



Classification of Acute Lymphoblastic Leukemia based on White Blood Cell Images using InceptionV3 Model

Rizki Firdaus Mulya¹, Ema Utami², Dhani Ariatmanto³

^{1,2,3}Magister of Informatics Engineering, Universitas AMIKOM Yogyakarta, Yogyakarta, Indonesia

¹rizkifirdausmulya@students.amikom.ac.id, ²ema.u@amikom.ac.id, ³dhaniari@amikom.ac.id

Abstract

Acute Lymphoblastic Leukemia (ALL) is the most prevalent form of leukemia that occurs in children. Detection of ALL through white blood cell image analysis can assist in prognosis and appropriate treatment. In this study, the author proposes an approach for classifying ALL based on white blood cell images using a Convolutional Neural Network (CNN) model called InceptionV3. The dataset used in this research consists of white blood cell images collected from patients with ALL and healthy individuals. These images were obtained from The Cancer Imaging Archive (TCIA), which is a service for storing large-scale cancer medical images available to the public. During the evaluation phase, the author used training data evaluation metrics such as accuracy and loss to measure the model's performance. The research results show that the InceptionV3 model is capable of classifying white blood cell images with a high level of accuracy. This model achieves an average ALL recognition accuracy of 0.9896 with a loss of 0.031. The use of CNN models like InceptionV3 in medical image analysis has the potential to enhance the efficiency and accuracy of image-based disease diagnosis.

Keywords: ALL; CNN; deep learning; inceptionV3; leukemia

1. Introduction

Cancer, a complex and devastating disease, continues to pose a significant threat to human health on a global scale. As the leading cause of death worldwide, it manifests through the abnormal and uncontrolled growth of cells, which can infiltrate and harm various tissues within the human body [1]. Among the diverse array of cancers, blood cancer stands out as a particularly distressing condition, mainly when it affects the vulnerable population of children. Shockingly, blood cancer is responsible for claiming the lives of numerous children, ranking it second only to accidents as the leading cause of death among individuals aged 5 to 14 years [2]. When we delve deeper into childhood blood cancer, one specific type looms prominently - acute lymphoblastic leukemia (ALL). This cancer primarily affects children, triggering an uncontrolled growth of immature white blood cells in the bone marrow [3]. The impact of ALL on young lives cannot be understated, as it inflicts immense physical, emotional, and psychological hardships on the affected child and their families. To comprehend the magnitude of blood cancer's prevalence and its toll on society, examining the incidence and mortality rates associated with leukemia is crucial. Based on recent data from 2015 to 2019, the

yearly incidence rate for leukemia is 14.1% per 100,000 individuals, encompassing both men and women. Additionally, 6.0% of every 100,000 individuals succumb to this relentless disease yearly [4]. It is worth noting that these rates have been adjusted for age, taking into account cases reported from 2015 to 2019 and deaths that occurred between 2016 and 2020.

In the field of medicine, the use of images as a diagnostic method has been practiced for a long time, and one of the technological solutions in the medical field is through deep learning [5]. Previous research has utilized machine learning [6]–[8] and deep learning [9]–[11] to detect Leukemia. Artificial Neural Networks (ANN), which serve as the foundation for the subfield of machine learning known as deep learning, have emerged as a transformative force in data analysis and pattern recognition. These intricate computational models are uniquely equipped to tackle diverse tasks, including direct classification of complex data such as images or sound, thereby revolutionizing various domains, including medicine [12]. One of the deep learning techniques that can be used in medical image analysis is Convolutional Neural Network (CNN).

A study conducted [13], a diagnostic aid system was suggested that classified ALL photos and images of

healthy cells using a CNN architecture with an ECA module, an improvement to the VGG16 CNN design. According to their experimental findings, the suggested CNN model extracts features well and achieves an accuracy of 91.1%. In another study [14], ALL images were classified into B-Cell and T-Cell using pre-trained DNN AlexNet and LeukNet models using the American Society of Hematology (ASH) dataset. Their classification accuracy was 94.12% using pre-trained CNN AlexNet, as well as LeukNet.

