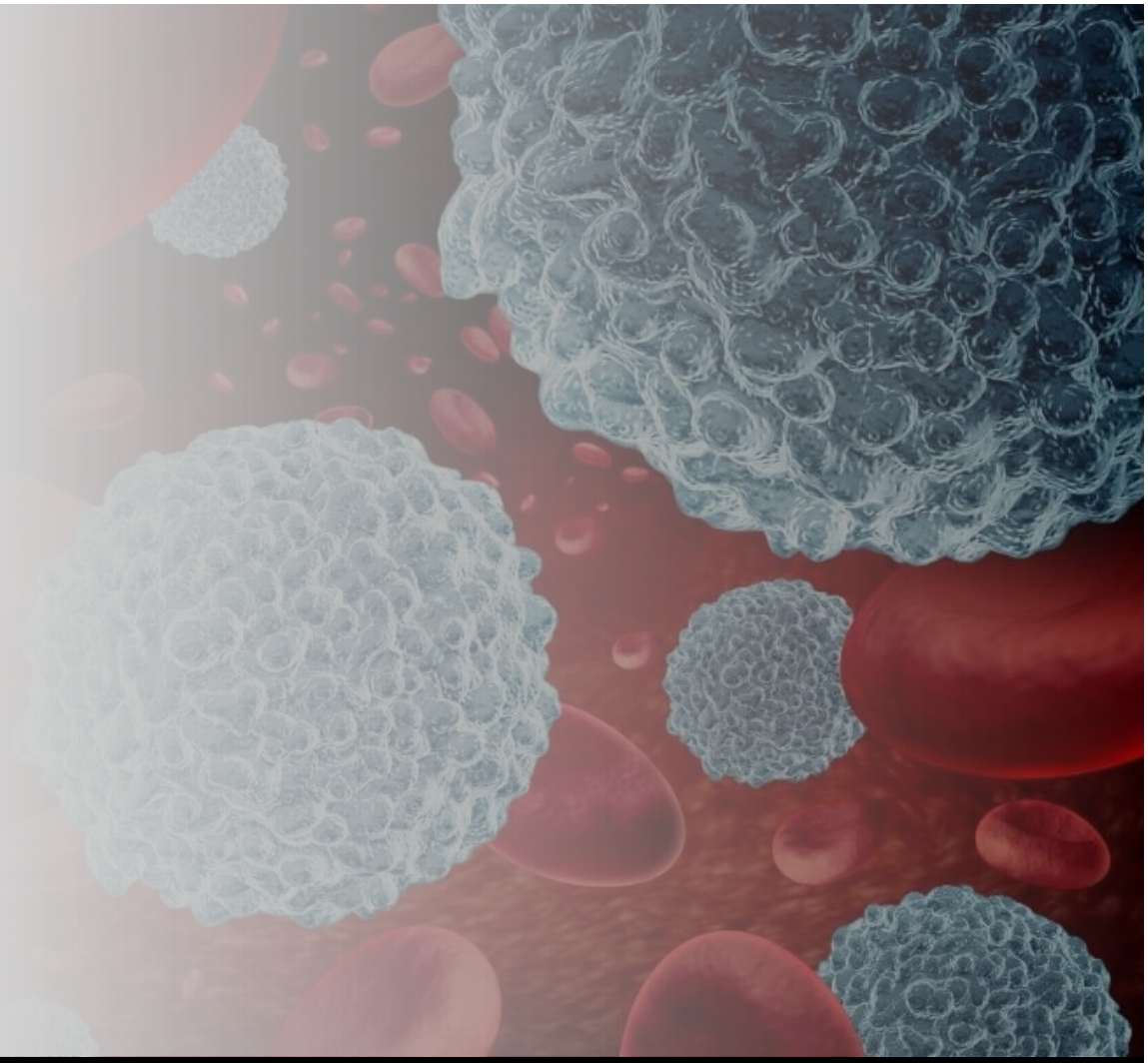
Machine Learning
Group Work – Assignment 2

BLOOD CELL AI –
CLASSIFICATION OF
BLOOD CELL SUBTYPES

Hendrik Künnemann - 57995
Luc Marcel Pellinger - 58611
Moritz Lind - 61230
Tomás Shoemaker Simão Gonçalves - 47016
Zeno Eule - 59093



NOVA SCHOOL OF BUSINESS & ECONOMICS

Transforming Blood-Based Disease Diagnostics with AI

BloodCell AI wants to revolutionize blood-based disease diagnostics through artificial intelligence

Our Purpose and Mission:

We are dedicated to saving lives by revolutionizing blood-based disease diagnostics by leveraging artificial intelligence. We strive to surpass CellVision Diagnostics in building the best model for blood-based disease diagnostics, ensuring our technology leads the way in reliable, efficient, and life-saving diagnostics.

The 4 Pillars of Our Work

Precise Classification	Clinical Significance	Efficiency and Reliability	Seamless Integration
• Blood Cell AI identifies and classifies blood cell subtypes such as Eosinophils, Lymphocytes, Monocytes, and Neutrophils • Utilizes cutting-edge AI technology that is trained on a comprehensive dataset of high-resolution blood cell images	• Essential for diagnosing conditions such as asthma (identified by elevated Eosinophil counts) and detecting infections or immune responses (noted by variations in Lymphocyte and Monocyte levels)	• Cutting diagnostic time from hours to minutes is a vital step that saves precious time, ultimately contributing to saving lives • Minimizes human error, ensuring rapid and informed treatment decisions	• AI models seamlessly integrate into existing healthcare systems • Makes state-of-the-art diagnostics accessible worldwide, ensuring precise and reliable blood sample analysis for a wide range of people

BloodCell AI empowers healthcare professionals to deliver superior patient care based on timely and accurate blood sample analysis. We invite you to join in transforming healthcare diagnostics with AI-driven insights!

