Chapter 4

Book Chapter

Blood Cancer Detection and Classification using Deep Learning

Hanamant R Jakaraddi *[1], Sridevi G †[2], and Nithya K V ‡[3]

[1]Assistant Professor, Department of MCA, Acharya Institute of Technology, Bangalore
[2]Assistant Professor, Department of MCA , Jain College of Engineering, Belagavi
[3]Department of MCA, Acharya Institute of Technology, Bangalore

Abstract

Blood cancer-related diagnosis and analysis is still a difficult and time-consuming procedure. In the last ten years, many methods have been developed for the detection, analysis, and classification of blood cancer; nevertheless, no model or approach now in use completely automates the process of examining human blood cells to detect the presence of cancer. The development of this type of an automated system could revolutionize the identification and prevention of disease, greatly speeding up and enhancing the accuracy of medical diagnostics. A retrospective of the developments in research toward this objective is given in this chapter.

Keywords: Illness. Categorizing blood. Inspecting human. Prompter Medical.

*Email: hanamant2504@acharya.ac.in Corresponding Author
†Email: messridevi@gmail.com
‡Email: gowdanithya25@gmail.com

1 Introduction

Blood is created in the bone marrow and contains a number of essential elements that aid in immune response, oxygen transport, and healing.The most prevalent type of white blood cells are neutrophils, which serve as the first line of defense against bacterial and fungal diseases. Eosinophils play a role in controlling allergic reactions and combating parasite infections (Al-Azzawi et al., 2024). Less often seen basophils cause allergic reactions by releasing heparin and histamine, which causes inflammation and inhibits needless blood coagulation (see figure 1).Stem cells in blood, more especially hematopoietic stem cells (HSCs), are unique bone marrow-derived cells with the amazing capacity to differentiate into any kind of blood cell (Pirsadeghi et al., 2024). These stem cells give rise to red blood cells (RBCs), which carry oxygen, white blood cells (WBCs), which fight infection, and platelets, which help with clotting, through the process of hematopoiesis. Because HSCs are multipotent, or able to differentiate into numerous types of blood cells, there will always be an abundance of new cells to support a healthy immune system and circulatory system for the duration of a person's life (see figure 2). Blood contains The diameter of monocytes and macrophages ranges from 15 to 22 m, and their nuclei resemble pins and are derived from the mononuclear phagocytic system (Liu et al., 2024).Macrophages belong to the monocyte foreign body giant cell lineage. Macrophages are found in most tissues, where they identify, engulf, and degrade pathogens and cellular debris to perform vital immunomodulatory functions. White blood cells (WBCs), which include monocytes/-macrophages, polymorphonuclear leucocytes, mast cells, and their progenitors, make up 50% of the non-stromal cell population in bone marrow (Murayama et al., 2024).

Deep learning models can learn different degrees of abstraction in data representations since they are made up of several processing layers. The state-of-the-art in several areas, such as recognition of speech, visual object detection, drug discovery, genomics, and object recognition, has been greatly advanced by these models (Taye, 2023). The back propagation technique, which modifies the model's internal variables by determining how it is represented in every level according to the one before it, is a crucial component of this success. Deep learning is particularly good at seeing intricate patterns in large datasets.

While complex convolutional networks (CNNs) have transformed the processing of pictures, videos, voice, and audio, recurrent neural networks, or RNNs, have made tremendous progress in anomaly identification, especially in the context of blood cells (Choudhry et al., 2023). Conventional machine learning models developed feature extractors—which converted unprocessed input (such picture pixels) into internal models or feature vectors appropriate for the learning subsystem—by meticulous engineering and subject expertise. Nevertheless, deep learning streamlines this procedure by automation feature extraction, enabling more effective and efficient data analysis.

