International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Automated Detection of Acute Lymphoblastic Leukemia Using AI

Dr. Naveenkumar N[1], Mr. Praveen Reddy C[2], Ms. Madhumathi P[3], Mr. Praveen Kumar P[4], Mr. Gnana Prakash T[5]

[1] Associate Professor, Dept. of Computer Science and Engineering, Muthayammal Engineering College, Rasipuram, India
[2–5] Students, Dept. of Computer Science and Engineering, Muthayammal Engineering College, Rasipuram, India

ABSTRACT :

In recent years, advancements in artificial intelligence—especially machine learning and deep learning—have brought substantial improvements to medical diagnostics. This study presents a system named AutoLeuk, designed to automate the diagnosis of Acute Lymphoblastic Leukemia (ALL). Utilizing Convolutional Neural Networks (CNNs) along with transfer learning, AutoLeuk can significantly reduce the cost and time required for diagnosis, especially compared to conventional methods like fluorescence in situ hybridization (FISH), which are often inaccessible in resource-limited healthcare environments.

The system introduces a two-stage fuzzy color segmentation technique to effectively isolate white blood cells (leukocytes) from other components in blood smear images. Key features such as the Hausdorff dimension and contour signatures are extracted and used to train the CNN for accurate classification. This AI-powered approach not only simplifies the diagnostic process but also makes it more affordable and accessible.

Keywords: CNN, Acute Lymphoblastic Leukemia (ALL), FISH, Transfer Learning, Image Segmentation

Introduction

AutoLeuk focuses on early detection of Acute Lymphoblastic Leukemia using deep learning, particularly CNNs and transfer learning. ALL is a severe form of blood cancer, most prevalent in children, and detecting it early is vital but challenging. Traditional diagnostic tools are often costly and labor-intensive.

Our system addresses these challenges by analyzing blood smear images with intelligent preprocessing techniques. The fuzzy-based segmentation isolates leukocytes, improving the detection process. Features like contour signatures and Hausdorff dimensions are used to train a fine-tuned CNN, helping the model distinguish between healthy and abnormal cells even with a limited dataset of 108 blood smear images.

Literature Review

The accuracy of leukemia detection greatly depends on effective white blood cell (WBC) segmentation. Techniques have evolved from basic image processing methods to sophisticated machine learning and deep learning strategies. Traditional methods include thresholding, morphological operations, and watershed algorithms. Recently, CNN-based models and encoder-decoder architectures such as U-Net have shown superior performance in complex segmentation tasks.

Proposed System

The proposed AI system incorporates the following pipeline:
- Preprocessing: Standardizing images through normalization, resizing, and applying data augmentation (rotation, flipping, zooming) to make the model more robust.
- Segmentation: Advanced methods isolate leukocytes from surrounding cells, improving feature detection.
- Feature Extraction: CNNs learn visual patterns essential for distinguishing leukemia cells.
- Classification: A deep learning model trained on processed data identifies signs of ALL.

Data Preprocessing

To prepare the data for analysis:
- Image Resizing: Standardized to dimensions like 224×224 to maintain uniformity.

- Normalization: Pixel values are scaled between 0 and 1 for consistency.
- Augmentation: Synthetic variations are created to avoid overfitting.
- Noise Reduction: Filters (e.g., Gaussian, median) help remove background artifacts and staining irregularities.

Dataset

We used a curated collection of blood smear images sourced from open medical repositories such as the Cancer Imaging Archive (TCIA). The dataset contains images stained using techniques like Wright's stain and is annotated to highlight leukocyte regions. It includes thousands of images showing various stages of leukemia, helping to train and validate the system under diverse conditions.

System Architecture

Our architecture supports both training and testing phases, integrating preprocessing, segmentation, CNN-based classification, and result interpretation. Transfer learning using models like ResNet and VGG16 plays a critical role in improving accuracy, especially with smaller datasets.

Methodology

Although deep learning offers impressive accuracy, the lack of interpretability hinders its adoption in clinical practice. Tools like Grad-CAM help visualize decision-making but aren't always sufficient. We address this gap by combining interpretable AI methods with robust performance, aiming to make our system trustworthy for medical professionals.

Feature Extraction

Advanced segmentation with attention-based U-Net models enhances the recognition of subtle features in white blood cells. Real-time processing is also explored for quicker diagnoses, and longitudinal image analysis is proposed for tracking disease progression over time.

Challenges in Detection

Despite its promise, deep learning still faces hurdles:
- High dependency on large labeled datasets
- Computational intensity
- Poor generalization across datasets
- Risk of overfitting, particularly with transfer learning

Future Enhancements

Enhancements will focus on:
- Leveraging transformer-based and multi-modal architectures that integrate image and clinical data
- Expanding the dataset with more diverse images to improve model generalizability
- Incorporating real-time capabilities for clinical deployment

Results

Our experiments demonstrate that using pre-trained networks like ResNet and VGG16, combined with thorough preprocessing, significantly improves detection accuracy. The results validate the effectiveness of the proposed approach in limited-data settings.

Conclusion

The AutoLeuk system showcases the potential of AI in transforming leukemia diagnosis. By replacing slow, manual processes with fast, automated analysis, it supports early intervention and better treatment outcomes. As deep learning continues to evolve, future systems can be more accurate, efficient, and interpretable—making them invaluable tools in modern healthcare.