Data at the Heart: Exploring the Blood Cell AI Dataset

Our AI models are trained on a meticulously curated dataset of high-resolution blood cell images

Dataset Acquisition:

- In collaboration with renowned research hospitals, BloodCell AI acquired a diverse dataset, including augmented pictures.
- This dataset underwent rigorous curation processes to ensure data quality and relevance to our diagnostic goals.

Dataset Composition:

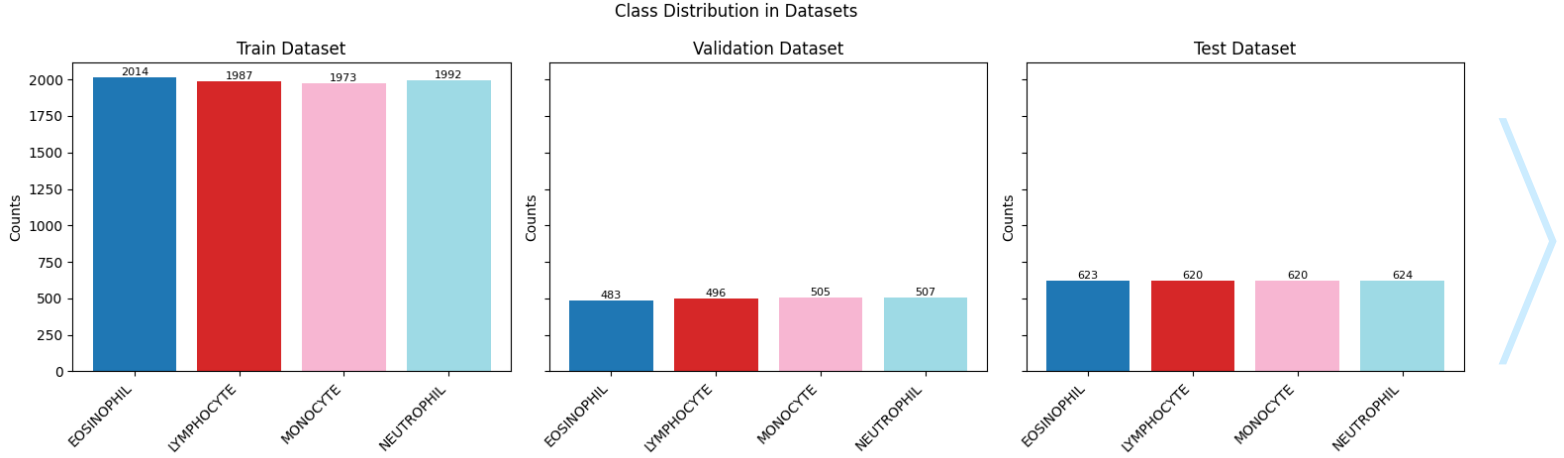
- The dataset comprises a comprehensive collection of blood cell subtypes, including Eosinophils, Lymphocytes, Monocytes, and Neutrophils.
- These subtypes are crucial indicators for various diseases and medical conditions.

Data Handling Protocol:

- To maintain the robustness of our AI models, we implemented a stringent data handling protocol.
- The dataset is meticulously divided into distinct training, validation, and test directories.

Training and Validation:

- Data is further split into training and validation subsets.
- The training subset enables comprehensive model learning, while the validation subset facilitates parameter optimization without compromising generalization.



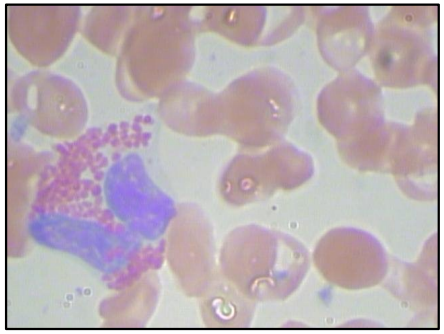
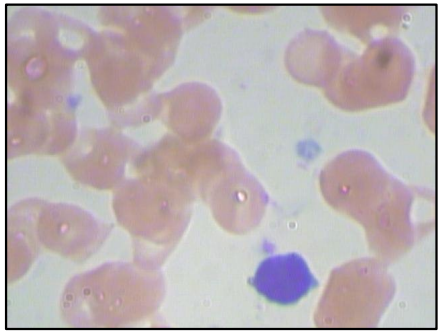
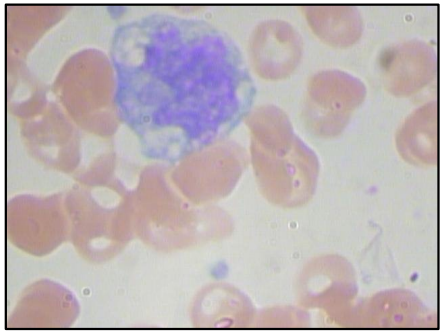
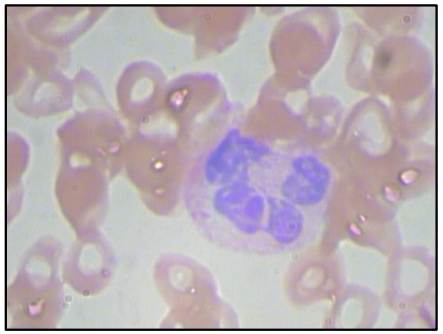
Conclusion

- The **quality and composition** of our dataset are integral to BloodCell AI's mission of transforming healthcare diagnostics.
- By **leveraging this diverse and meticulously curated dataset**, we ensure that each analysis is precise, dependable, and contributes to superior patient care.