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

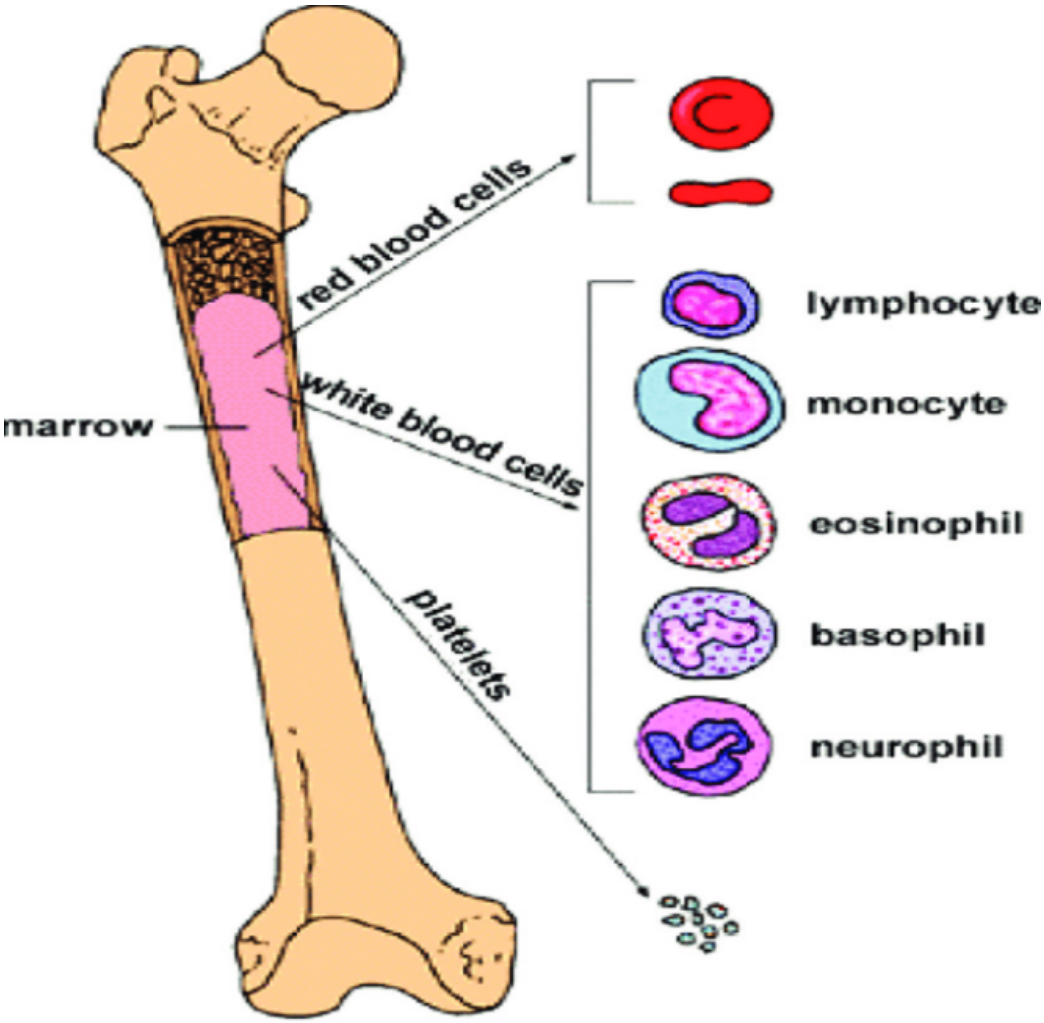

Figure 1. what is in our bones?

