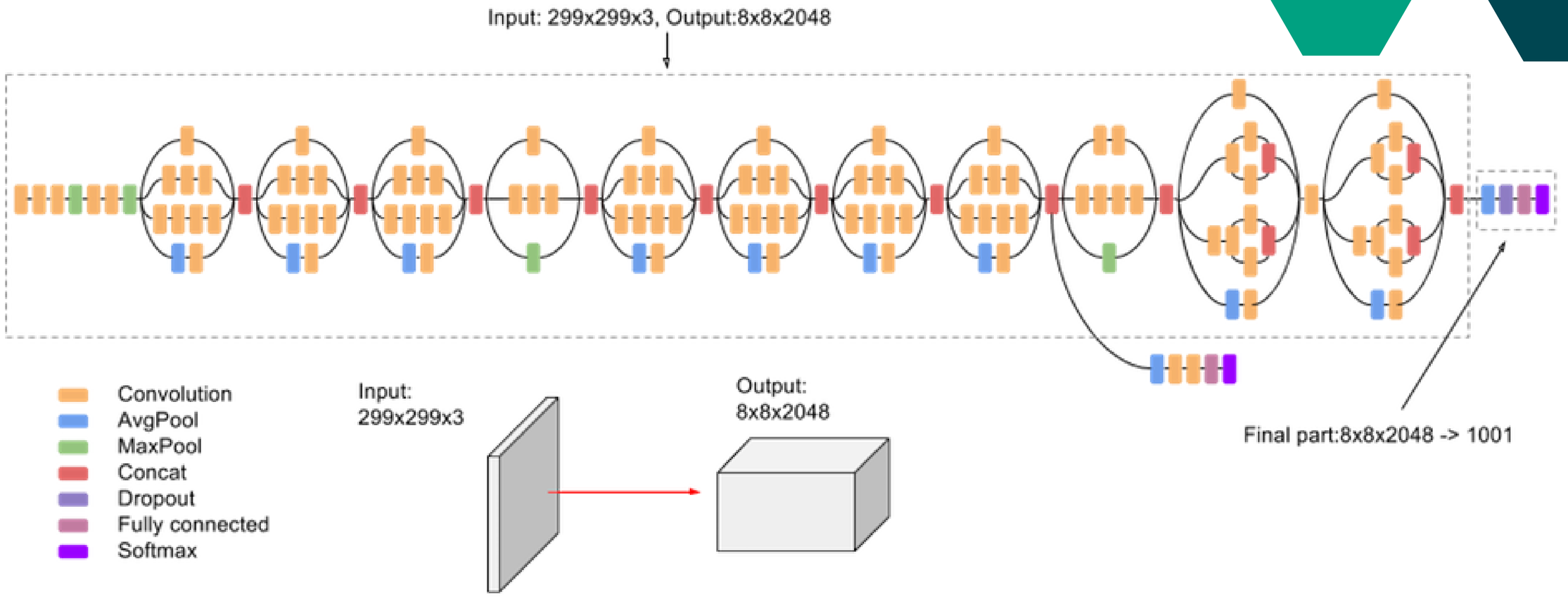
InceptionV3

Inception V3 has a 48-layer structure

Mainly it consists of 3 blocks, inception block A, block B, and block C. It adds convolutional layers to reduce computation complexity.

Inception V3 uses auxiliary classifier activation function to reduce the vanishing gradient problem

InceptionV3 Architecture



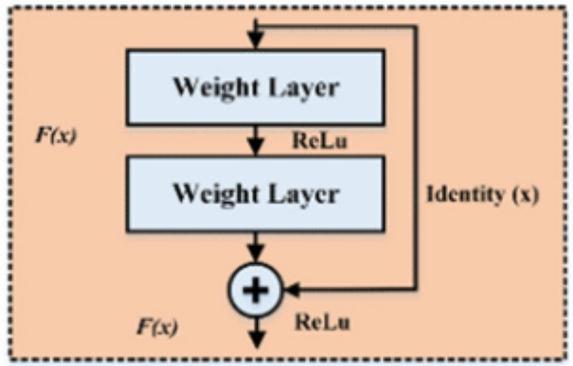
ResNet 50

50-layer structure containing, 48 Convolution layers, 1 MaxPool layer, and 1 Average Pool layer

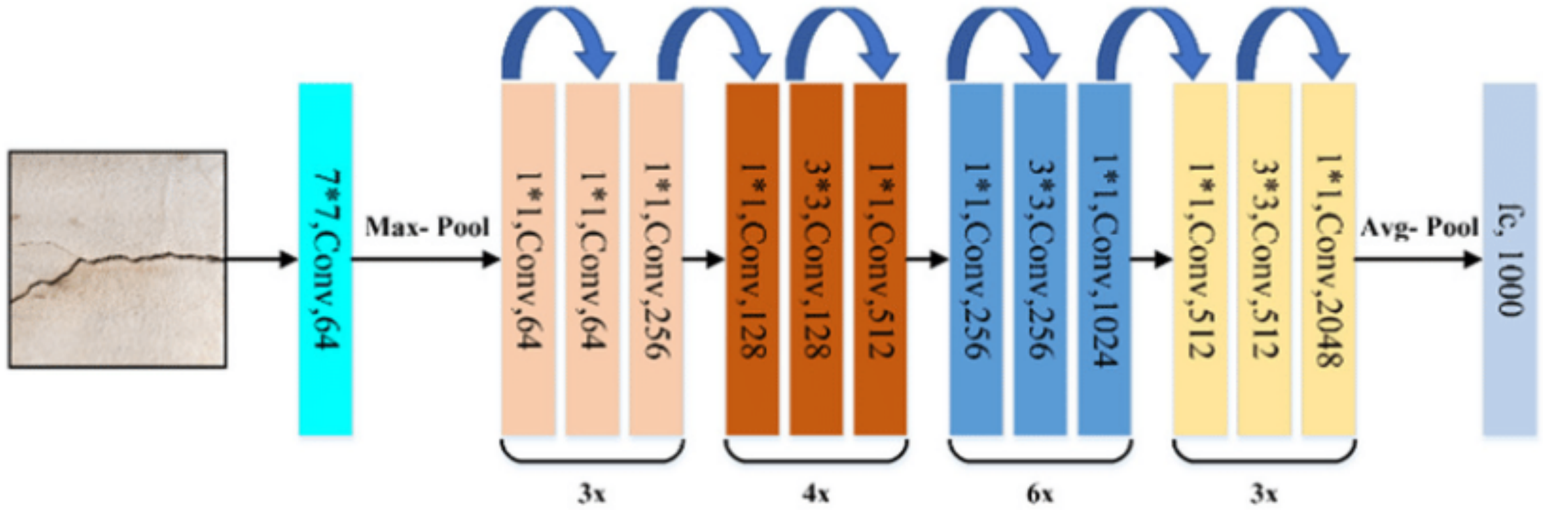
---

The output from an earlier layer is added to a later layer via a skip connection in ResNet

ResNet 50 Architecture



Residual Learning Block



Ensemble Model

1. Stacking individual models in a list
2. Takes a list of tensors as input, all of the same shape
3. Collects outputs of models in a list
4. Averaging outputs

Ensemble Model Structure

Model: "model_3"

Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	[(None, 224, 224, 3)]	0	[]
model1 (Functional)	(None, 4)	14815044	['input_4[0][0]']
model2 (Functional)	(None, 4)	23989124	['input_4[0][0]']
model3 (Functional)	(None, 4)	22007588	['input_4[0][0]']
average (Average)	(None, 4)	0	['model1[0][0]', 'model2[0][0]', 'model3[0][0]']