Data at the Heart: Understanding Benefits of Recognizing & Classifying

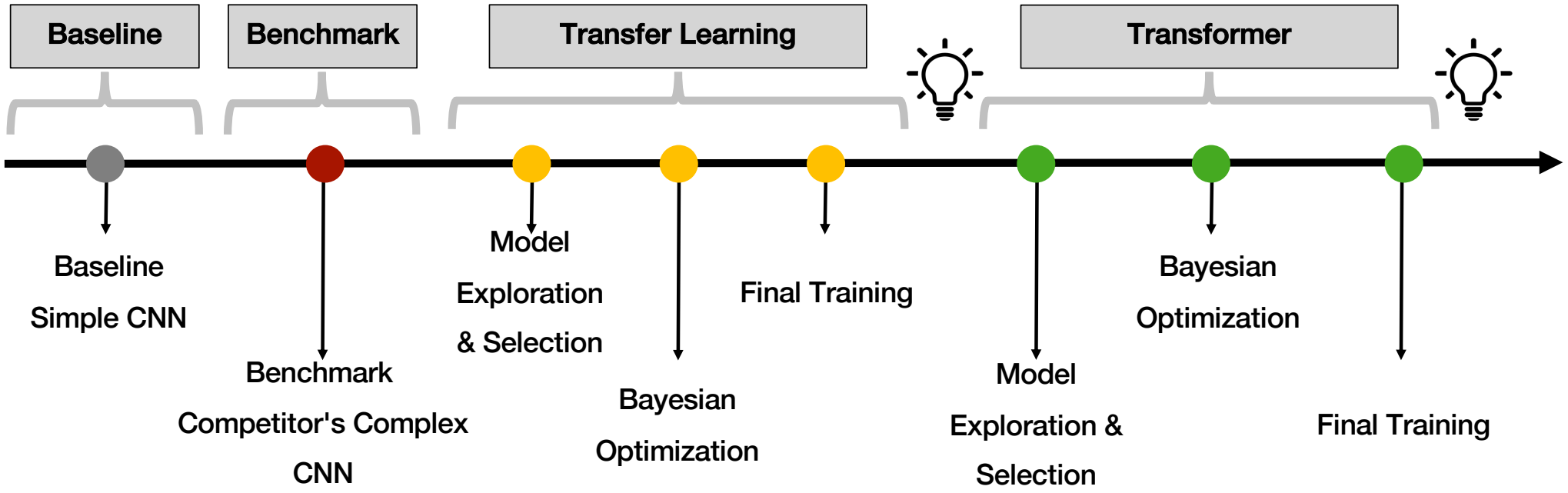
Early detection leads to timely diagnosis and treatment, treatment monitoring tracks effectiveness, patient management improves care based on cell changes, and risk assessment evaluates infection and complication risks.

Eosinophils	Lymphocytes	Monocytes	Neutrophils
• Allergic Reactions: Indicates asthma, eczema, hay fever. • Parasitic Infections: Points to parasitic infections. • Autoimmune Disorders: Signals conditions like eosinophilic esophagitis.	• Viral Infections: Sign of viral infections (e.g., mono, hepatitis). • Chronic Inflammatory Conditions: Indicates diseases like rheumatoid arthritis. • Immune Response: Assesses immune system status, especially in HIV/AIDS.	• Chronic Infections: Indicates chronic infections (e.g., tuberculosis). • Inflammatory Conditions: Associated with inflammatory diseases. • Cancer Detection: Linked to leukemia and lymphoma.	• Bacterial Infections: Suggests bacterial infections (e.g., pneumonia). • Acute Inflammation: Marker for acute inflammation (e.g., heart attack). • Neutropenia: Low counts indicate bone marrow issues or chemotherapy effects.

Modeling – Road Map

Path to Optimal Blood Cell Classification: From Baseline to Advanced Models

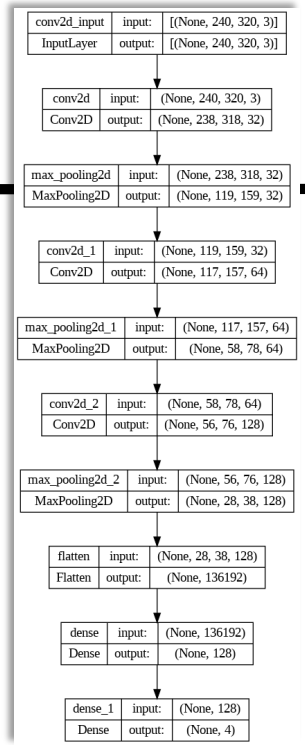


Modeling – Baseline & Benchmarking

Establishing the Foundation: Baseline and Benchmark Models

Basline

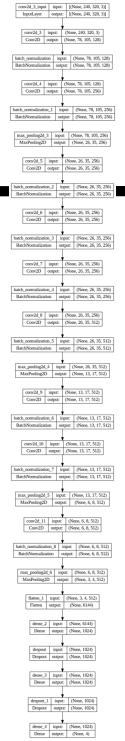
- Feature Extraction with 3 CNN Layers
- Preparational Flatten layer
- Dense layer with 128 units.
- Final layer for classification.
- **Test Accuracy: 24.41%**



Benchmark

CellVision Diagnostics

- Feature Extract with multiple CNN Layers
- Max pooling layers to reduce dimensions.
- Preparational flatten layer.
- 2 Dense Layers with dropout
- Final layer for classification.
- **Test Accuracy: 60.35%**