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

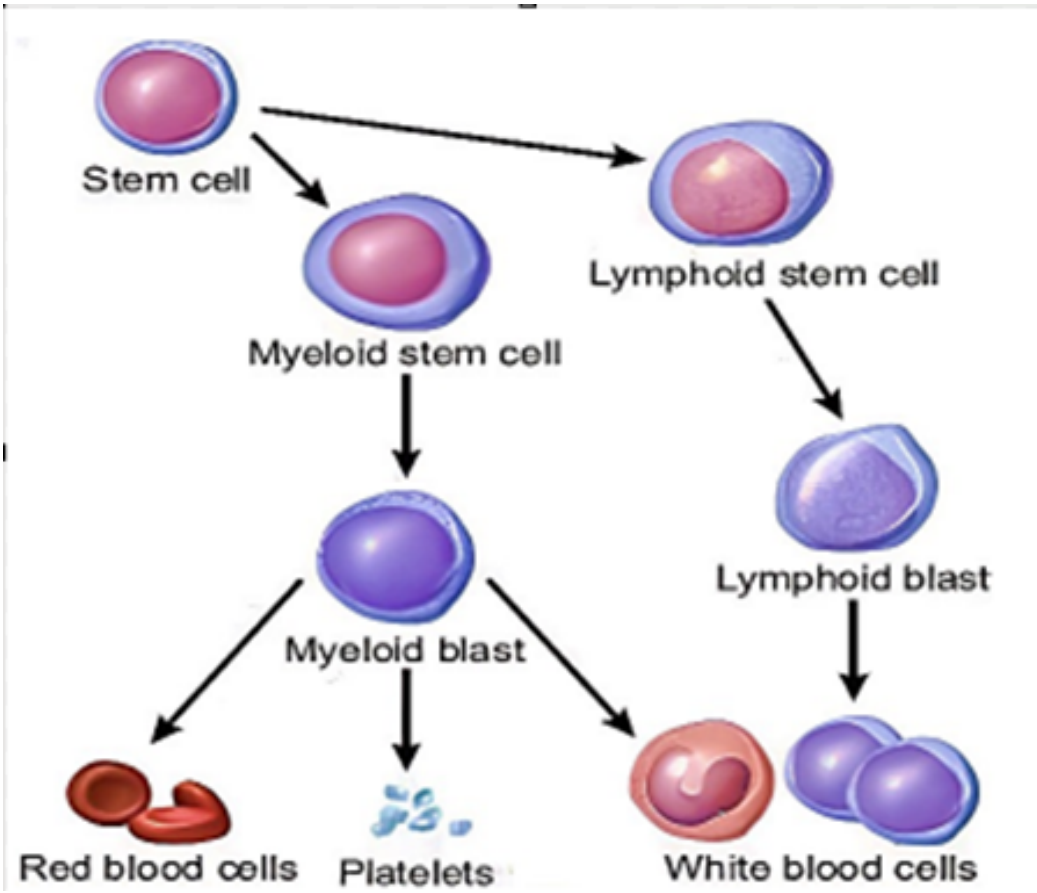

Figure 2. Progress in stem cell research

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

2 Deep Learning Models and Their Applications

2.1 Matrix Metalloproteinases (MMPs) and Carcinogenesis

One of the most well-known proteinase classes connected to the initiation of cancer is matrix metalloproteinases, or MMPs. New developments in technology have expanded our knowledge of MMPs' role as tumor microenvironment modulators. In addition to their function in the migration of cancer cells and the turnover of extracellular matrix, MMPs also affect signaling pathways linked to angiogenesis, inflammation, and cell proliferation. Interestingly, MMPs have non-proteolytic roles as well, which is making scientists reevaluate conventional cancer therapies (Koistinen et al., 2023).

2.2 Acute Leukemia Diagnosis Using Deep Learning

Both children and adults can be affected by acute leukemia, a potentially fatal disease that progresses quickly if treatment is not received. In youngsters, acute lymphoblastic leukemia (ALL) in particular spreads quickly and can be fatal in a matter of weeks (Faust et al., 2023). By using deep learning and image processing techniques, this research improves the diagnostic procedure for ALL and yields very accurate results. Using stained bone marrow images, a novel approach was presented to categorize ALL subtypes and reactive bone marrow (normal). In conjunction with sophisticated segmentation methods, a convolutional neural network (CNN) was utilized to generate accurate classification results.

2.3 Automatic Classification of Acute Lymphoblastic Leukemia

Efficient diagnostic approaches are necessary for white-blood cell leukemia, a serious disease that affects bone marrow and blood. Deep convolutional neural networks are used in this work to automatically identify acute lymphoblastic leukemia and classify its subtypes (L1, L2, L3, and Normal), which have been neglected in previous studies (Hassan Abbas Gondal et al., 2023). Rather of starting from zero and training an AlexNet model from scratch, data augmentation techniques were used to fine-tune the model for the dataset. In order to evaluate efficiency across different image formats, the dataset was also compared across different color models. In terms of diagnosing ALL, the system achieved 98.11% specificity, 80.50% accuracy, and 100% sensitivity. Furthermore, the model showed 100% specificity and sensitivity in categorizing EVERY subtype.

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

2.4 Multiple Myeloma (MM)

The clonal B-cell tumor known as multiple myeloma (MM) affects the plasma cells, a type of B-cells that have undergone terminal differentiation. Three distinct phases can be seen in the progression of multiple myeloma (MM): an active phase with a tiny proportion (<1%) of proliferating grow plasma cell types, an inactive phase with non-proliferating mature the plasma, cells, and a fulminant phase with increased plasmablastic cells and extramedullary proliferation (Bhaumik et al., 2023). Recent years have seen significant progress that has illuminated the essential components that cause neoplastic transformation in MM.