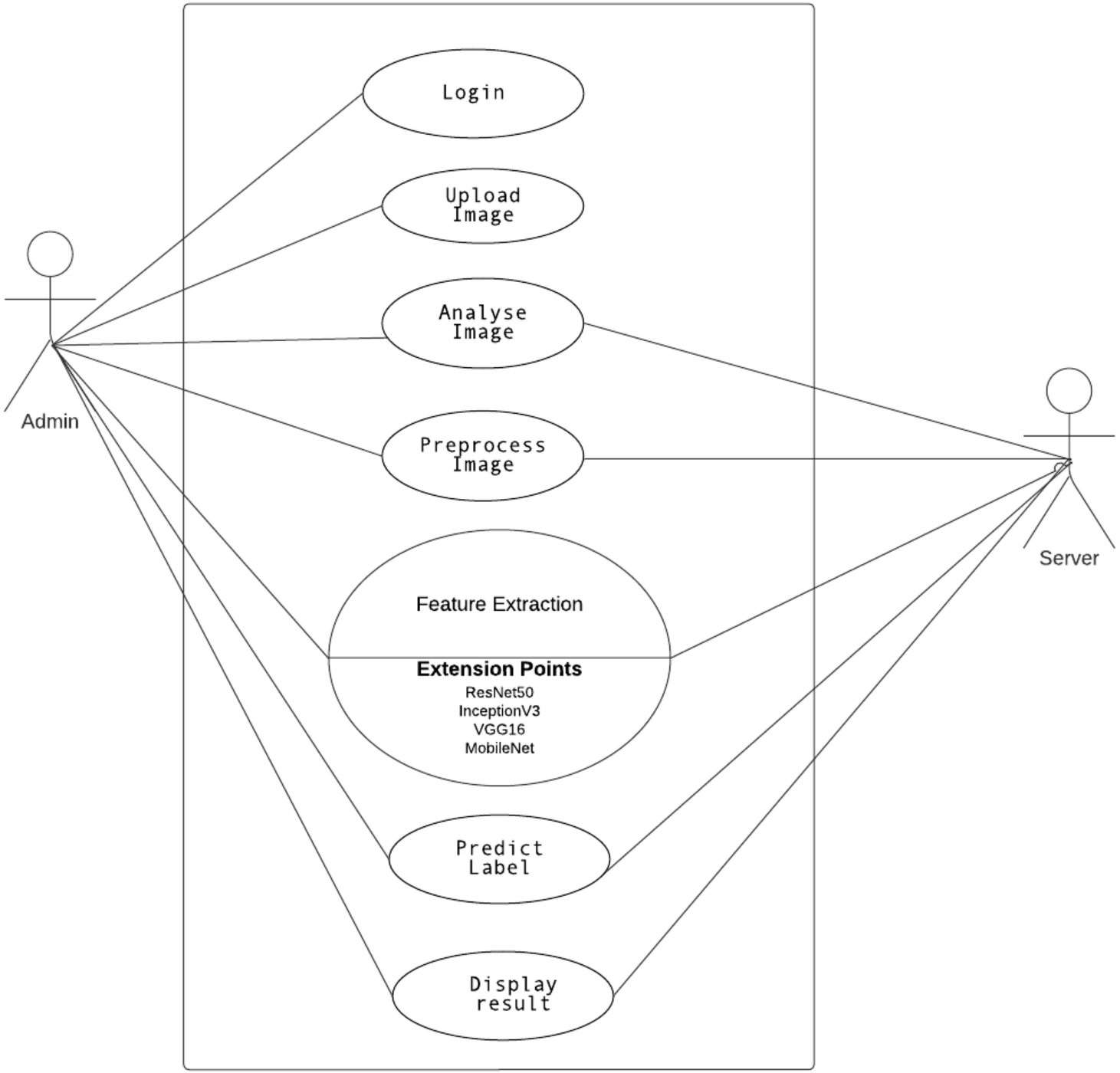
Total params: 60,811,756
Trainable params: 706,572
Non-trainable params: 60,105,184

Implementation

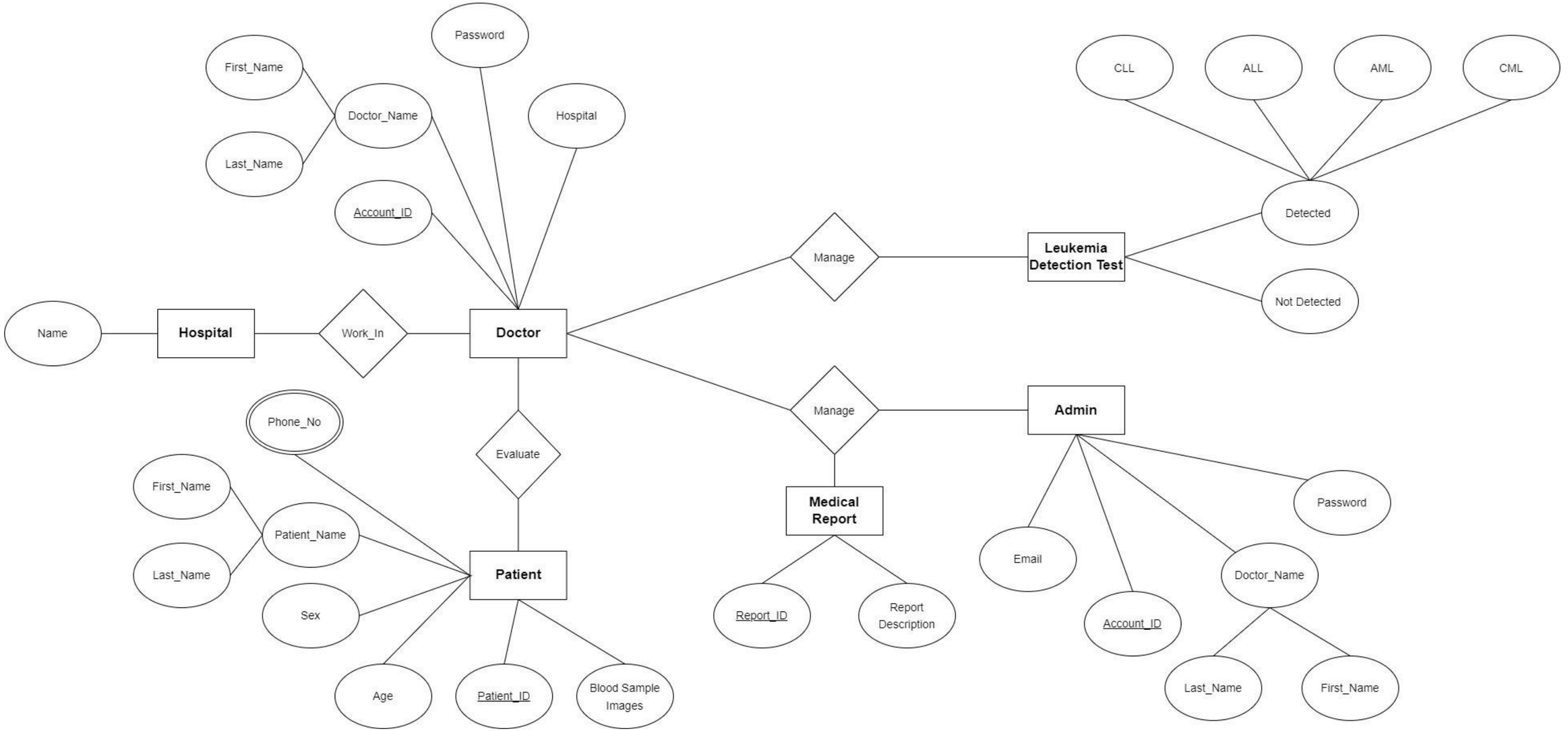
Fit Model

- EarlyStopping is called to stop the epochs early based on some metric It helps to avoid overfitting the model. Over here we are telling you to stop based on val_loss metric, we need it to be minimum.
- Patience says that after a minimum val_loss is achieved then after that in the next iterations if the val_loss increases in any of the 5 iterations then the training will stop at that epoch.