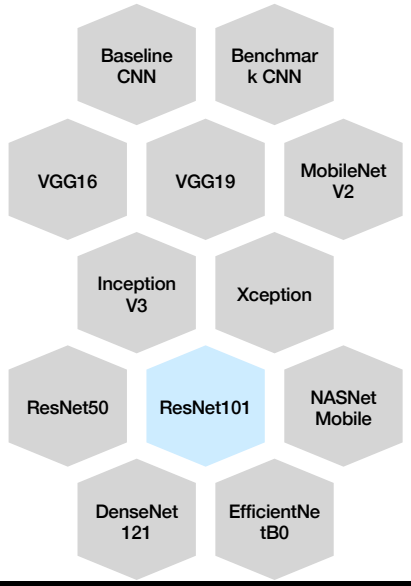
NOVA SCHOOL OF BUSINESS & ECONOMICS

Blood Cell AI

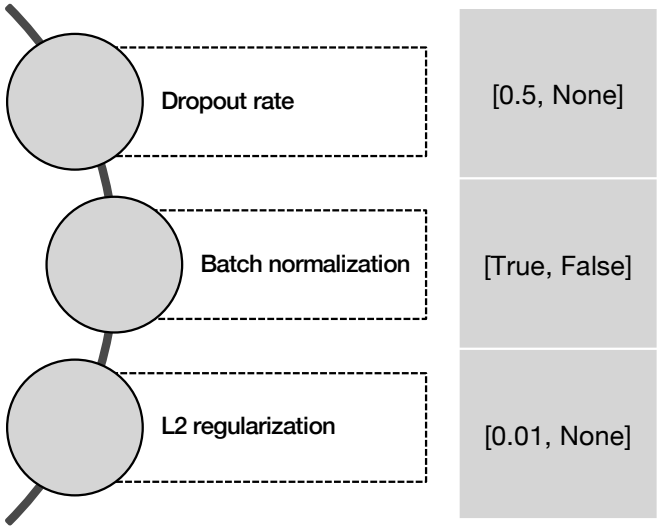
Modeling – Transfer Learning

Leveraging Pre-Trained Models: Exploring Transfer Learning

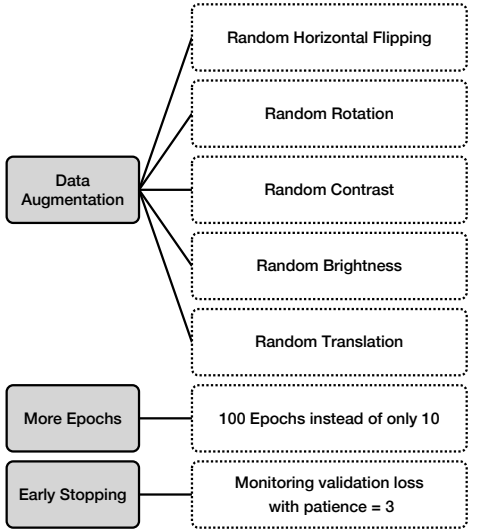
Model Exploration & Selection



Bayesian Optimization

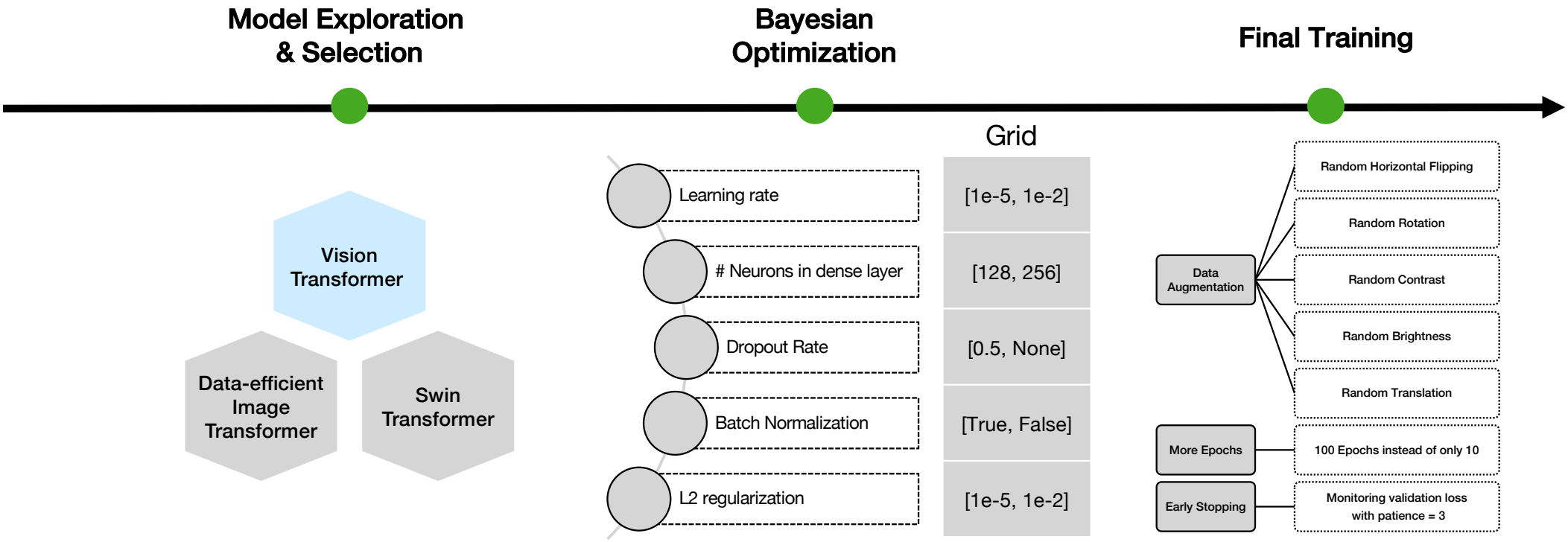


Final Training



Modeling – Transformer

Further Boosting our Model's Performance with a State-of-the-Art Transformer



Performance

Metrics for Evaluation

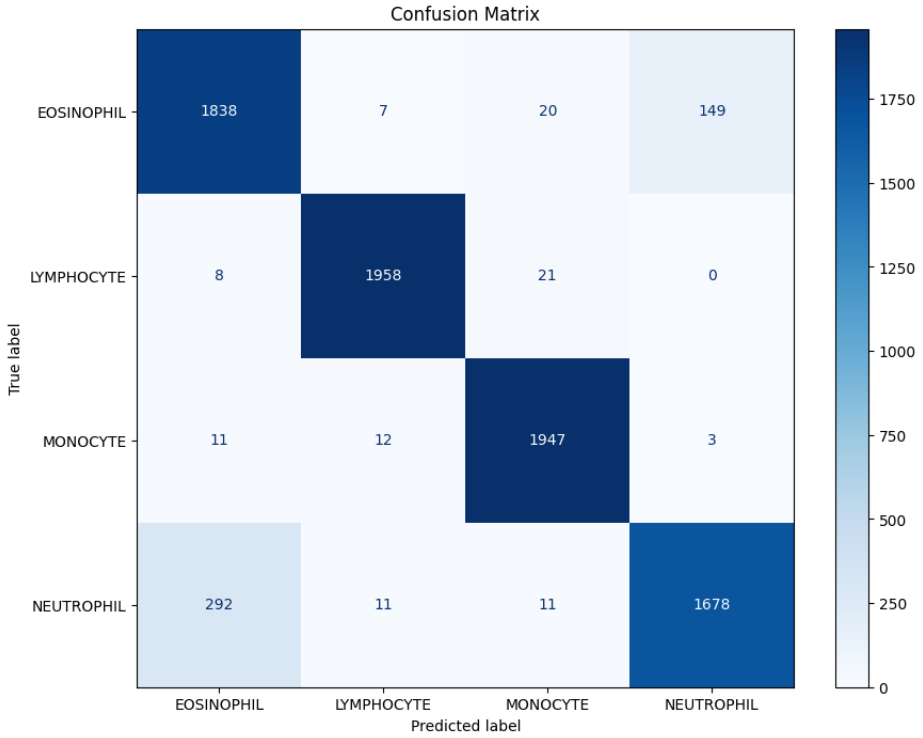
Metric	Accuracy	Precision	Recall	F1-Score
Definition	• Proportion of true results (both true positives and true negatives) among the total number of cases examined. $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$	• Proportion of true positives among the total positive predictions. $Precision = \frac{TP}{TP + FP}$	• Proportion of true positives among the total actual positives. $\text{Recall} = \frac{TP}{TP + FN}$	• Harmonic mean of precision and recall. $F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$