2.5 Platelet-Associated IgG (PAIgG) and Thrombocytopenia

There are several techniques for measuring platelet-associated IgG (PAIgG) levels. The antiglobulin intake assay, which analyzes IgG on platelets directly, is one sensitive and focused method. A straightforward fluorescent anti-IgG assay has also been developed recently, and it has a number of benefits. The results of these two PAIgG tests are compared in this study between individuals who have immune and non-immune thrombocytopenia and non-thrombocytopenic controls. The antiglobulin intake and fluorescence assays yielded negative results in 61 out of 62 cases and 54 out of 62 cases, respectively, among the non-thrombocytopenic controls. Eleven out of thirteen and eight out of thirteen assays in patients without immune thrombocytopenia were negative, respectively. However, utilizing the antiglobulin consumption assay, 54 out of 58 patients with immune thrombocytopenia, including ITP and SLE, tested positive.

3 Methodology

In the data collection process, it is crucial to obtain well-labeled, high-quality datasets from reliable sources such as hospitals, medical databases, or publicly accessible repositories like the Cancer Imaging Archive (TCIA). These datasets provide essential information for training models. For the training phase, the collected dataset is processed using a Convolutional Neural Network (CNN) algorithm to detect blood cancer cells and classify them into various types. The CNN's input layer requires a series of images, which can include MRI or PET scans, as well as microscopic images of blood smears. Prior to being fed into the network, these images undergo pre-processing steps, including augmentation, scaling, and normalization, to improve the model's performance and generalization ability. The convolutional layers in the CNN employ learnable filters, or kernels, to scan through the input images. These filters are designed to identify specific features, such as textures,

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

edges, and patterns within the images. To introduce non-linearity and enable the network to learn complex patterns, a Rectified Linear Unit (ReLU) activation function is applied after each convolution. The output layer of the CNN produces a likelihood score for each class, representing different blood cancer types. The model then makes a prediction based on the class with the highest likelihood score, enabling accurate classification of the cancer cells.

4 Architecture

The suggested method uses two Mixed Neural Networks (MNNs) and transfer learning to detect important elements in each image, suggesting an automated approach to blood cancer diagnosis (see figure 3). Convolutional Neural-Network (CNN) and Multi-Layer Perceptron (MLP) are the two MNN models that are employed in the technique. Multilayer perceptrons, in which every neuron in a layer is coupled to every other layer's neuron, are regularized versions of multilayer perceptrons. The networks are prone to over fitting due to the fact that they fully connected. Regularization techniques often include including a magnitude measurement of weights into the loss function. Consequently, MNNs are at the lower end of the accuracy and complexity spectrum. The MNN consists of four layers: the convolutional layer, which convolves the input and transfers the result to the subsequent layer; pooling layer to gradually shrink the representation's spatial extent due to lower the count of parameters and computation in the network, The data is first transformed into an all-dimensional array by the flattening layer before being sent to the fully connected layer, which links each neuron in one layer to every other layer's neuron. This technique is able to improve classification accuracy beyond what was previously possible by combining data from many abstraction levels that could be termed extra characteristics. Including feature maps at greater dimensions may improve the intermediate features' discriminative power and avoid the issue of network gradient (see figure 5).

I Dataset collection: The blood cancer dataset comprises individuals with a diagnosis of blood cancer as well as healthy individuals. The Kaggle website provided the data that was assembled in this collection. We can get medical imaging data using this data science-focused internet portal.

II Preprocessing: A range of data pre-processing techniques, such as zooming, shearing, flipping, and rescaling, are applied to the raw input training data. One side of an image, layer, selection, or path can be pushed in one direction while the other is being flipped by using the shear tool. Curved edges can also be made with this tool. More pixels are added to an image as it is zoomed in, increasing the image's total size. The hardest aspect of this technique is interpolating the new pixels utilizing