Inception V3 was utilized in research [15] to extract features, and neural networks were employed to categorize photos into six sports categories based on context and surroundings. The effectiveness of the framework and neural networks has been tested through comparisons with several classifiers, including Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). The findings had an average accuracy rate of 96.64%. Another study [16] developed and proposed a method for real-time detection of Acute Lymphoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML) while maintaining high accuracy. Convolutional Neural Networks (CNN) and an image segmentation technique named AlexNet from MATLAB were employed in the process. The dataset used was C-NMC_Leukemia, which contained approximately 15,000 microscopic images of white blood cells. The proposed approach using CNN (AlexNet) + SVM achieved an accuracy of over 98%. According to research by [17], Acute Lymphocytic Leukemia (ALL) was found by the object detection approach, in which abnormal blood components were recognized by matching them with training data and then using the Faster-RCNN model to predict the risk of cancer cell development. They collected a dataset of 256 primary data from leukemia patients and determined 4 main attributes (Neutrophils, Lymphocytes, Monocytes, and Eosinophils) in determining leukemia patients. The results obtained with average mean precision were 0.10, 0.16, and 0 for each epoch of 40, 60, and 120, respectively.

Based on existing research, there is still no research that classifies ALL using the InceptionV3 architecture for such identification. Instead, the previously described research uses deep learning models with CNN architectures such as VGG16, AlexNet, LeukNet, SVM and Faster-RCNN. In addition, this research tries to develop and see the performance of the Convolutional Neural Network (CNN) on the InceptionV3 model to classify the type of acute lymphoblastic leukemia based on white blood cell images. The researcher will assess the accuracy of leukemia disease recognition using the InceptionV3 architecture for identification purposes.

The content of this paper is organized as follows. The first section discusses the background of leukemia disease with models that can detect ALL using machine

learning models. Section 2 discusses the proposed research method to detect ALL using the InceptionV3 CNN model. Section 3 contains the model training results using scenarios performed at three folds. Finally, Section 4 shows the conclusion of this research.

2. Research Methods

2.1 Research Workflow

The stages of research flow are shown in Figure 1, which are divided into five stages: data collection, pre-processing, CNN model, training, and analysis and decision. In the first stage, the data used is the C-NMC_Leukemia dataset obtained from The Cancer Imaging Archive. The second stage is pre-processing, where the processes involved are data merging, labeling, and augmentation. The third stage is the CNN model, where scenarios are conducted to determine the best accuracy using the CNN algorithm from the InceptionV3 architecture. The fourth stage is training, which aims to train the appropriate algorithm to achieve better accuracy. Finally, the fifth stage is analysis and decision, which involves evaluating the results, conducting analysis, and drawing conclusions based on the experiments conducted.

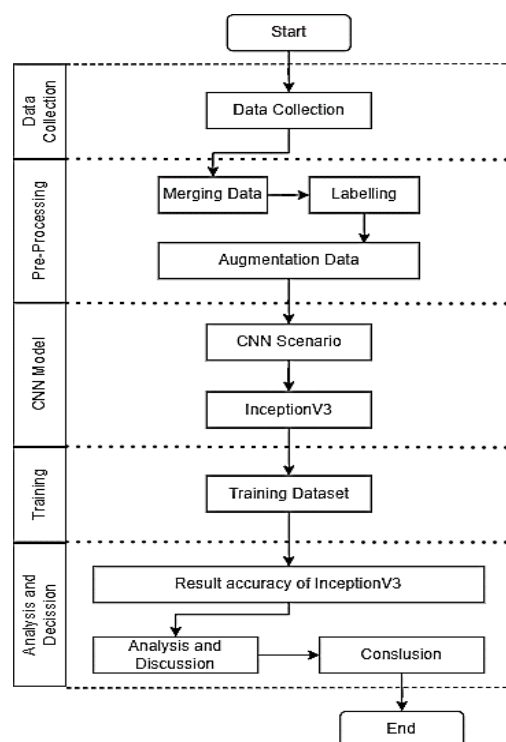


Figure 1. Research Flow Diagram

2.2 Dataset Collection

The Cancer Imaging Archive (TCIA) is a service that stores large-scale cancer medical images available to the public [18]. The dataset obtained by the researcher consists of 15,135 images from 118 patients with ALL, labeled as normal cells and leukemia blasts [19]. The

dataset used is called C-NMC_Leukemia, which consists of microscopic images of white blood cells, specifically lymphoblasts, in bitmap (bmp) format with dimensions of 450 x 450 pixels for each image.