Importance	• General Measure of model performance • Insufficient in case of class imbalance due to skewed results • Risk of disease cells being misclassified	• Critical for minimizing false positives, which is especially important for medical diagnostics to prevent inappropriate treatments.	• Ensures that as many actual positive cases as possible are detected, essential for not missing critical diagnostic information.	• Balanced measurement of the overall model's performance in handling both false positives and false negatives.
Approach	Rather than evaluating each model immediately after its implementation, we plan to systematically log their performance metrics during training using TensorBoard. This allows us to efficiently compile and review all relevant performance data at the end of our modeling phase.			

In summary, while accuracy gives a general idea of model performance, it does not provide detailed insights into our model's behavior concerning false positives and false negatives. Precision and recall are more informative metrics in medical diagnostics, ensuring that the model is both specific and sensitive to the presence of diseases.

Results – Training

Performance of the best model

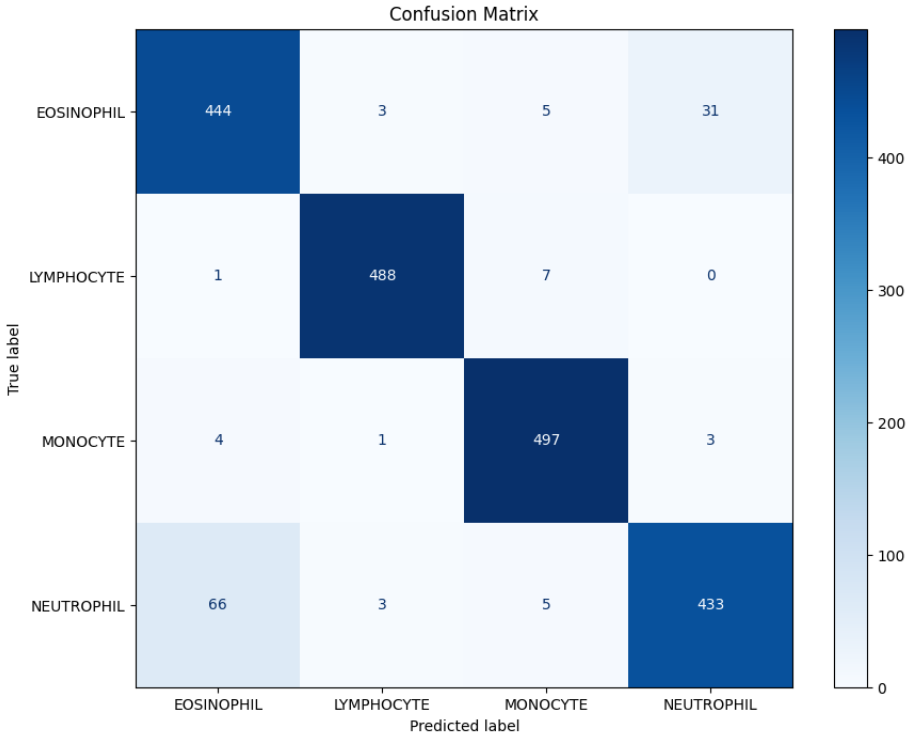
Accuracy = 0.93	Precision	Recall	F1-Score
Eosinophil	0.86	0.91	0.88
Lymphocyte	0.98	0.99	0.99
Monocyte	0.97	0.99	0.98
Neutrophil	0.92	0.84	0.88