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

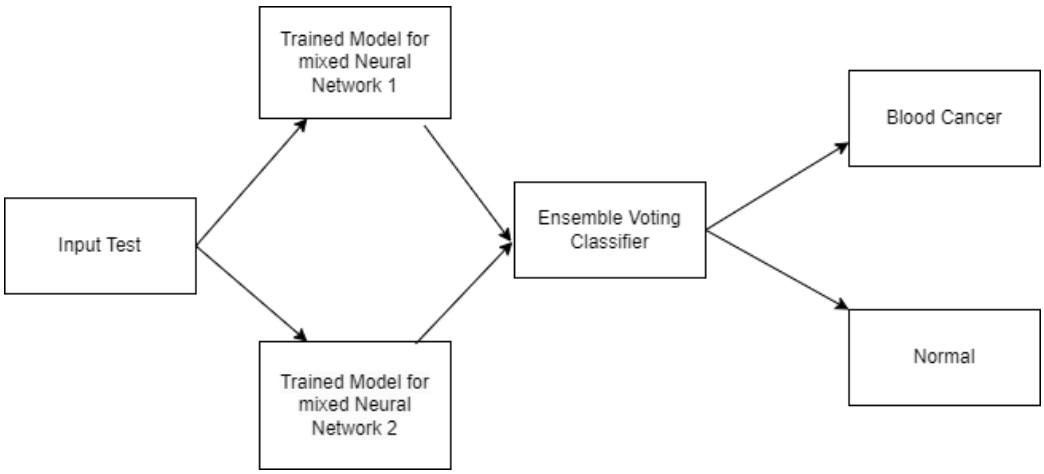

Figure 3. Beginning to Test

the surrounding region's previously existing pixels. There is a chance that if you flip a picture, it will appear mirrored or upside down. For instance, the pixel that was formerly located at (x, y) will now be found at (width - x -1), y) following a horizontal flip.

III Training Process: A method known as rescaling is used to rescale the data in each spectrum dimension. The input is then fed into the Mixed Neural Network Model (MNN). The model's accuracy enhanced by each layer of the MNN, including the convolutional, polling, flattening, and fully connected layers. An ensemble voting classifier was used to provide the final prediction following the application of the two training models to the test data. This followed the use of the training models (see figure 4). Whether or not blood cancer is present in the case is determined by the conclusion.

5 Result

The process of developing a deep learning online application for blood cancer detection and classification includes three stages: web building, model training, and data preparation (see figure 6). Using medical photos, our sophisticated web tool uses deep learning to precisely identify and categories blood cancer. Send in your blood smear slide for a thorough analysis (see figure 7). The approach produces a confidence score, indicates whether