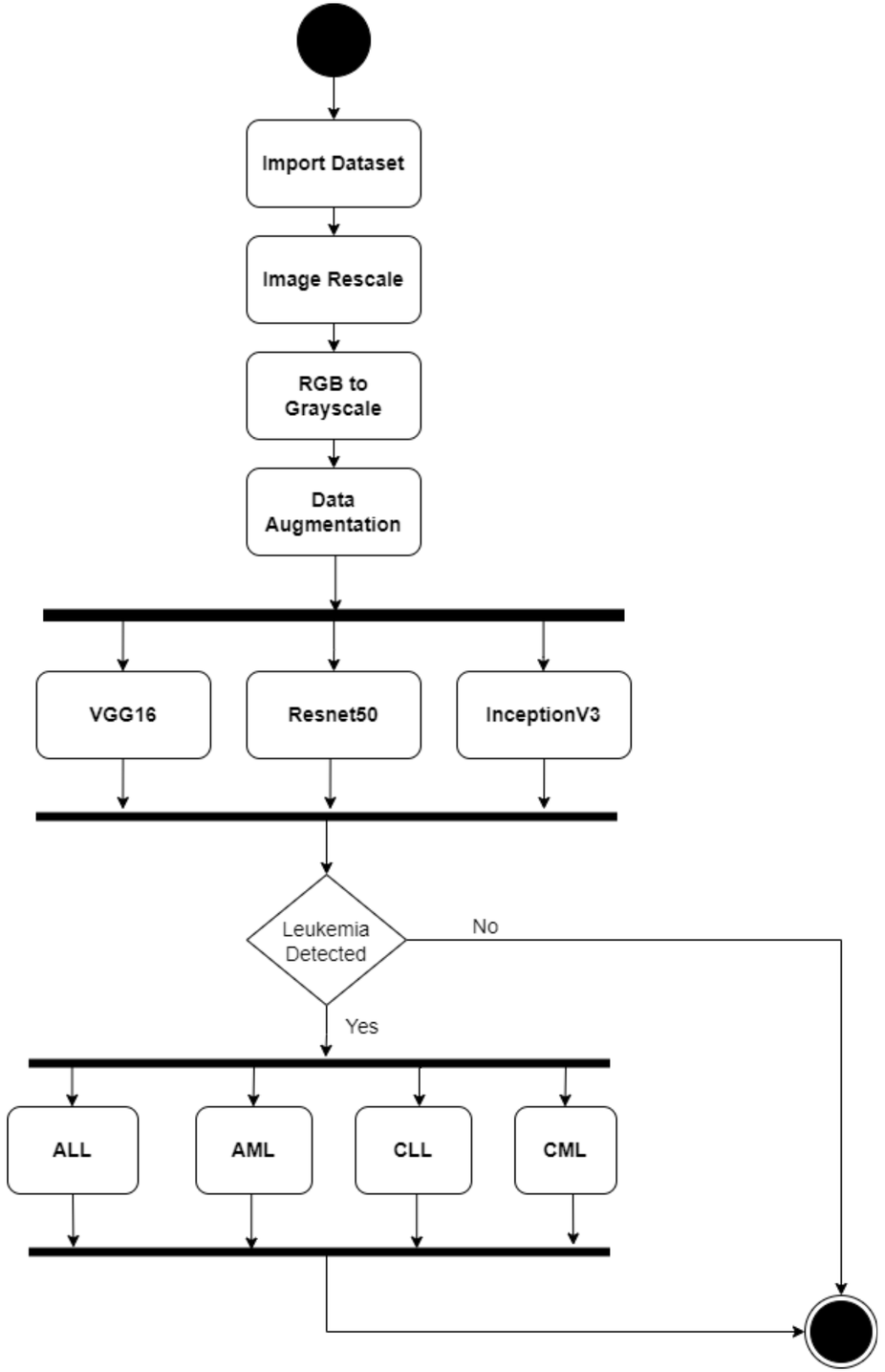
Use Case Diagram



ER Diagram


Password
First_Name
Doctor_Name
Hospital
Last_Name
Account_ID
CLL
ALL
AML
CML
Detected
Manage
Leukemia Detection Test
Not Detected
Name
Hospital
Work_In
Doctor
Manage
Admin
Phone_No
Evaluate
First_Name
Patient_Name
Medical Report
Password
Email
Last_Name
Patient
Doctor_Name
Sex
Report_ID
Report Description
Account_ID
Last_Name
First_Name
Age
Patient_ID
Blood Sample Images

Activity Diagram



Hardware Specifications

- Processor – i5 and more
- RAM - 8 GB
- Hard Disk - 2 TB

Software Specifications

- Database - Image Dataset consisting all four classes images
- Frontend - Streamlit
- Backend – Ensemble Model (VGG16 + Resnet50 + InceptionV3)

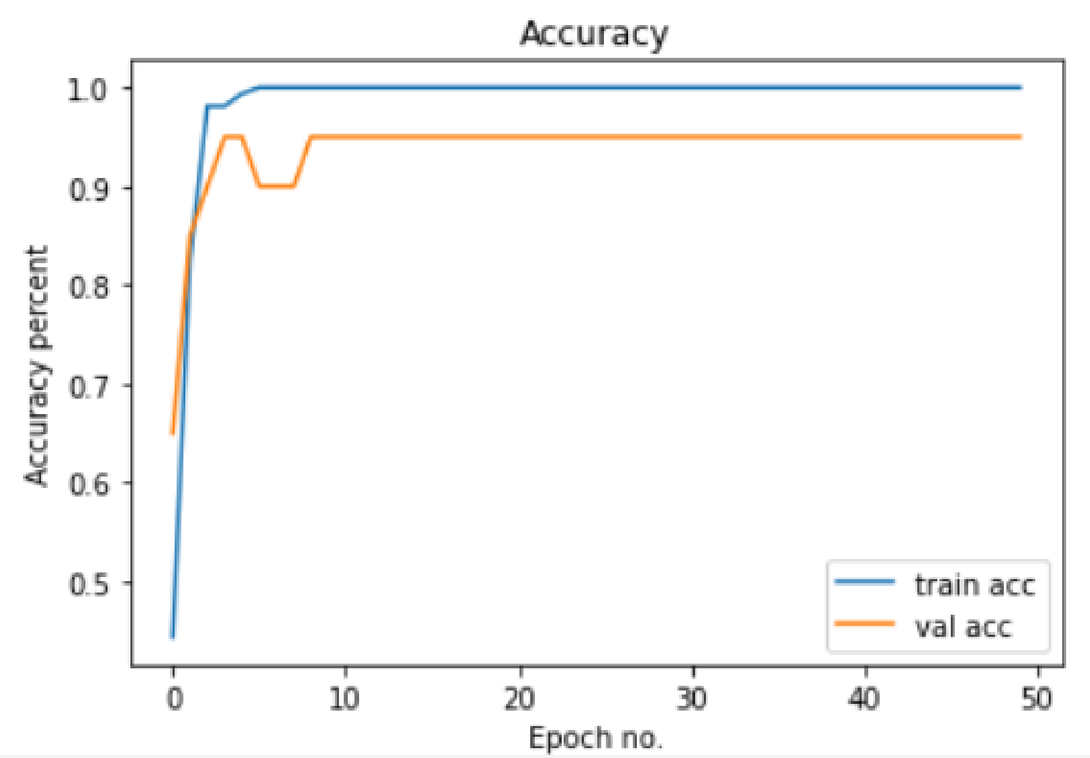
Results

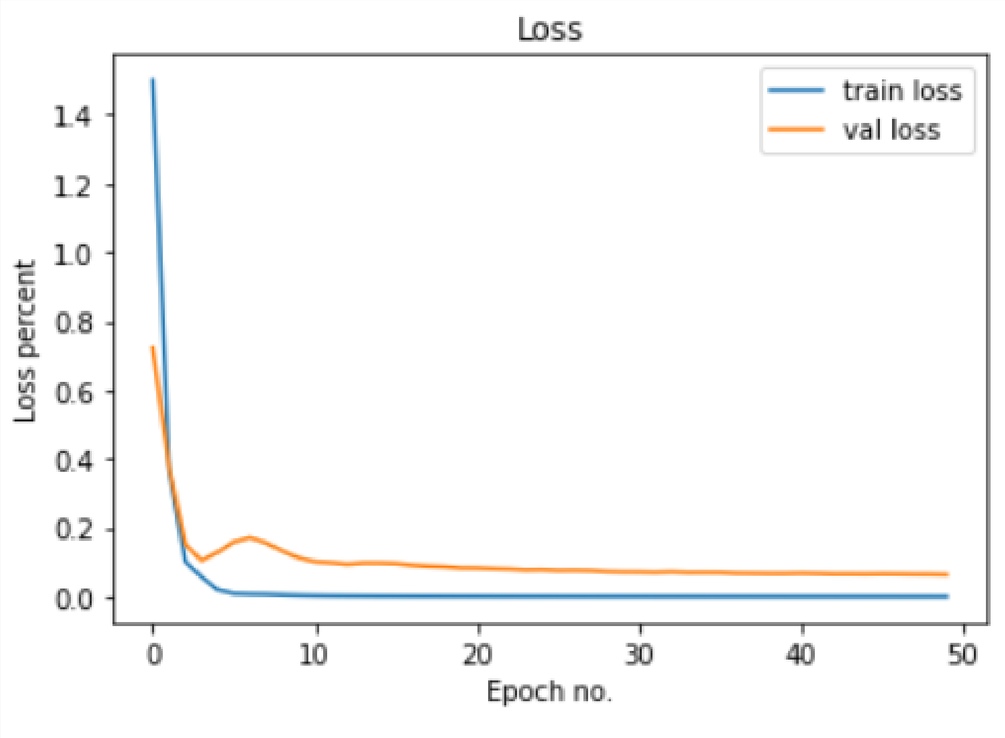
Different Models with their Accuracies

Model	Accuracy
VGG16	98%
Resnet50	90%
InceptionV3	92.5%
VGG16 + Resnet50 + InceptionV3	99.8%
VGG16 + Resnet50	25%
VGG16 + InceptionV3	93%
Resnet50 + InceptionV3	91%