Results – Validation

Performance of the best model

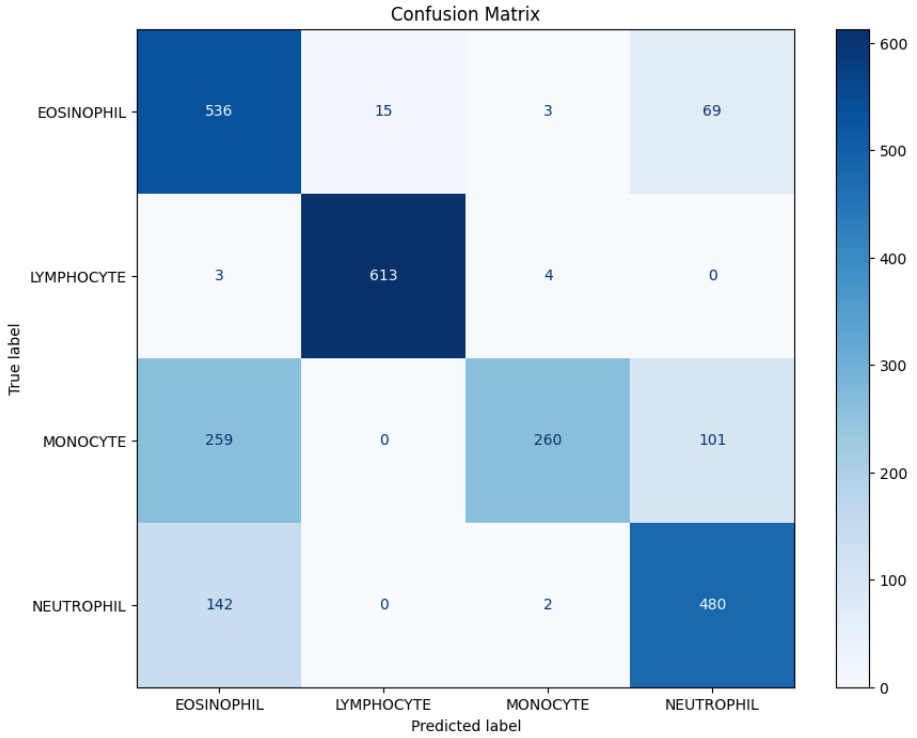
Accuracy = 0.94	Precision	Recall	F1-Score
Eosinophil	0.86	0.92	0.89
Lymphocyte	0.99	0.98	0.98
Monocyte	0.97	0.98	0.98
Neutrophil	0.93	0.85	0.89



Results – Test

Performance of the best model

Accuracy = 0.76	Precision	Recall	F1-Score
Eosinophil	0.57	0.86	0.69
Lymphocyte	0.98	0.99	0.98
Monocyte	0.97	0.42	0.58
Neutrophil	0.74	0.77	0.75



Conclusion

Outperforming CellVision Diagnostics

	CellVision	BloodCell
Accuracy	0.60	0.76
Precision	0.65	0.81
Recall	0.65	0.76
F1-Score	0.65	0.78

Performance

BloodCell AI outperforms competitors, offering superior results in cell classification and disease diagnosis.

Metrics

The superiority in all metrics ensures quicker and better diagnostics, enhancing medical decision-making and saving valuable time for healthcare professionals.

Quality of Patient Care

Adoption of BloodCell AI by medical facilities enhances the quality of patient care while optimizing resources, resulting in improved financials and operational efficiency.

Strategic Investment

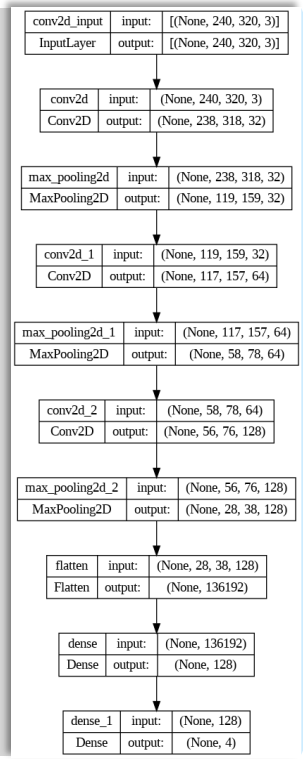
Incorporating BloodCell AI into medical practice represents a strategic investment in both patient outcomes and organizational effectiveness.

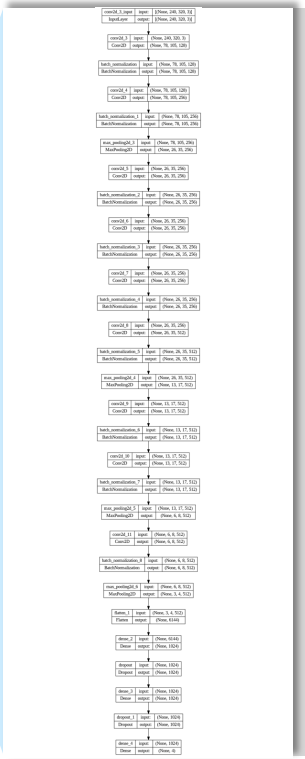
Modeling

Model Selection (1/2 – Baseline and Benchmark)

Baseline Model

- **Convolutional Layers**: Three layers with filters increasing from 32 to 128, all using 3×33×3 kernels and ReLU activation. Each convolutional layer is followed by a 2×22×2 max pooling layer to reduce spatial dimensions.
- **Flatten Layer**: Converts the 3D output to a 1D vector for the dense layers.
- **Dense Layers**: One dense layer with 128 units and ReLU activation, followed by an output layer with softmax activation designed for multi-class classification based on the number of classes (num_classes).



