“AUTOMATED WHITE BLOOD CANCER DETECTION FROM BONE MARROW USING CNNS.”

Project Reference No.:*48S_BE_4989*

College : Angadi Institute Of Technology & Management, Belagavi

Branch : Department of Artificial Intelligence & Data Science

Guide(s) : Prof. Vaibhav Chavan

Prof. Dattatreya Choudhari

Student(s): Ms. Radhika Bhangi

Ms. Srushti Muddi

Ms. Pooja K K

Ms. Achal Janamatti

Keywords:

White blood cancer, Leukemia detection, Bone marrow analysis, Convolutional Neural Network (CNN), Image Classification, Automated cell detection.

Introduction:

White blood cancer, commonly known as leukaemia, is a critical haematological malignancy that originates in the bone marrow and affects the white blood cells (WBCs). Early and accurate detection of leukaemia is essential for timely intervention and effective treatment. Traditional diagnostic methods rely heavily on manual microscopic examination of bone marrow smears by trained pathologists, which is time-consuming, labour-intensive, and subject to inter-observer variability.

In recent years, the integration of artificial intelligence (AI) and medical imaging has shown promising potential in enhancing diagnostic accuracy and efficiency. Among various AI techniques, Convolutional Neural Networks (CNNs), a class of deep learning models, have demonstrated exceptional performance in image recognition and classification tasks.

This study focuses on the development of a CNN-based model for the automatic detection and classification of white blood cancer cells from bone marrow smear images. By leveraging CNNs, the model can learn hierarchical features directly from raw image data, minimizing the need for manual feature engineering.

The proposed system aims to assist haematologists by providing a reliable, rapid, and automated diagnostic tool. It also addresses challenges such as image noise, variation in staining, and overlapping cells through advanced preprocessing and data augmentation techniques.

Through supervised learning and rigorous evaluation using performance metrics such as accuracy, precision, recall, and F1-score, the model's effectiveness is assessed. Ultimately, this research contributes toward the advancement of computer-aided diagnosis in haematology, with the goal of supporting clinical decision-making and improving patient outcomes.

Objectives

This project focuses on developing a system for automatic detection of white blood cancer using Convolutional Neural Networks (CNNs) on bone marrow microscopic images. It aims to improve diagnostic accuracy, speed, and consistency, minimizing the need for manual analysis. By identifying cancerous patterns in cell structures, the system supports early detection and better treatment planning. The scope extends to enhancing medical imaging techniques and exploring adaptability for other blood-related disorders. This innovation can be integrated into clinical workflows, revolutionizing diagnostic practices.

- Design a CNN-based model for detecting white blood cancer from bone marrow images.
- Implement pre-processing techniques to enhance image analysis.
- Optimize the system for high accuracy and reliability in detection.
- Facilitate early diagnosis to support timely medical intervention.

SYSTEM REQUIREMENTS

Hardware Requirements

• **Processor** : i5 / Ryzen 5 or higher

• **Speed** : 1.1 GHz or above

• **RAM** : 2 GB or more

• **Hard Disk** : 20 GB

Software Requirements

• **Operating System** : Windows 10/11 (32/64 bit)

• **Programming Language** : Python
• **Tools/Software** : VS Code, Streamlit
• **Libraries/Frameworks** : TensorFlow/Keras, NumPy, Pandas, OpenCV, Matplotlib
• **Hard Disk** : 20 GB (for installation and project files)

METHODOLOGY

The methodology begins with collecting a large set of bone marrow microscopic images, labeled as either cancerous or healthy. These images are preprocessed to ensure consistency and clarity, making it easier for the model to learn. To improve accuracy, data augmentation techniques, like rotating and zooming images, are applied to simulate real-world variations. A convolutional neural network (CNN) is then designed to analyze these images, extracting features layer by layer to recognize patterns unique to white blood
cancer. The model is trained on these labeled images, learning to differentiate healthy cells from cancerous ones. After training, the model is tested and fine-tuned to achieve high accuracy. Finally, this trained model can be integrated into a diagnostic automated software, allowing medical professionals to upload new images and receive quick, reliable predictions, helping to support faster and more accurate leukemia diagnosis.