Results

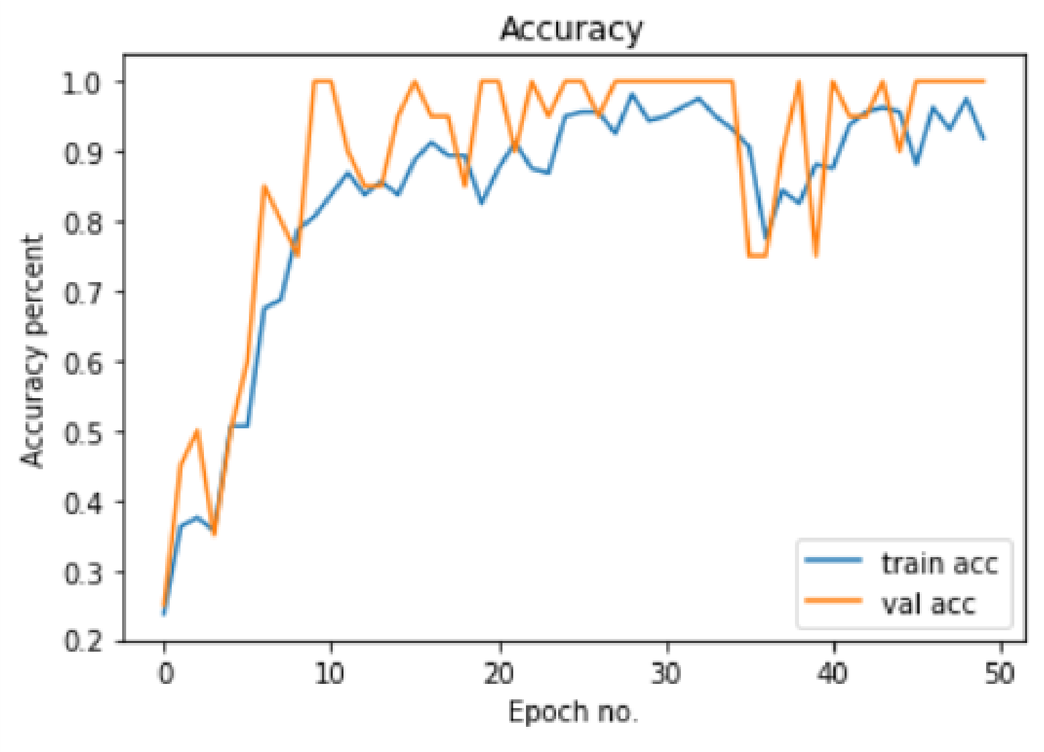
- **VGG16**

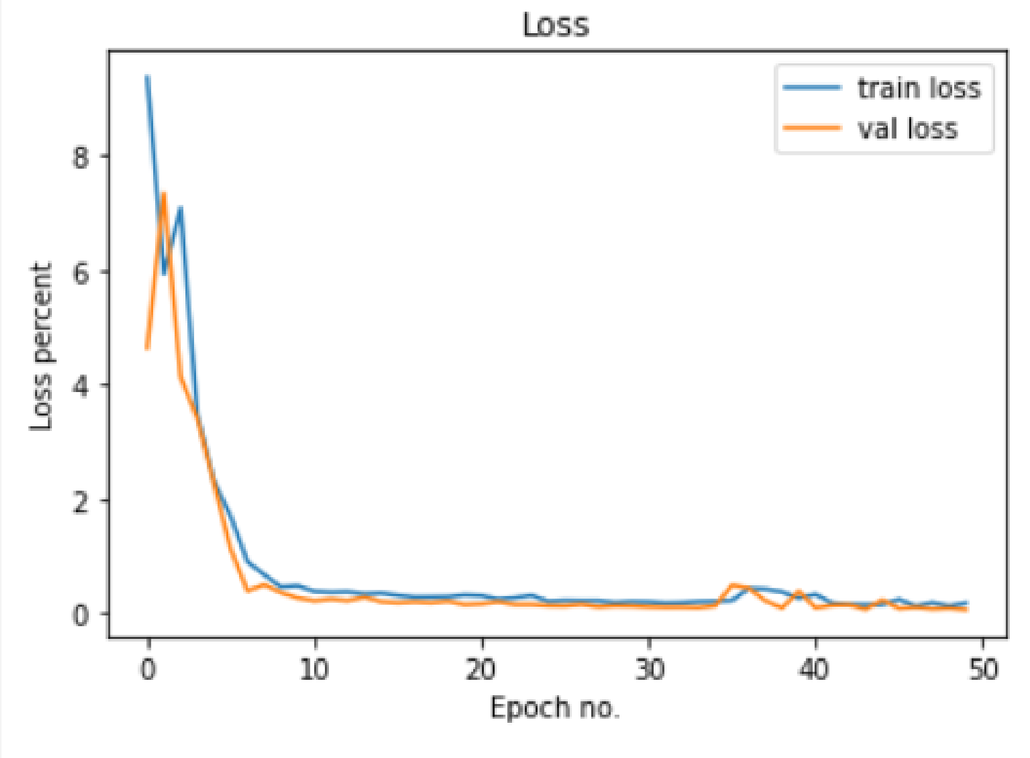




Results

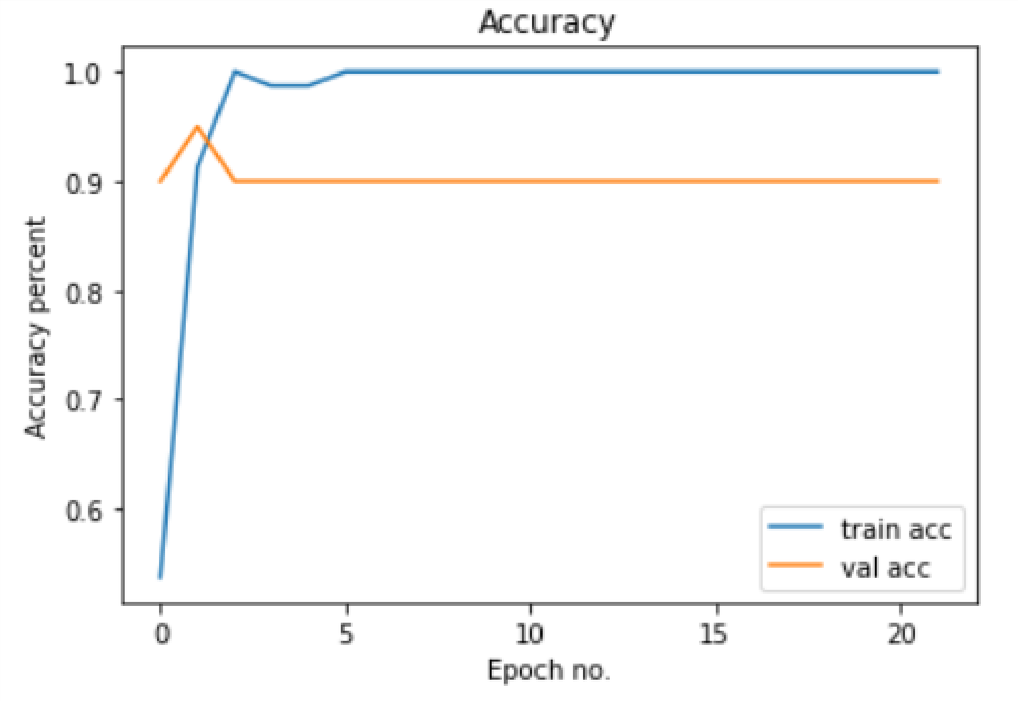
- **ResNet50**

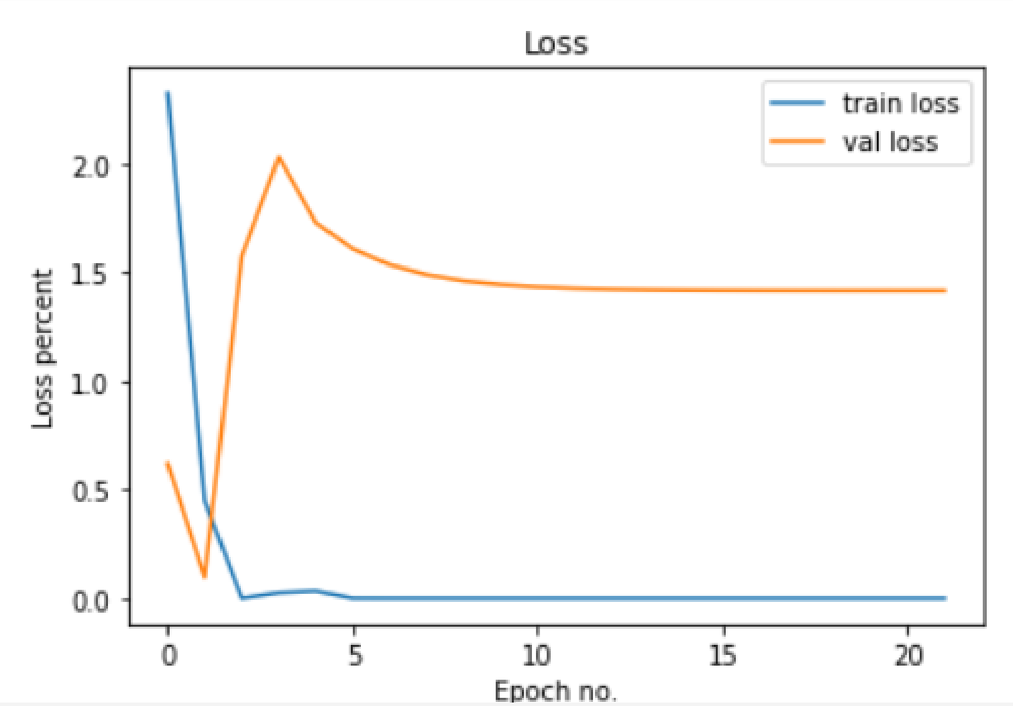




Results

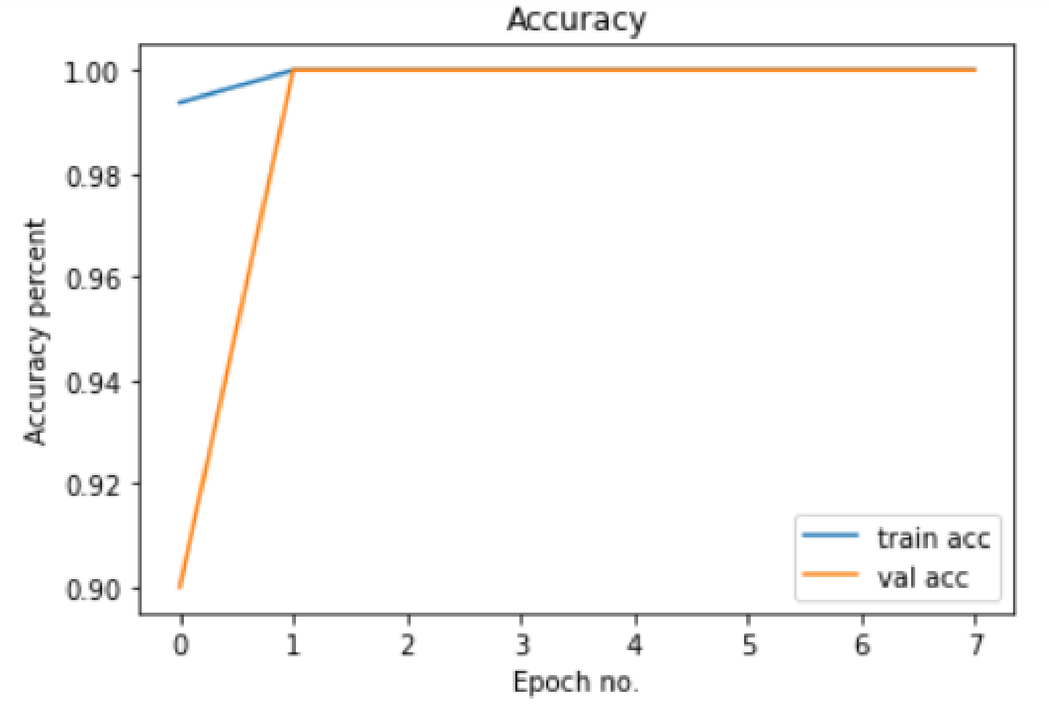
- **InceptionV3**

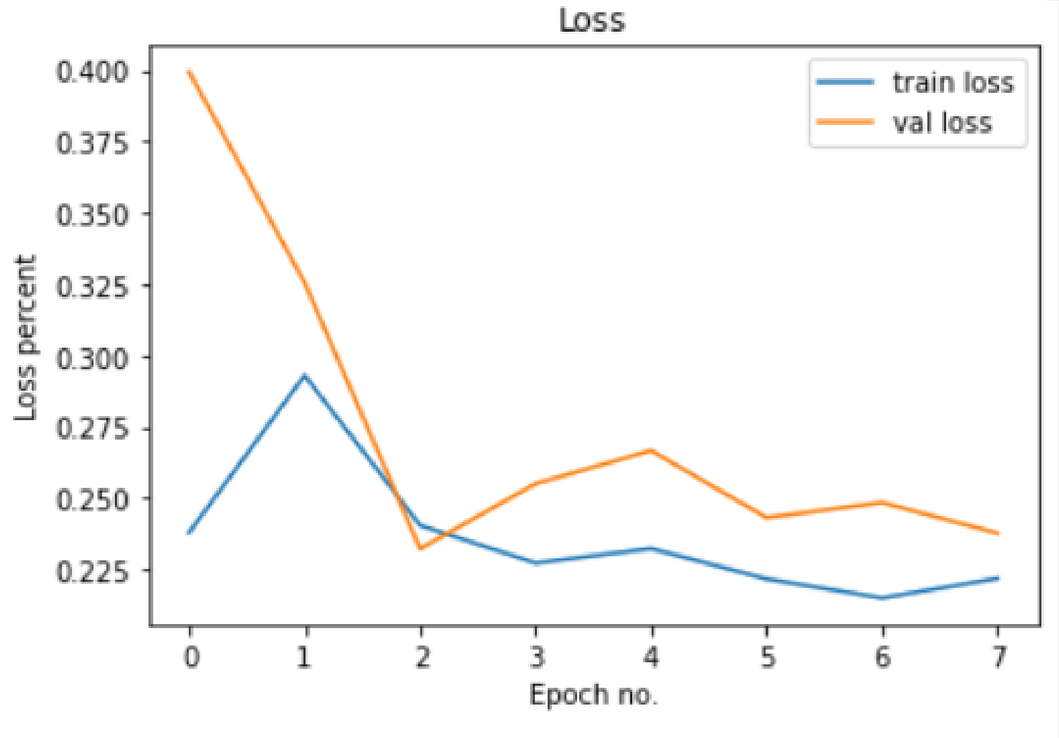




Results

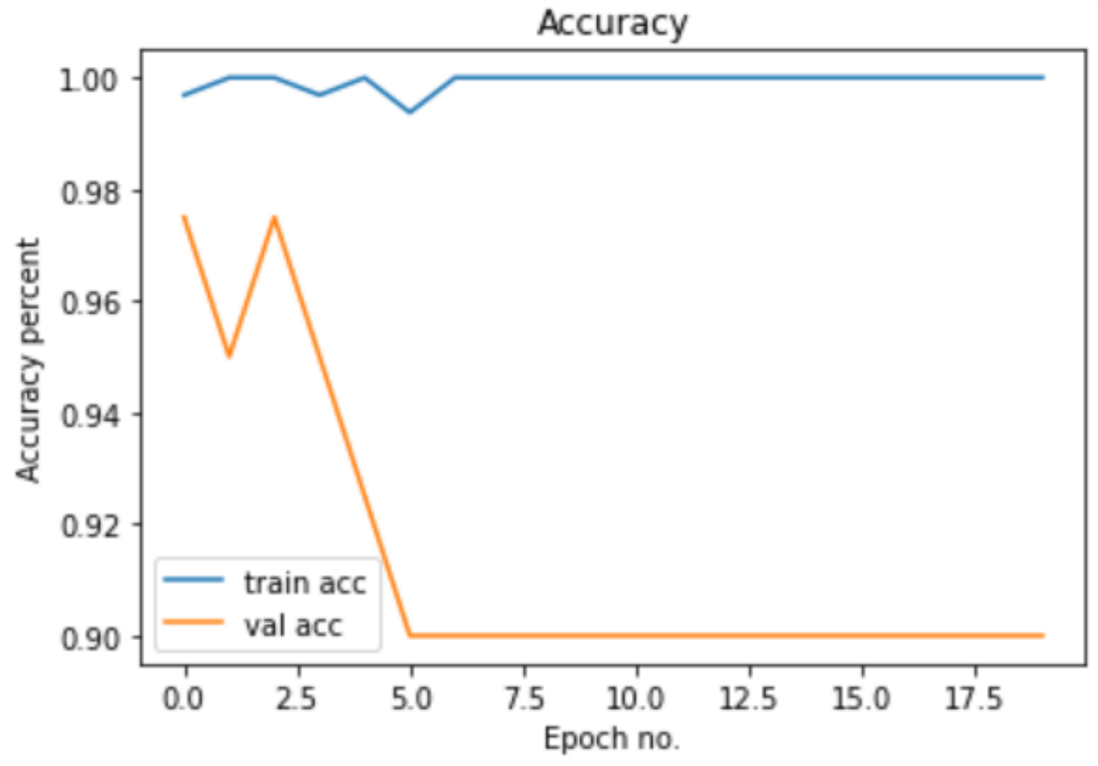
- **VGG + Resnet50 + InceptionV3**

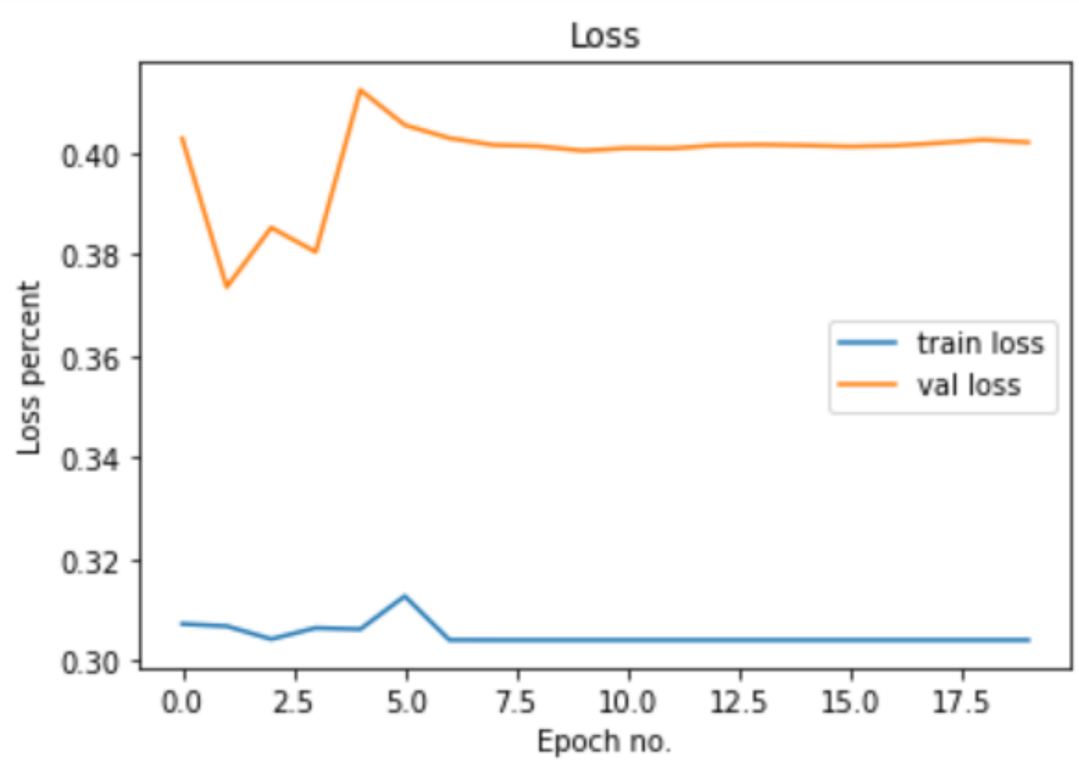




Results

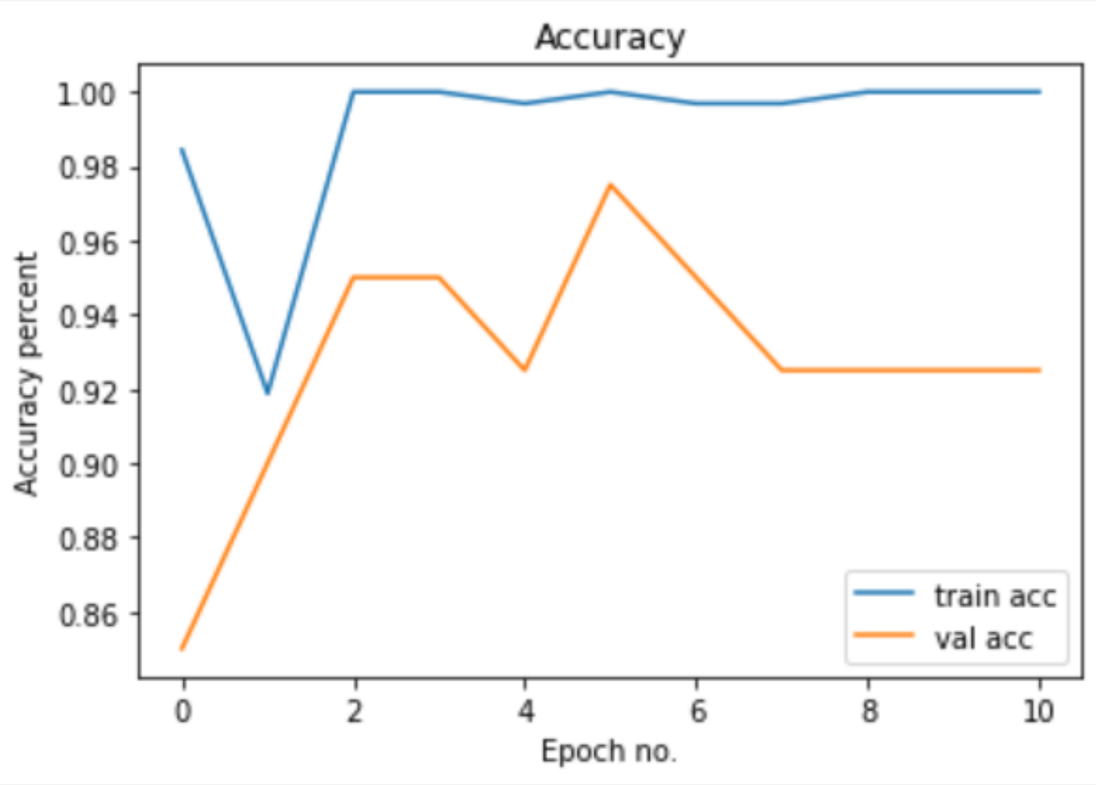
- **VGG + Resnet 50**

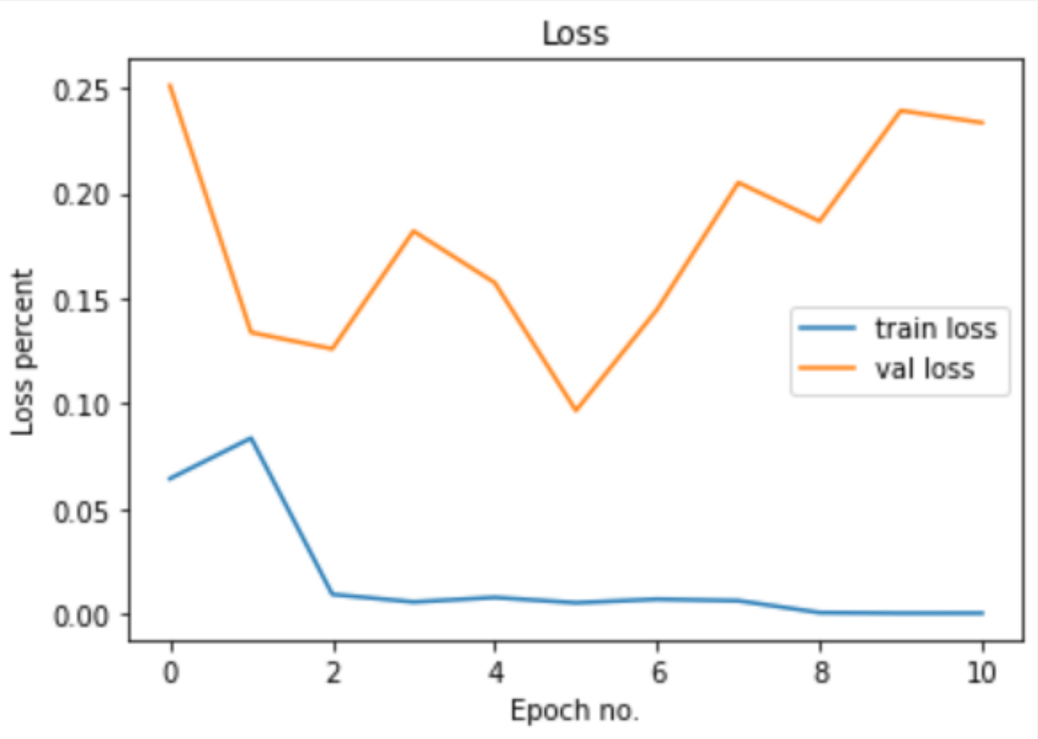




Results

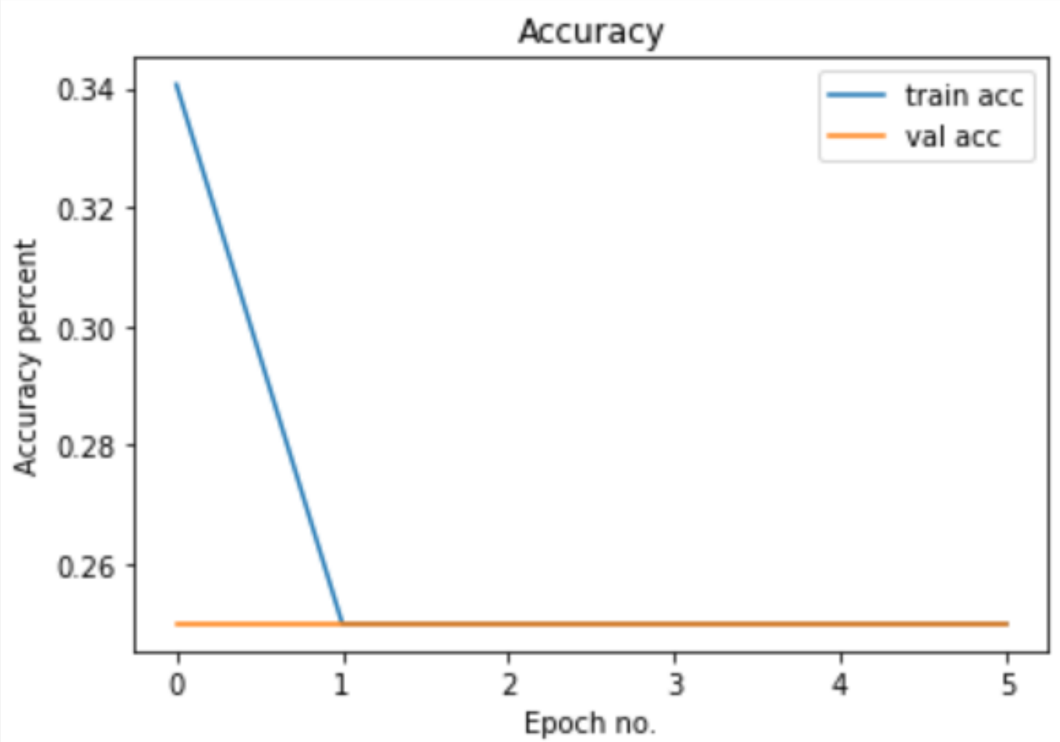
- **VGG + Inception V3**

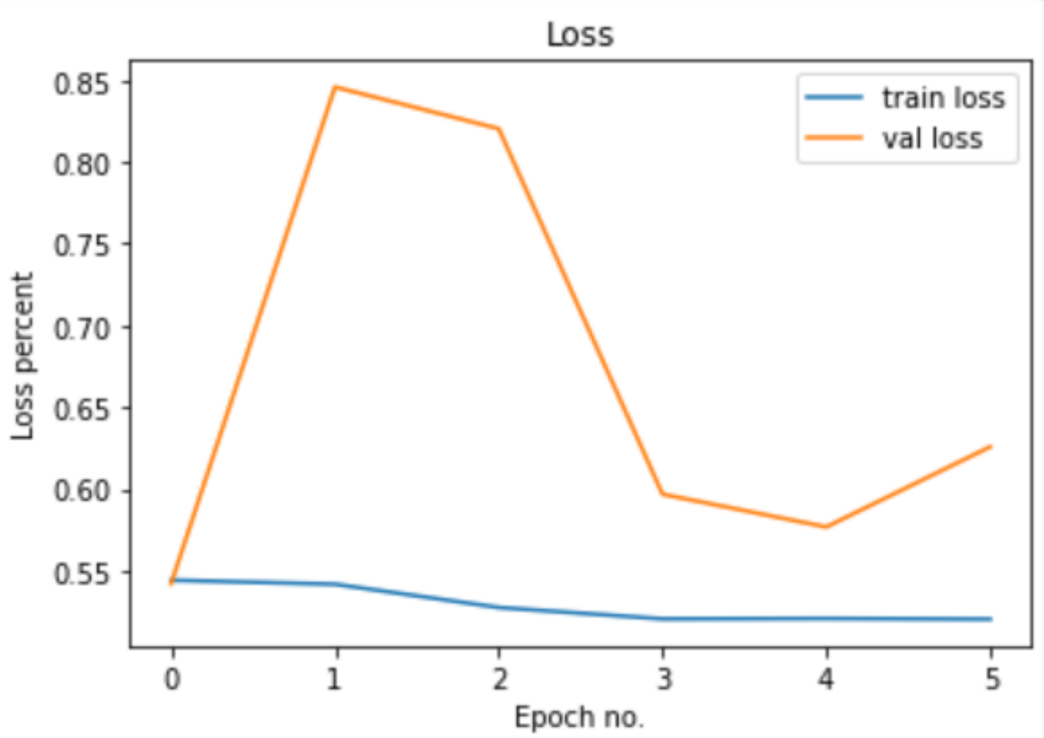