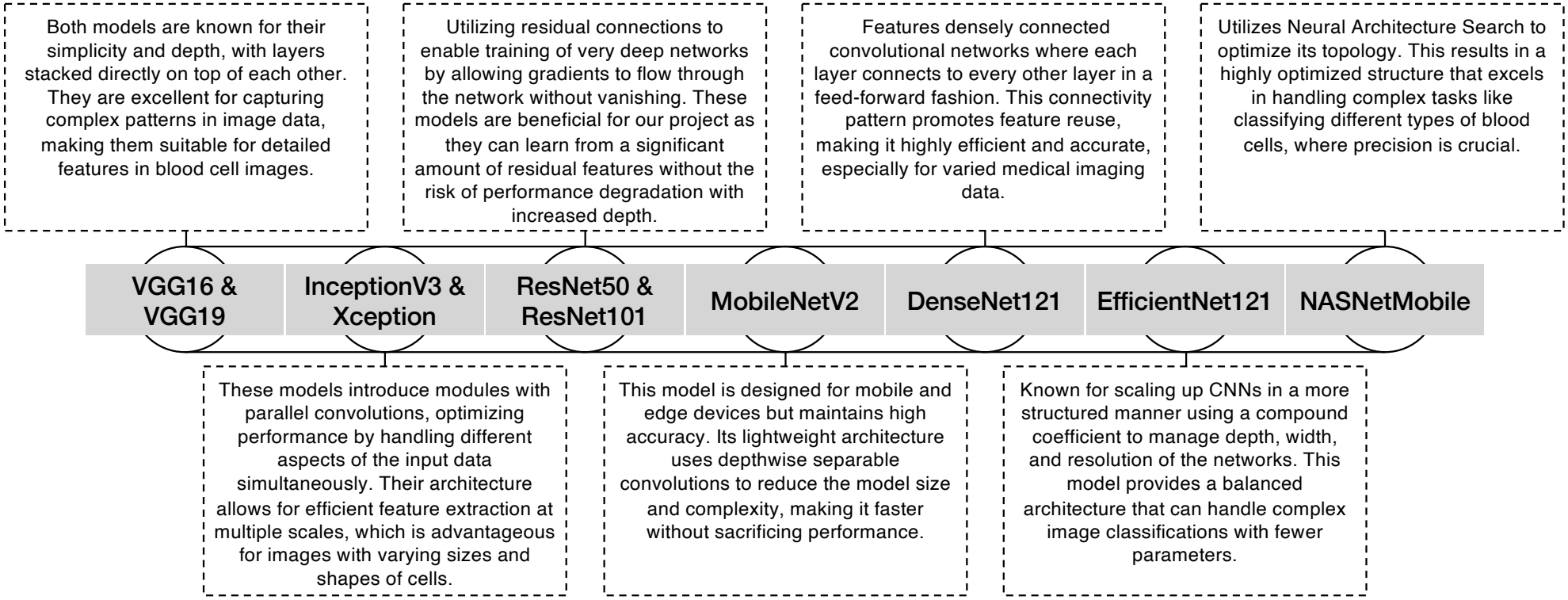


Benchmark Model (Kaggle)

- **Convolutional Layers**: Multiple layers with increasing filter sizes from 128 up to 512, using 3×33×3 to 8×88×8 kernels, ReLU activation, and same padding. Each layer is followed by batch normalization.
- **Pooling Layers**: Max pooling applied after specific convolution layers to reduce spatial dimensions.
- **Flatten Layer**: Converts the multi-dimensional output of the last convolutional layers into a flat vector.
- **Dense Layers**: Two dense layers, each with 1024 units and ReLU activation, separated by 50% dropout layers to prevent overfitting.
- **Output Layer**: Softmax layer with 4 outputs for multi-class classification.

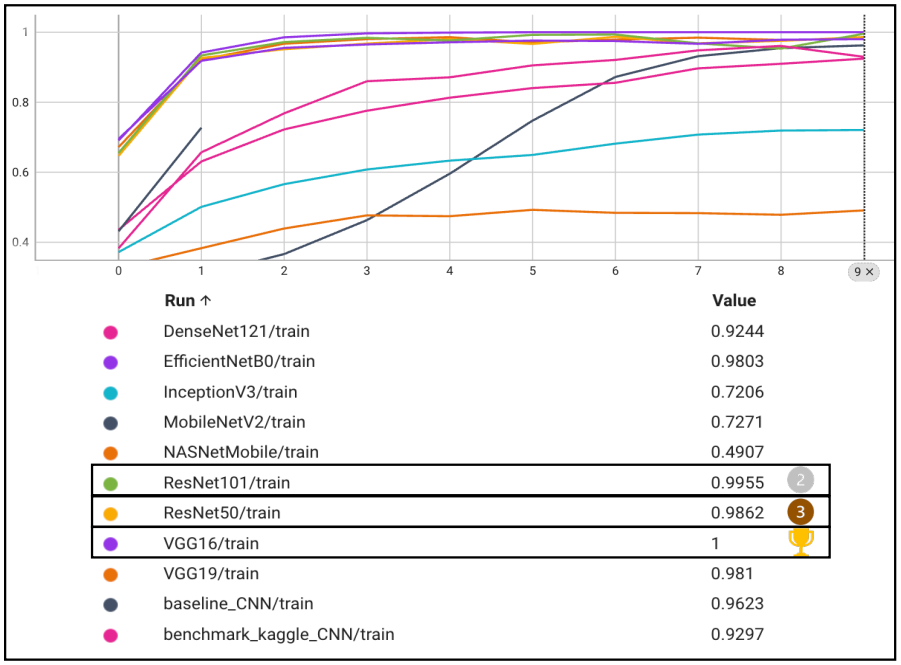
Modeling

Model Selection (2/2 – Transfer Learning)