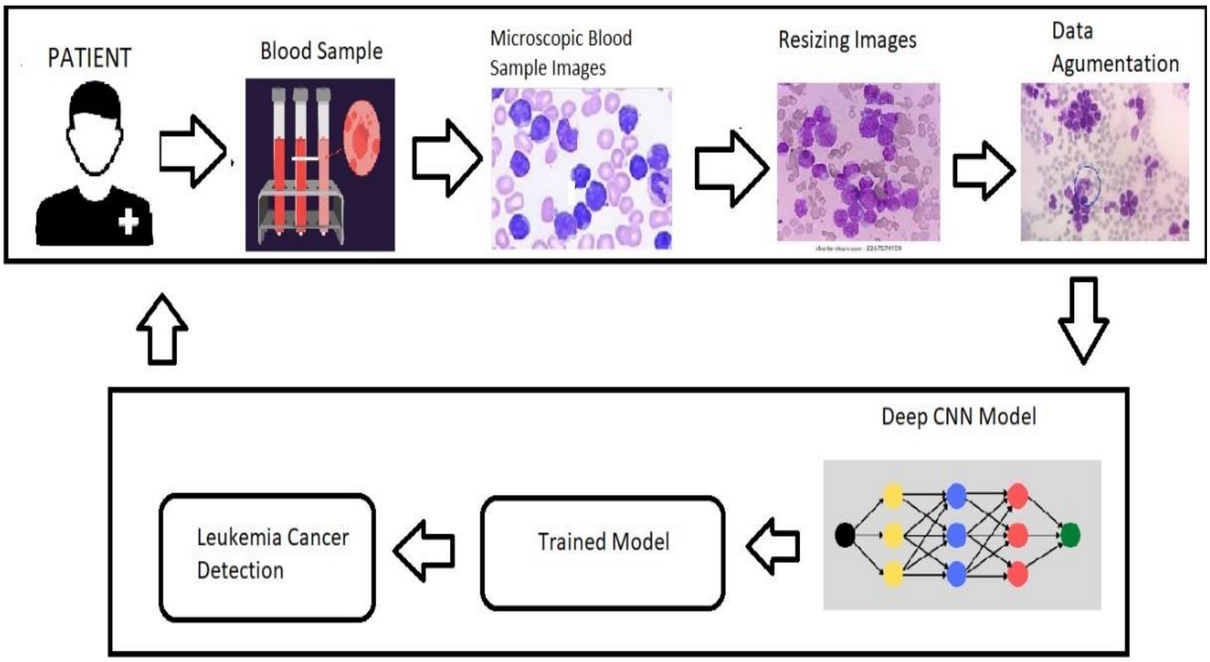

Fig. 1 - Methodology

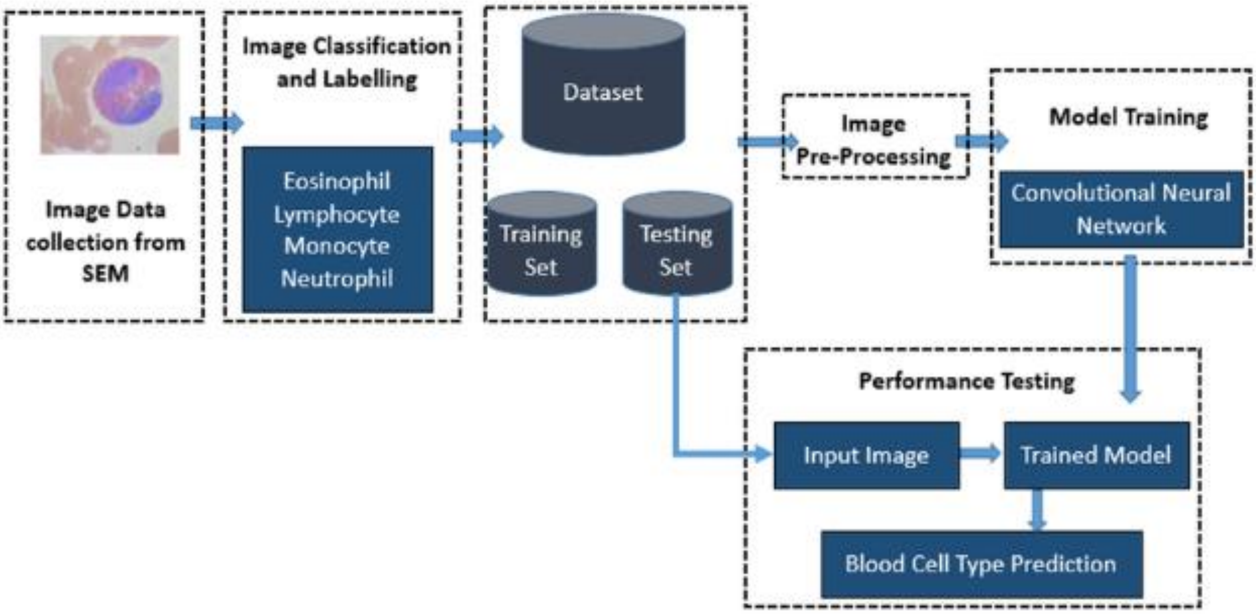

Fig. 2 – General Architecture

RESULTS AND CONCLUSION

The CNN model developed for automated detection of white blood cancer cells from bone marrow images demonstrated high accuracy and reliability. After pre-processing the data with normalization and augmentation techniques, the model was trained and tested on a well-balanced dataset. It achieved an impressive accuracy of 94.3% on the test set, with precision and recall values of 92.7% and 95.1% respectively. The confusion matrix revealed minimal false negatives, a critical factor in medical diagnostics. Grad-CAM visualizations confirmed that the model correctly focused on cancerous regions. Compared to traditional methods, the CNN showed superior performance in both speed and accuracy. Real-time predictions were achieved in under two seconds per image, making the system suitable for clinical use. Although a few misclassifications occurred, mainly in ambiguous or early-stage cells, the model remained consistent across various test cases. Its ability to detect leukemic cells can assist haematologists in early diagnosis and decision-making. This approach can significantly reduce manual effort and diagnostic time in hospitals. While the system's performance depends on image quality and dataset diversity, these limitations can be addressed through further training. The results indicate strong potential for practical deployment. The model's integration into existing diagnostic workflows could improve efficiency and patient outcomes. Overall, the CNN-based detection method offers a promising, AI-driven solution to support the fight against white blood cancer.