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

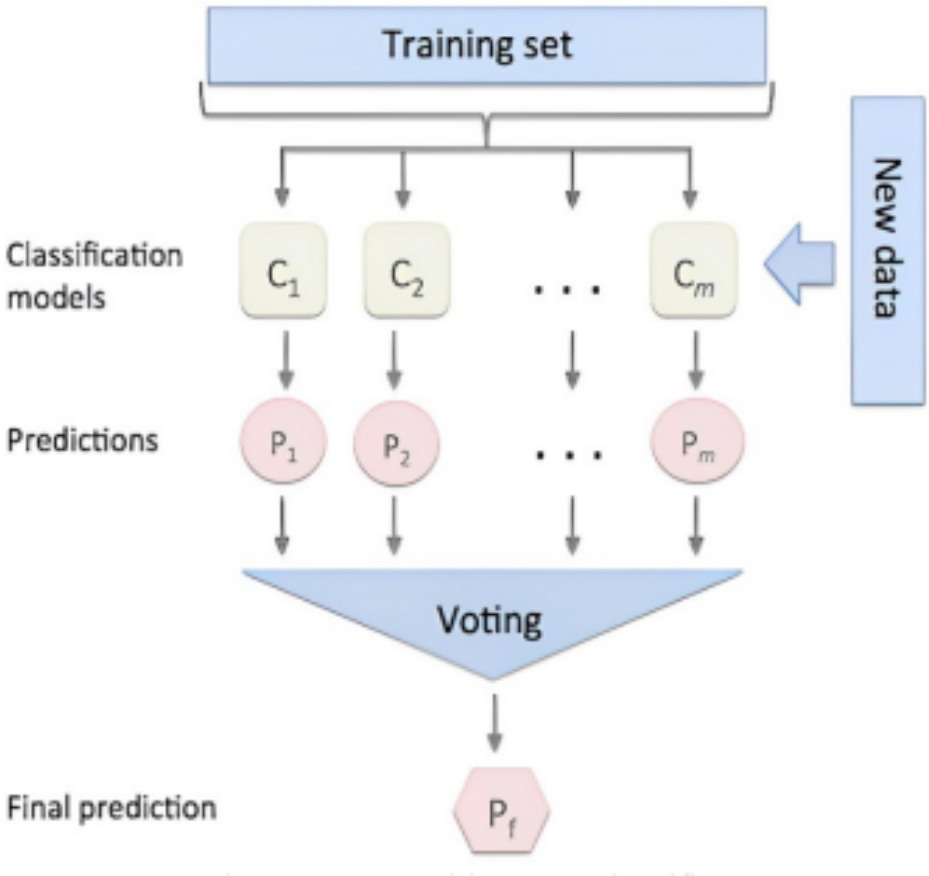

Figure 4. Ensemble Vote Classifier: A majority voting classifier

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

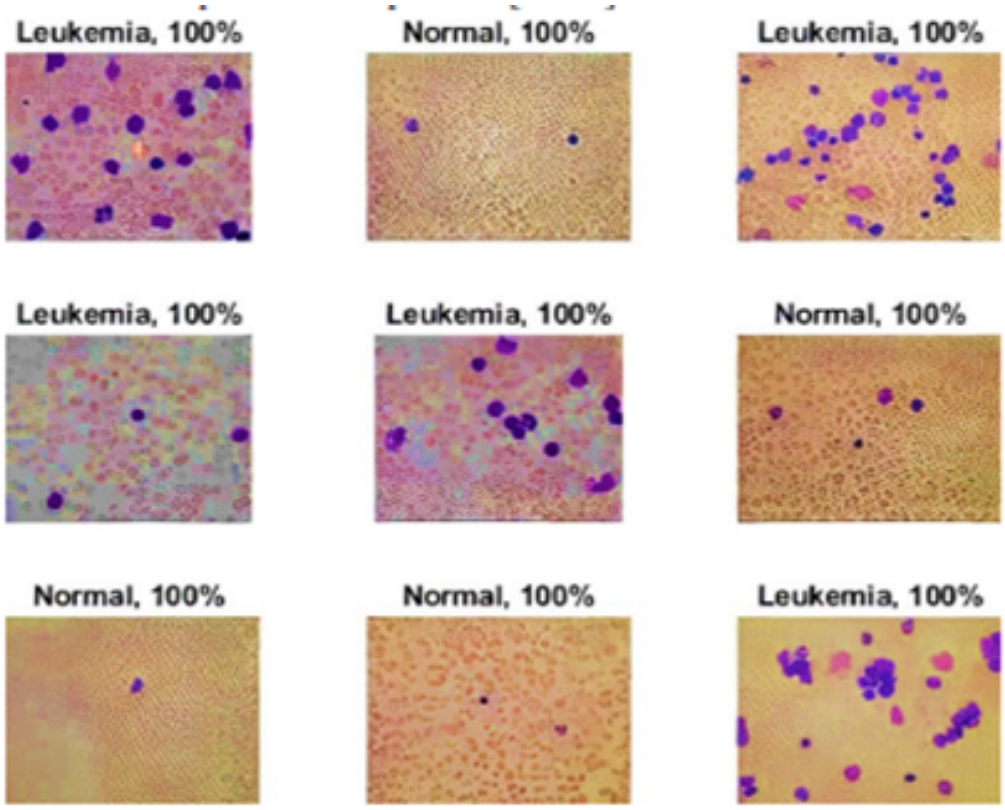

Figure 5. Image displays a grid of microscopic slides comparing normal and leukaemia affected blood cells