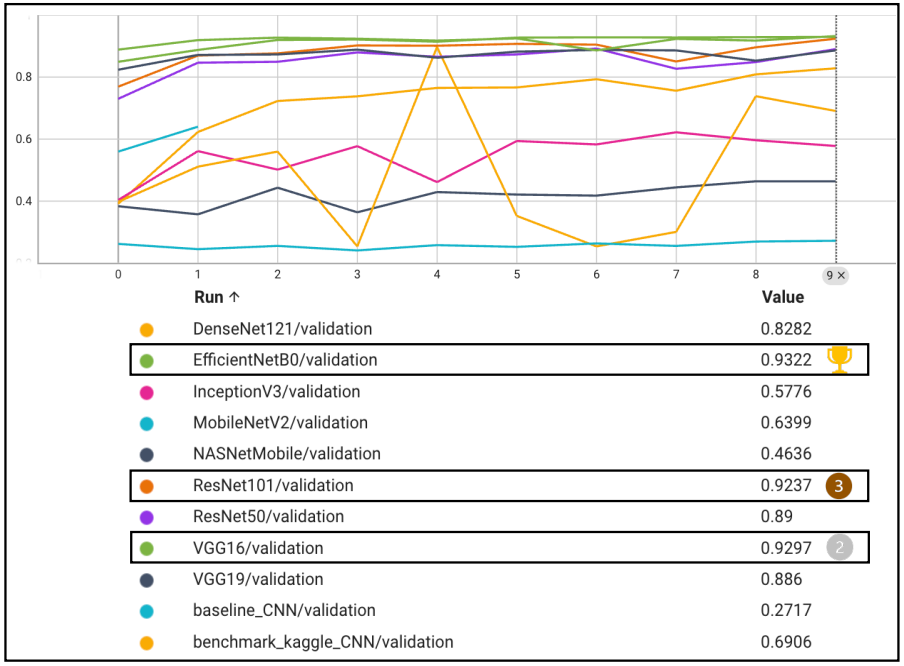


Modeling

Model Selection (2/2 – Transfer Learning)



Run ↑	Value
DenseNet121/train	0.9244
EfficientNetB0/train	0.9803
InceptionV3/train	0.7206
MobileNetV2/train	0.7271
NASNetMobile/train	0.4907
ResNet101/train	0.9955
ResNet50/train	0.9862
VGG16/train	1
VGG19/train	0.981
baseline_CNN/train	0.9623
benchmark_kaggle_CNN/train	0.9297

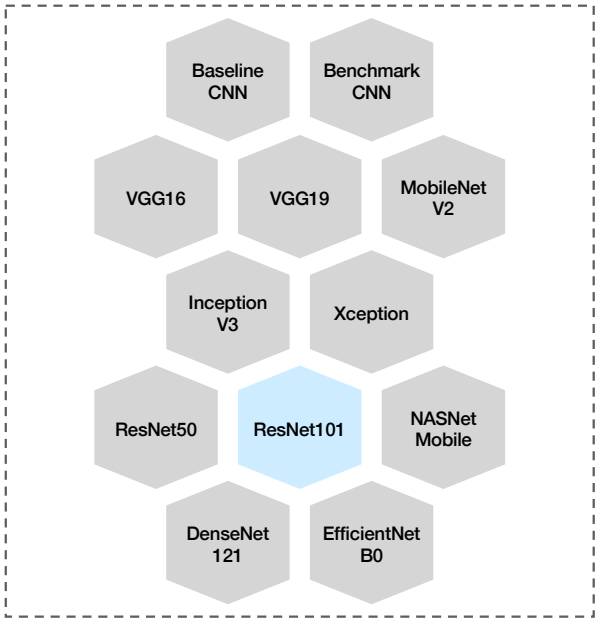


Run ↑	Value
DenseNet121/validation	0.8282
EfficientNetB0/validation	0.9322
InceptionV3/validation	0.5776
MobileNetV2/validation	0.6399
NASNetMobile/validation	0.4636
ResNet101/validation	0.9237
ResNet50/validation	0.89
VGG16/validation	0.9297
VGG19/validation	0.886
baseline_CNN/validation	0.2717
benchmark_kaggle_CNN/validation	0.6906

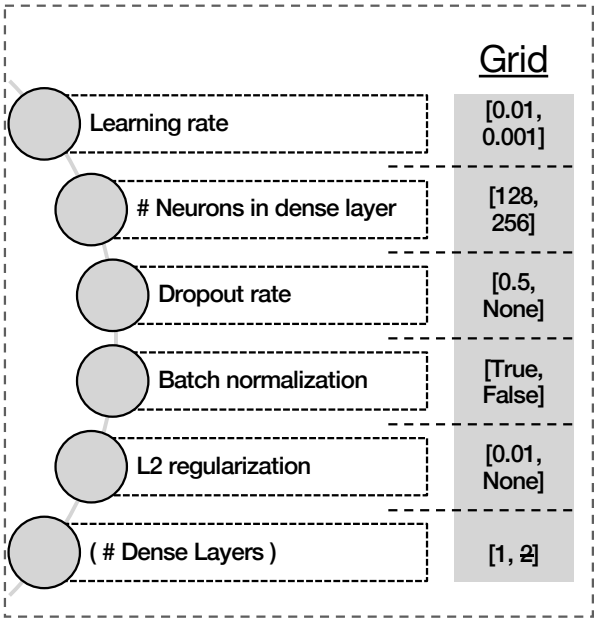
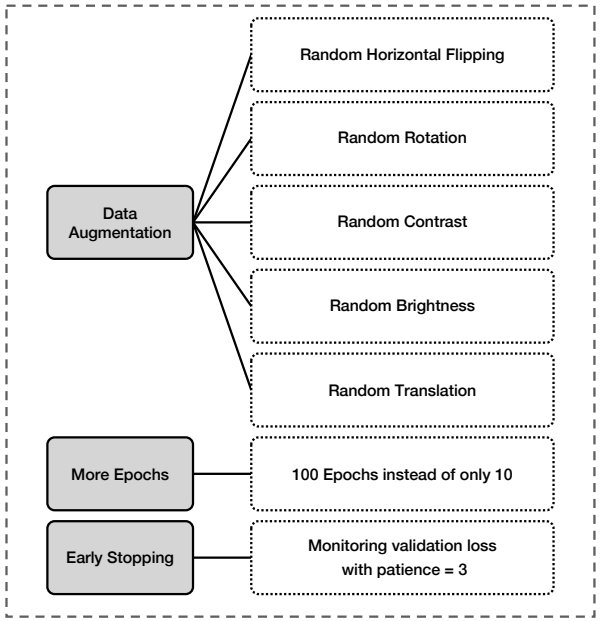
Modeling

Fine Tuning the Best Model

Model Selection



Hyperparameter Optimization



Final Training