2.3. Data Preprocessing

At this stage, there are several processes carried out, starting from merging data, labeling, and data augmentation so that it can be continued into the next process, namely creating a CNN scenario.

Merging data is done to combine existing data on the C-NMC_Leukemia dataset. The purpose of merging data in this study is that the C-NMC_Leukemia dataset has a dataset that has been divided so that in this process merging is needed to combine these datasets.

Labeling is done on the white blood cell image data that has been collected, then labeled according to the image information. The label performed in this study is a label for white blood cell images based on patients labeled normal cell and blast leukemia.

Data augmentation is done by modifying the original image by changing its shape or position. The purpose of data augmentation is to improve the performance of the model because the machine will be able to recognize more objects of various shapes and patterns. The data augmented in this research will make changes to the brightness of the image such as saturation, contrast, brightness, and gamma in the image.

2.4 CNN Model

Convolutional Neural Networks (CNN), a remarkable innovation inspired by the intricate workings of the

human brain, have revolutionized the field of deep learning and emerged as a potent tool in various image processing and computer vision applications [20]. The versatility and adaptability of CNN models have allowed them to be extensively employed in numerous tasks, ranging from image segmentation and feature extraction to text-to-image synthesis and face recognition, among countless others [21].

A typical CNN model comprises four essential elements, each contributing to its formidable computational prowess. The first element is the convolutional layer, which performs localized feature extraction by convolving input data with a set of learnable filters, enabling the network to discern relevant patterns and structures within images. Following the convolutional layers, the pooling layers help downsample the spatial dimensions of the feature maps, reducing computational complexity and facilitating translation invariance. Activation functions, such as the popular Rectified Linear Unit (ReLU), introduce non-linearity into the network, enabling the modeling of complex relationships between features and improving the network's expressive capabilities. Finally, the fully connected layers integrate the learned features and generate the final predictions based on the extracted information [22]. A distinctive characteristic of CNNs is their ability to organize and shape neurons into three dimensions, forming a multi-layered architecture capable of capturing intricate details and spatial relationships within images [23]. The architecture of the CNN model is depicted in Figure 2.

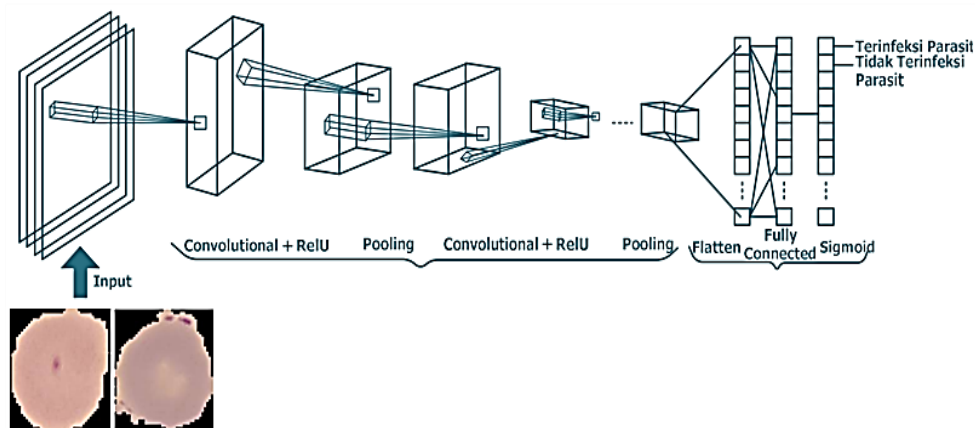


Figure 2. CNN Architecture Model

The proposed pre-trained Inception-V3 model, as described by [24]. One of the top hardware experts in the industry created this model, which has over 20 million parameters. Each model's symmetric and asymmetric building block comprises different convolutional, average, max pooling, concatenated,

dropout, and fully connected layers. The input activation layers of this model also undergo batch normalization, which is standard practice. Using Softmax for classification is used. The diagram of the Inception-V3 CNN model is shown in Figure 3.