Results

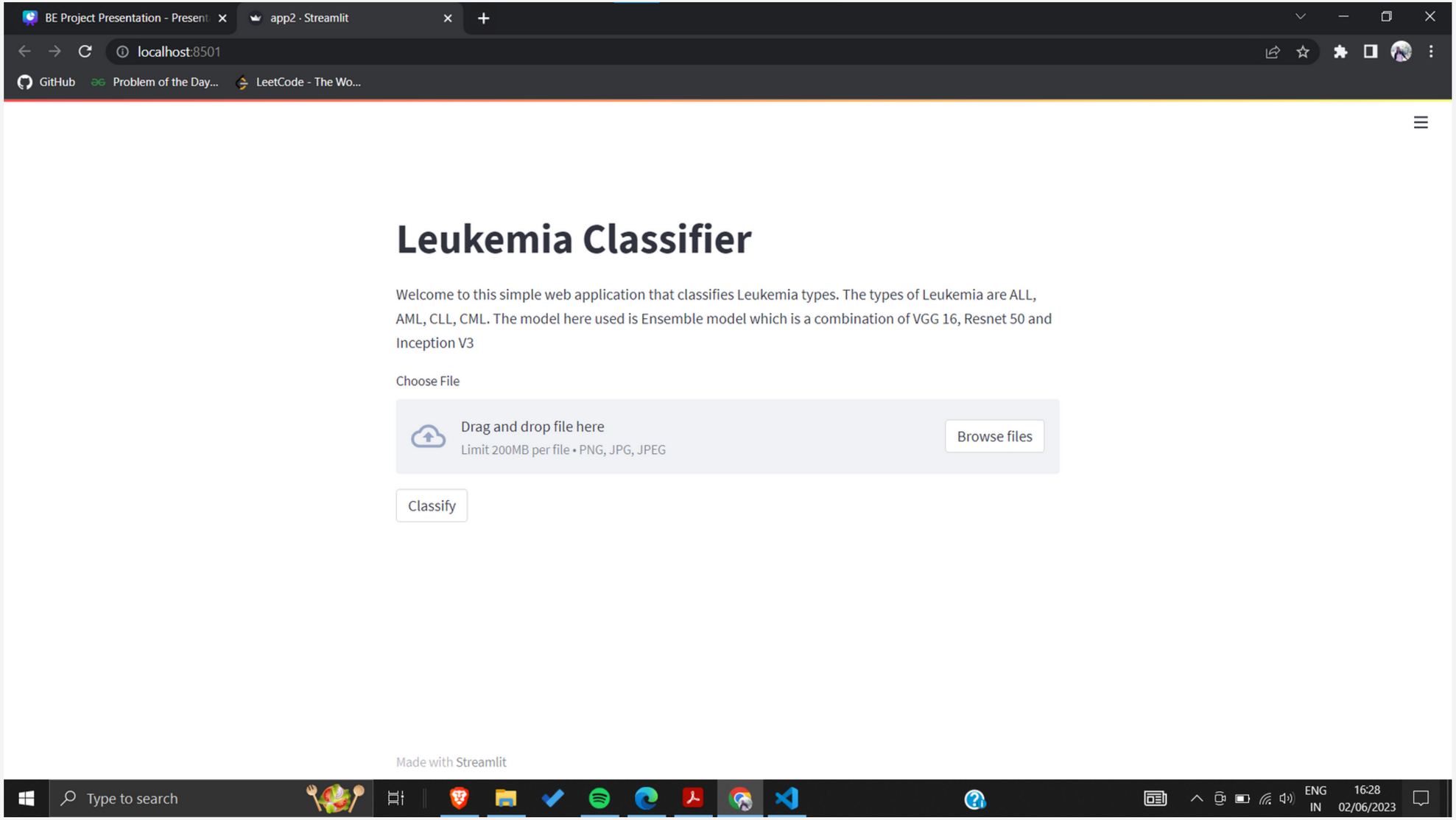
- **Resnet50 + Inception V3**



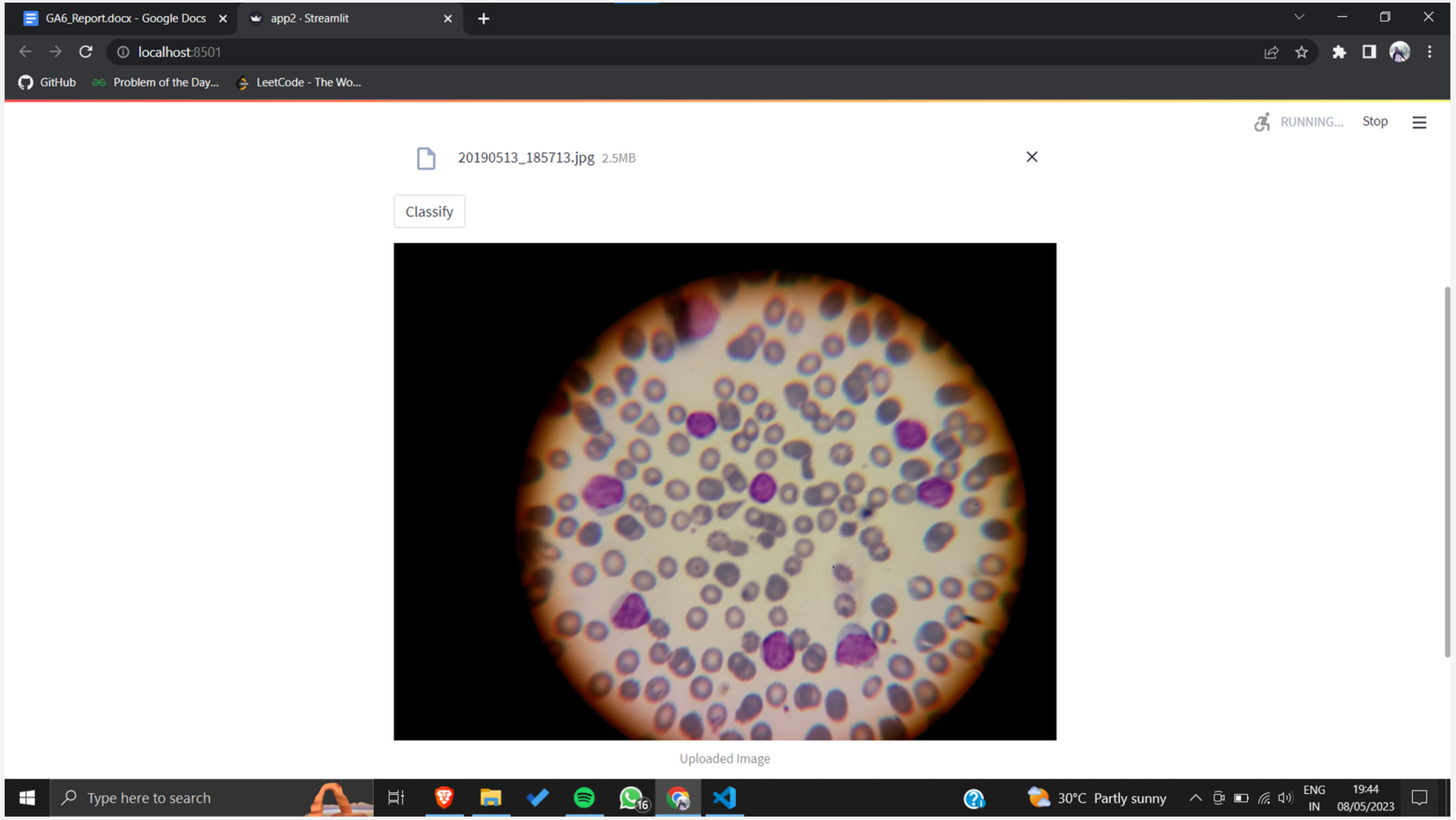


Testing

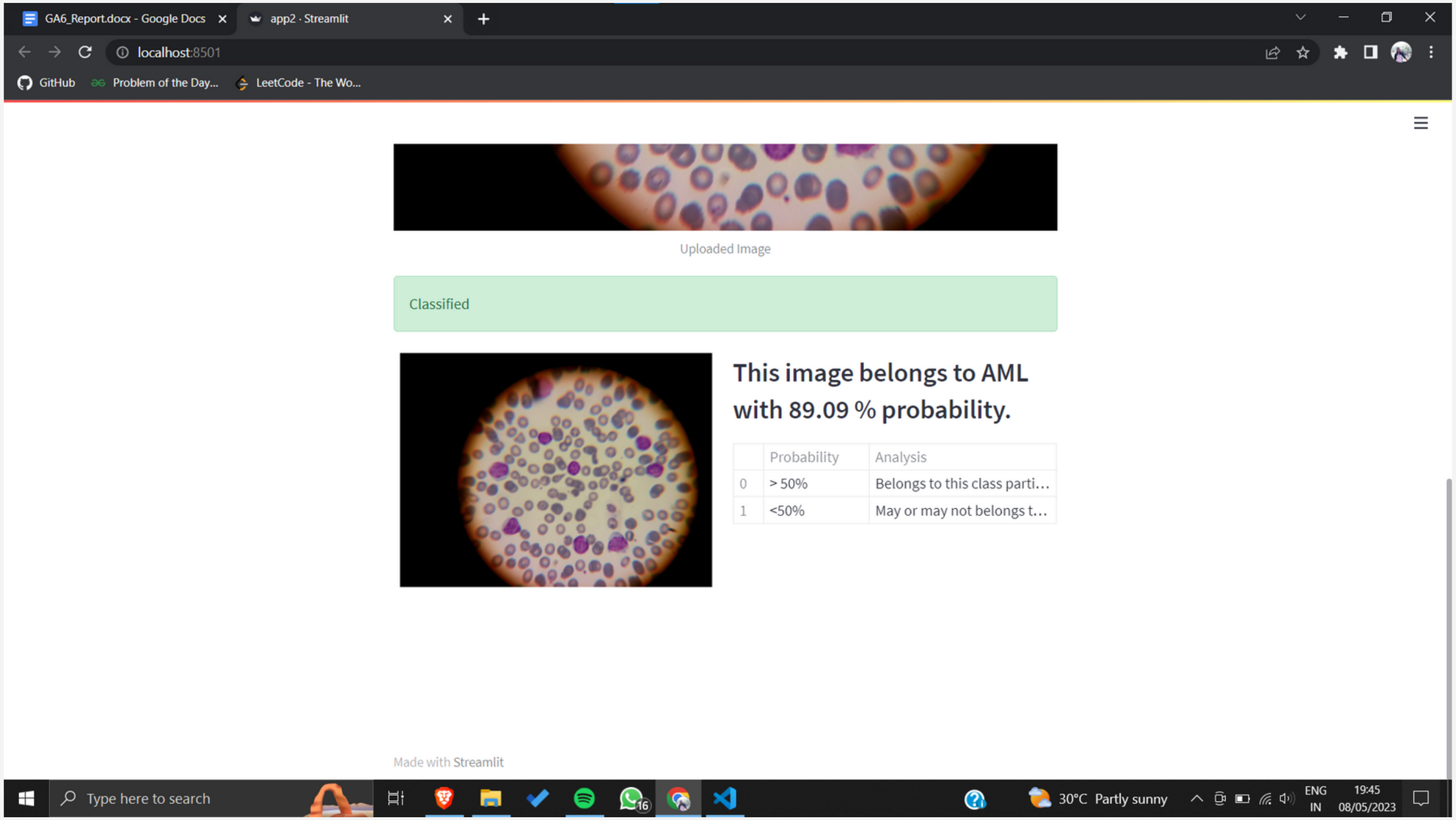
GUI Screenshot 1



GUI Screenshot 2



GUI Screenshot 3



Conclusion

- Performed leukemia types detection using pre-trained models like VGG16, ResNet50, InceptionV3, and an ensemble of these models.
- Used the ensemble model of VGG 16, ResNet50, and Inception V3 which gives us the most appropriate results with higher 99.8% accuracy.
- Ensemble Model takes more computation time but gives high accuracy compared to individual ones.

References

