Leukemia Detection Using Convolutional Neural Networks

Jumana Lightwala

[C0892358@mylambton.ca](mailto:C0892358@mylambton.ca) Bibek Shiwakoti

[C0898100@mylambton.ca](mailto:C0898100@mylambton.ca)Ronak Shah

[C0891297@mylambton.ca](mailto:C0891297@mylambton.ca)Pratikkumar Mishra

[C0891298@mylambton.ca](mailto:C0891298@mylambton.ca)

***Abstract***

***This paper presents a deep learning-based approach for detecting leukemia using convolutional neural networks (CNNs). Leveraging a dataset provided by Mehradaria on Kaggle, we preprocess, visualize, and model the image data using TensorFlow and Keras. The proposed CNN models achieve high accuracy, demonstrating the potential for effective automated leukemia detection. Three models were developed and evaluated, each showcasing various strengths and weaknesses.***

***Keywords***

***Leukemia Detection, Convolutional Neural Networks, Deep Learning, TensorFlow, Keras.***

**I. Introduction**

Leukemia is a type of cancer that affects blood and bone marrow. Early detection is crucial for effective treatment and improved patient outcomes. Traditional methods of leukemia detection rely heavily on manual examination by experts, which can be time-consuming and prone to human error. In this study, we propose a convolutional neural network (CNN) model to automate the detection process, enhancing accuracy and efficiency.

**II. Dataset**

The dataset used for this project was sourced from Kaggle, provided by Mehradaria. It consists of images labeled into four classes: Benign, Early Pre-B, Pre-B, and Pro-B. The data was split into training and validation sets to train and evaluate the models.

A white screen with black text

Description automatically generated

*Fig. 1. About the Dataset*

**III. Related Studies**

Numerous studies have leveraged deep learning techniques to improve cancer detection, particularly leukemia. Zhang et al. (2020) provided a comprehensive review on deep learning applications in cancer detection and diagnosis, highlighting various neural network architectures and their effectiveness across different medical imaging domains [1]. The pioneering work by Krizhevsky et al. (2012) introduced deep convolutional neural networks (CNNs) for large-scale image classification, which laid the foundation for subsequent advancements in medical image analysis [2].

Specifically for leukemia detection, Shafique and Tehsin (2018) demonstrated the use of CNNs to accurately distinguish between different types of leukemia in blood samples. Their work emphasized the capability of deep learning models in automating diagnostic tasks in hematology [3]. Kumar et al. (2021) further expanded on this by proposing an automated leukemia detection system using deep CNNs, integrating image processing and classification techniques tailored for blood smear analysis [4].

Moreover, Litjens et al. (2017) conducted a thorough survey on deep learning methodologies in medical imaging, underscoring how these approaches have revolutionized image analysis in various medical fields, including oncology and pathology [5].

**IV. Methodology**

**A. Data Preprocessing**

The images were preprocessed using the following steps:

1. **Resizing**: Images were resized to 224x224 pixels to standardize input dimensions.
2. **Normalization**: Pixel values were normalized to scale the data between 0 and 1, aiding in faster convergence during training.
3. **Data Augmentation**: Techniques such as rotations, shifts, and flips were applied to increase data diversity and reduce overfitting.

**B. Model Architectures**

Three CNN models were designed with varying architectures to explore their effectiveness:

1. **Original Model**: A simple CNN with multiple convolutional layers followed by max-pooling layers. Dense layers are added towards the end, concluding with a softmax activation function for classification. This model serves as a baseline.
2. **Segmented Model**: Similar to the original but with slight modifications to the architecture for improved feature extraction and better generalization.
3. **Combined Model**: Integrates features from both the original and segmented models, aiming to leverage the strengths of both architectures for improved performance.

All models included batch normalization and L2 regularization to enhance training stability and prevent overfitting. Dropout layers were also added to further mitigate overfitting.

**C. Training**

The models were trained using the Adam optimizer with an initial learning rate of 0.001. A ReduceLROnPlateau callback was employed to reduce the learning rate when the validation loss plateaued. The training was conducted for a maximum of 25 epochs with early stopping based on validation loss to prevent overfitting.

**V. Results**

**A. Model Performance**

1. **Original Model**: The Original Model ran for 25 epochs, achieving a training accuracy of 99.99% and a validation accuracy peaking at 95.84%. Training loss decreased from 6.2 to 0.1948, and validation loss decreased from 14 to 0.4442. Learning rate adjustments occurred at epochs 8 and 21, leading to improved validation performance. Notable improvements in validation accuracy were observed after learning rate reductions.