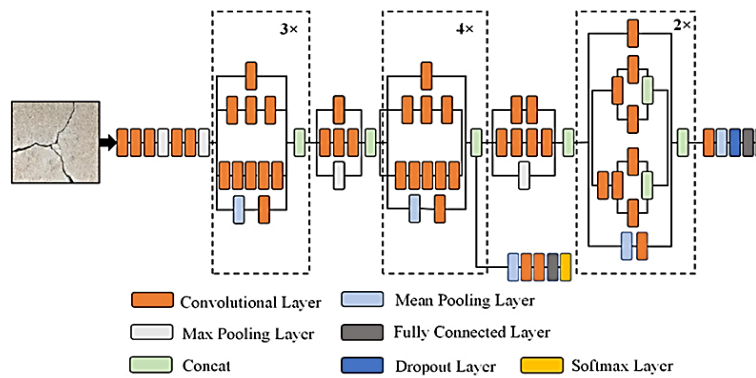


Figure 3. InceptionV3 CNN Architecture Model

3. Results and Discussions

This section is the result of the training data stage which aims to determine the performance of the InceptionV3 architecture by looking at the accuracy of the architecture. This is followed by analyzing and giving conclusions on the experiments that have been carried out.

In order to achieve the highest level of accuracy attainable, the research procedure will be meticulously conducted, incorporating a carefully designed experimental framework that integrates a well-defined scenario and the powerful CNN algorithm of the InceptionV3 architecture. This comprehensive approach aims to harness the model's full potential and ensure a robust evaluation of its performance across multiple folds. To facilitate a thorough assessment of the model's capabilities, the scenario will involve a particular configuration of the experimental parameters, including the number of epochs utilized for each fold. In this study, an empirically determined choice has been made to utilize 50 epochs for fold 1, fold 2, and fold 3. Implementing different folds can enhance the performance of the model, as shown in Table 1.

Table 1. Comparison of Accuracy for Each Fold

Fold	Accuracy	Loss
Fold 1	0.978	0.061
Fold 2	0.996	0.0137
Fold 3	0.995	0.019

By implementing multiple folds within the experimental framework, the researcher aims to capitalize on the benefits of cross-validation, which has been widely recognized as a crucial technique for assessing the robustness and generalizability of machine learning models. The inclusion of different folds not only helps mitigate the potential bias and variability inherent in a single fold but also allows for a more comprehensive analysis of the model's performance across diverse subsets of the dataset.

In the comprehensive series of tests conducted as part of this research study, the dataset employed for experimentation and analysis was thoughtfully partitioned into two subsets. Namely, a substantial 70% of the data training, which equated to approximately 10.500 datasets, was exclusively reserved for the critical task of training the machine learning model, while the remaining 30% of the data testing, comprising an impressive count of roughly 4.500 datasets, was thoughtfully earmarked for the pivotal role of serving as the test data during the rigorous evaluation process.

In the first scenario, the researcher reached an early stopping at epoch 24 of the 50 epochs set at the beginning, with the highest accuracy obtained being 0.978 and a corresponding loss of 0.061. The graphical representation of the training accuracy and loss values can be seen in Figure 4.