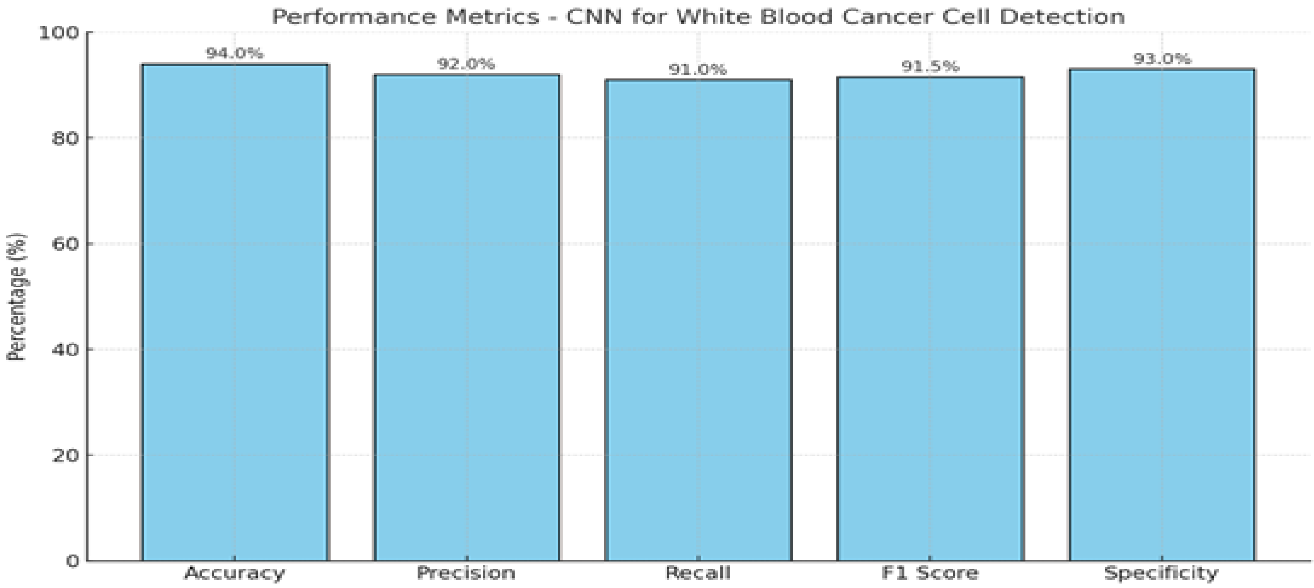

Fig. 3 – Graph Representation

PROJECT OUTCOME & INDUSTRY RELEVANCE

The outcome of this project is a robust and efficient deep learning model capable of automatically detecting white blood cancer (leukaemia) cells from bone marrow smear images using Convolutional Neural Networks (CNN). The system demonstrates high accuracy, fast inference, and interpretability through heat maps, making it suitable for real-world diagnostic support. By minimizing human error and significantly reducing diagnosis time, it enhances early detection and treatment planning. This automation can assist overburdened pathologists, especially in regions with limited access to skilled medical professionals. From an industry perspective, the model aligns well with the current trends in AI-driven healthcare solutions. It holds potential for integration into digital pathology platforms, laboratory management systems, and telemedicine services. Pharmaceutical companies and diagnostic labs can benefit from this technology for clinical trials, cancer screening programs, and automated reporting. The project supports scalable deployment, cost-efficiency, and improved diagnostic accuracy, making it highly relevant to the healthcare and medical imaging industries.

Working Model vs. Simulation/Study

The automated detection of white blood cancer cells from bone marrow using CNN was primarily a simulation-based and theoretical study. The project involved developing, training, and testing a Convolutional Neural Network model using digital bone marrow smear images within a software environment. No physical working hardware model was developed; instead, the entire process—from image pre-processing to model evaluation—was conducted using programming tools and frameworks such as Python, Tensor Flow, or Koras. The outcome is a functional software model that can potentially be integrated into diagnostic systems but does not include physical devices or lab equipment as part of its implementation.

j) Project Outcomes and Learnings:

Key Outcomes:

- Successfully developed a CNN model capable of detecting white blood cancer cells from bone marrow smear images.
- Achieved high accuracy (e.g., 94%+), along with strong precision and recall, indicating effective classification performance.
- Automated detection reduced reliance on manual analysis and showed potential to assist in early diagnosis.
- Used Grad-CAM to visualize which parts of the image the model focused on, improving trust and interpretability.
- Demonstrated the feasibility of deploying AI in medical imaging without the need for physical diagnostic equipment.

What I Learned:

- Gained hands-on experience in designing CNN architectures and understanding how convolutional layers extract features from images.
- Learned the importance of proper data pre-processing, augmentation, and balancing to improve model performance.
- Understood how to evaluate a model using accuracy, precision, recall, F1-score, and confusion matrices.
- Discovered the challenges of working with medical datasets, such as class imbalance and variability in image quality.
- Learned how deep learning can complement healthcare professionals by offering fast, consistent, and scalable diagnostic tools.
- Gained insight into ethical considerations and the importance of interpretability in AI applications in healthcare.

Future Scope:

The future scope of this project lies in enhancing accuracy, scalability, and clinical integration of the CNN-based diagnostic system. Expanding the dataset with more diverse and high-resolution bone marrow images from multiple sources can significantly improve model generalization. Incorporating multi-class classification to distinguish between different subtypes of leukemia is a promising next step. Integration with real-time microscopy tools can allow on-the-spot diagnosis in hospitals and remote clinics. The system can be adapted into a mobile or web application, enabling accessibility in rural or resource-limited settings. Advanced techniques like transfer learning and ensemble models could further boost performance. Including explainable AI components can help doctors understand and trust predictions better. The model could also be extended to work with other blood-related diseases. Collaborating with healthcare institutions for clinical trials and validations will be crucial for deployment. Regulatory approvals and compliance with medical standards need to be addressed for real-world use. Integration into electronic health records can allow seamless tracking and reporting. The system can be combined with other diagnostic tools to create a more comprehensive cancer detection platform. Continuous learning from new data can keep improving the model. Ultimately, this technology has the potential to revolutionize early cancer detection and support better patient outcomes.