1. Aftab M O, Javed Awan M, Khalid S, Javed R, Shabir H. Executing Spark BigDL for Leukemia Detection from Microscopic Images using Transfer Learning. 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA). 2021 (pp. 216-220)
2. Luis H S Vogado, Rodrigo M S Veras, Flavio H D Araujo, Romuere R V Silva, Kelson R T Aires. Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. Engineering Applications of Artificial Intelligence. 2018; 72: 415-422.
3. Ahmed, Yigit, Isik, Alpkoçak. Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network. Diagnostics. 2019; 9(3):104.
4. Nighat Bibi, Misba Sikandar, Ikram Ud Din, Ahmad Almogren, Sikandar Ali. IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning. Journal of Healthcare Engineering. 2020; 2020(1):1-12
5. Mohamed Loey, Mukdad Naman, Hala Zayed. Deep Transfer Learning in Diagnosing Leukemia in Blood Cells. Computers. 2020; 9(2):29.
6. Mustafa Ghaderzadeh, Farkhondeh Asadi, Azamossadat Hosseini, Davood Bashash, Hassan Abolghasemi and Arash Roshanpour. Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images: A Systematic Review. Scientific Programming.2021;2021(5):1-14.
7. A Ratley, J Minj, P Patre. Leukemia Disease Detection and Classification Using Machine Learning Approaches: A Review. First International Conference on Power, Control and Computing Technologies (ICPC2T). 2020 (pp. 161-165).
8. G Sriram, T R Ganesh Babu, R Praveena, J V Anand. Classification of Leukemia and Leukemoid Using VGG-16 Convolutional Neural Network Architecture. MCB Molecular and Cellular Biomechanics.2022; 19(1):29-40.
9. Bodzas A, Kodytek P, Zidek J. Automated Detection of Acute Lymphoblastic Leukemia From Microscopic Images Based on Human Visual Perception. Frontiers in Bioengineering and Biotechnology. 2020; 8:1005.
10. Chayan Mondal, Md Kamrul Hasan, Mohiuddin Ahmad, Md Abdul Awal, Md Tasnim Jawad, Aishwariya Dutta, Md Rabiul Islam, Mohammad Ali Moni. Ensemble of Convolutional Neural Networks to diagnose Acute Lymphoblastic Leukemia from microscopic images. Informatics in Medicine Unlocked. 2021; 27(12):100794.

THANK YOU!