A graph of loss and loss

Description automatically generated with medium confidence

*Fig. 2. Original Model Training and Validation Curves*

1. **Segmented Model**: The Segmented Model accurately classified blood cell development stages and benign cells over 25 epochs. Training accuracy improved from about 70% to nearly 100%, and validation accuracy increased from around 25% to about 85%. Training loss decreased from about 6.2 to 0.27, while validation loss decreased from about 14 to 0.7. One misclassification was observed, with the model confidently predicting most samples.

A graph of different sizes and colors

Description automatically generated with medium confidence

*Fig. 3. Segmented Model Training and Validation Curves*

1. **Combined Model**: The Combined Model also accurately classified blood cell development stages and benign cells over 25 epochs. Training accuracy improved from 68% to nearly 100%, and validation accuracy increased from 15% to about 90%. Training loss decreased from 5.5 to 0.09, while validation loss decreased from 3.4 to 0.48. Learning rate adjustments occurred at epochs 6, 10, and 17, aiding in performance improvements.

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

Fig. 4. Combined Model Training and Validation Curves

**B. Visualization of Predictions**

Sample predictions made by the models are shown in Figure 5. The models correctly classified the majority of the samples in the final Combined model, demonstrating their effectiveness.

A collage of images of blood cells

Description automatically generated

*Fig. 5. Sample Predictions from the Combined Model*

**C. Correlation Analysis**

A correlation matrix was created to analyze the similarities between the three models' predictions. The matrix revealed perfect self-correlation (1.00) for all models. There was a weak positive correlation (0.16) between the Original and Segmented models and a strong positive correlation (0.65) between the Original and Combined models. The Segmented and Combined models showed a moderate positive correlation (0.48). The Original model had the highest correlation with the Combined model, while the Segmented model had the lowest correlation with the other models. Overall, the Combined model exhibited moderate to strong correlation with both the Original and Segmented models..

A screenshot of a computer screen

Description automatically generatedFig. 6. Correlation Matrix of Model Predictions

**D. Comparison of Accuracy and Loss**  
All models showed improvement in accuracy and a decrease in loss over time. The Combined model achieved the highest stable validation accuracy and the lowest overall loss. The Segmented model exhibited the most erratic validation accuracy and the highest initial loss spike. Training accuracy reached near-perfect levels for all models, with the Original and Combined models improving the fastest. Validation metrics displayed high variability across all models, indicating signs of overfitting, especially in the Segmented model. By the final epochs, all models converged to similar low loss values, with the Combined model demonstrating the best overall performance and stability.

A graph of different colored lines

Description automatically generated with medium confidence

Fig. 7. Comparison of Accuracy and Loss Across Models

**VI. Discussion**

The results indicate that the proposed CNN models are highly effective in detecting leukemia. However, there are areas for improvement, particularly in enhancing the models' performance on the validation set. The

correlation analysis shows that the combined model is more similar to the segmented model than the original model, suggesting that the combined model leverages features from both to achieve better performance.

**VII. Conclusion**

This study demonstrates the potential of CNNs for automated leukemia detection. The proposed models achieve high accuracy and can significantly aid in the early detection of leukemia, potentially saving lives through timely intervention. Future work will focus on optimizing the model architecture and exploring more advanced techniques such as transfer learning.

**References**

[1] H. Zhang, et al., "Deep Learning for Cancer Detection and Diagnosis: A Comprehensive Review," *IEEE Transactions on Medical Imaging*, vol. 39, no. 11, pp. 3734-3747, 2020.  
[2] A. Krizhevsky, et al., "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012.  
[3] S. Shafique and S. Tehsin, "Detection of Leukemia and Its Types in Human Blood Sample Using Convolutional Neural Networks (CNN)," in *IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1-6, 2018.  
[4] S. Kumar, et al., "Automated Detection of Leukemia in Blood Smear Images Using Deep Convolutional Neural Networks," in *Lecture Notes in Computational Vision and Biomechanics*, Springer, 2021.  
[5] G. Litjens, et al., "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017.  
[6] M. Mehradaria, "Leukemia Detection Dataset," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/mehradaria/leukemia-detection>. [Accessed: Aug. 10, 2024].