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

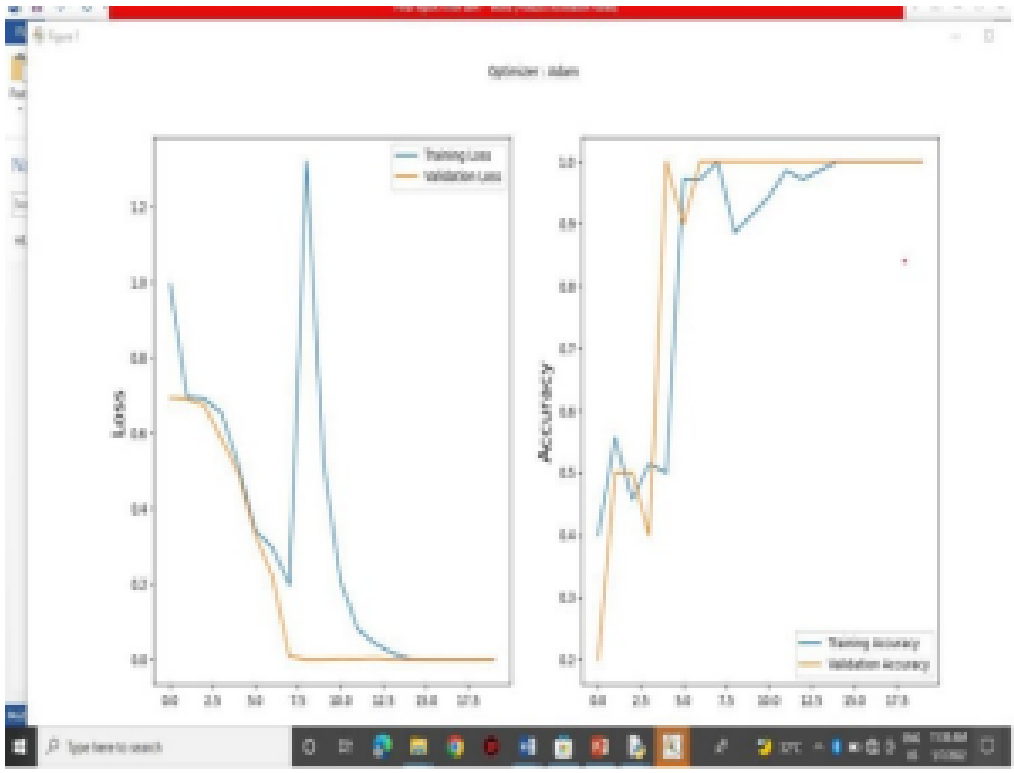

Figure 6. Training Set 1

blood cancer is present, and describes the type (such as leukemia). Tools for visualization draw attention to aberrant areas, which facilitates understanding. With the help of cutting-edge CNN models like ResNet50, our platform guarantees excellent accuracy and dependability (see figure 8). This tool is quick, easy to use, safe, and helps doctors make well-informed diagnostic decisions that improve patient care and results.

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

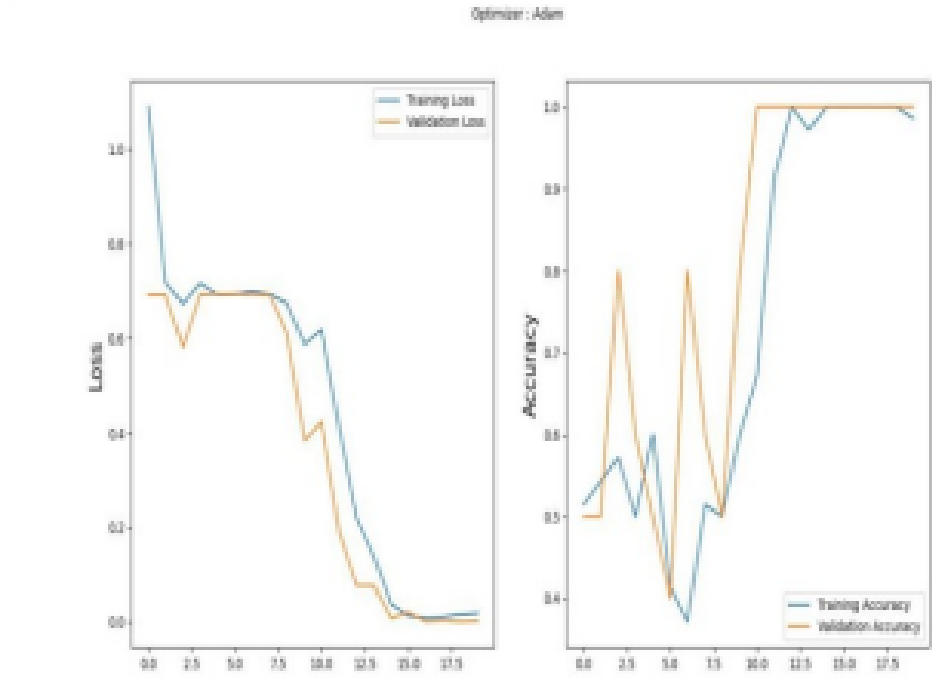

Figure 7. Training Set 2

Editors: S. Pandikumar, Manish Kumar Thakur, Pallavi M O
DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5 |

Figure 8. Result of the application

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

6 Conclusion

The paper developed an advanced diagnostic system using MobileNetV2 architecture and Python. This system surpasses the earlier binary classification model, which only distinguished between "Normal" and "Leukemia" blood cells, by enabling multi-class classification. It now identifies various subtypes of leukemia such as "Benign," "[Malignant] early Pre-B," "[Malignant] Pre-B," and "[Malignant] Pro-B." This improvement provides healthcare professionals with a detailed understanding of the disease, leading to more informed treatment decisions. The system's accuracy enhances blood cancer diagnosis precision, contributing to earlier interventions and better patient outcomes. Additionally, its adaptability ensures it can evolve with advancements in blood cancer research, maintaining its relevance and effectiveness. Overall, the paper significantly advances medical image analysis and has the potential to improve patient lives through earlier and more precise blood cancer diagnoses.

References

Al-Azzawi, H. M. A., Paolini, R., Cirillo, N., O'Reilly, L. A., Mormile, I., Moore, C., Yap, T., & Celentano, A. (2024). Eosinophils in Oral Disease: A Narrative Review. International journal of molecular sciences, 25(8). https://doi.org/10.3390/ijms25084373

Bhaumik, S., Lockett, J., Cuffe, J., & Clifton, V. L. (2023). Glucocorticoids and Their Receptor Isoforms: Roles in Female Reproduction, Pregnancy, and Foetal Development. Biology, 12(8). https://doi.org/10.3390/biology12081104

Choudhry, N., Abawajy, J., Huda, S., & Rao, I. (2023). A Comprehensive Survey of Machine Learning Methods for Surveillance Videos Anomaly Detection. IEEE Access, 11. https://doi.org/10.1109/ACCESS.2023.3321800

Faust, C., Auquier, P., Hamidou, Z., Bertrand, Y., Tabone, M. D., Ansoborlo, S., Baruchel, A., Gandemer, V., Dalle, J. H., Chastagner, P., Kanold, J., Poirée, M., Sirvent, N., Plat, G., Pellier, I., Michel, G., & Berbis, J. (2023). Brothers and sisters of childhood acute leukemia survivors: Their long-term quality of life and its determinants. Cancer Medicine, 12(5), 6200–6212. https://doi.org/10.1002/cam4.5355

Hassan Abbas Gondal, C., Irfan, M., Shafique, S., Salman Bashir, M., Ahmed, M., M.Alshehri, O., H. Almasoudi, H., M. Alqhtani, S., M. Jalal, M., A. Altayar, M., & F. Alsharif, K. (2023). Automated Leukemia Screening and Sub-types Classification Using Deep Learning. Computer Systems Science and Engineering, 46(3), 3541–3558. https://doi.org/10.32604/csse.2023.036476

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5

Koistinen, H., Kovanen, R. M., Hollenberg, M. D., Dufour, A., Radisky, E. S., Stenman, U. H., Batra, J., Clements, J., Hooper, J. D., Diamandis, E., Schilling, O., Rannikko, A., & Mirtti, T. (2023). The roles of proteases in prostate cancer. IUBMB Life, 75(6), 493–513. https://doi.org/10.1002/iub.2700

Liu, S., Lin, Z., Qiao, W., Chen, B., & Shen, J. (2024). Cross-talk between biometal ions and immune cells for bone repair. Engineered Regeneration, 5(3), 375–408. https://doi.org/10.1016/j.engreg.2024.01.003

Murayama, M., Chow, S. K., Lee, M. L., Young, B., Ergul, Y. S., Shinohara, I., Susuki, Y., Toya, M., Gao, Q., & Goodman, S. B. (2024). The interactions of macrophages, lymphocytes, and mesenchymal stem cells during bone regeneration. Bone and Joint Research, 13(9), 462–473. https://doi.org/10.1302/2046-3758.139.BJR-2024-0122.R1

Pirsadeghi, A., Namakkoobi, N., Behzadi, M. S., Pourzinolabedin, H., Askari, F., Shahabinejad, E., Ghorbani, S., Asadi, F., Hosseini-Chegeni, A., Yousefi-Ahmadipour, A., & Kamrani, M. H. (2024). Therapeutic approaches of cell therapy based on stem cells and terminally differentiated cells: Potential and effectiveness. Cells and Development, 177. https://doi.org/10.1016/j.cdev.2024.203904

Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. Computers, 12(5). https://doi.org/10.3390/computers12050091

DOI:10.48001/978-81-966500-2-5-4 | ISBN: 978-81-966500-2-5