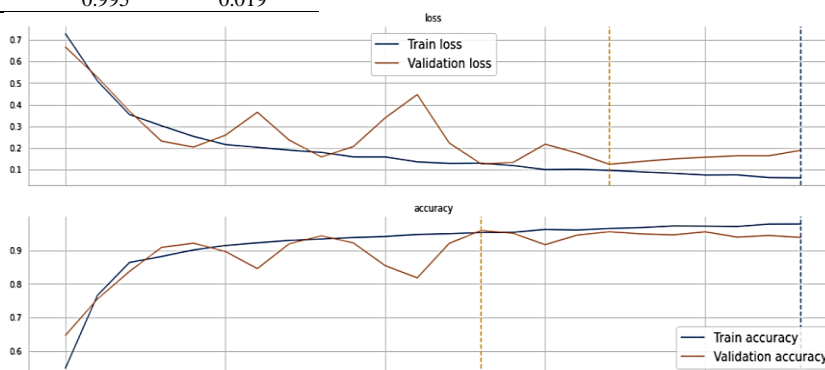


Figure 4. Accuracy and loss results for fold 1

In the second scenario, the researcher obtained an accuracy result at epoch 50, with a value of 0.996 and a corresponding loss of 0.0137. In this scenario, higher accuracy results are obtained with lower loss compared

to the first scenario. The graphical representation of the training accuracy and loss values can be seen in Figure 5.

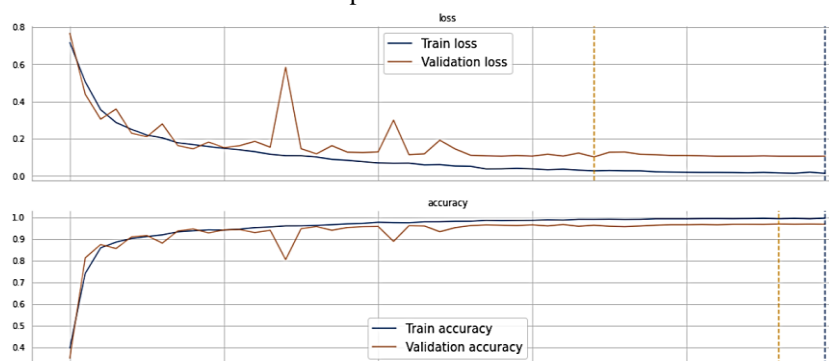


Figure 5. Accuracy and loss results for fold 2

In the third scenario, the researcher obtained an accuracy result at epoch 50, with a value of 0.995 and a corresponding loss of 0.019. In the second scenario, higher accuracy results were obtained with lower loss

compared to the first scenario. However, the accuracy and loss results are almost comparable to the second scenario. The graphical representation of the training accuracy and loss values can be seen in Figure 6.



Figure 6. Accuracy and loss results for fold 3

From the three testing scenarios conducted, the researcher obtained the highest accuracy result of 0.996 and a loss value of 0.0137 in the second fold. The average results from the three folds yielded an accuracy of 0.9896 and an average loss of 0.031. This indicates that the testing conducted using the InceptionV3 architecture is effective in achieving high accuracy in image classification [15]. This is due to the InceptionV3 architecture's ability to better recognize patterns and objects at different scales. Additionally, the transfer learning approach used to train the InceptionV3 model has shown superior performance compared to models such as LSTM [25].

From the results and analysis of the research mentioned, the researcher contributes to the development of a more effective and non-invasive diagnostic solution for the recognition and classification of acute lymphoblastic leukemia through white blood cell image analysis. It also encourages the use of image processing and deep learning technologies in the field of medicine and

hematology to improve the testing and diagnosis of acute lymphoblastic leukemia.

4. Conclusion

The utilization of CNN models such as InceptionV3, in medical image analysis represents a transformative leap forward in enhancing the efficiency and accuracy of image-based disease diagnosis. By leveraging this advanced technology, the automatic detection and classification of acute lymphoblastic leukemia (ALL) based on white blood cell images can be seamlessly accomplished, equipping medical experts with invaluable insights for more precise prognoses and tailored treatment strategies. The research findings from this study unequivocally demonstrate the remarkable effectiveness of the CNN InceptionV3 model in the context of ALL classification. The results revealed an impressive average accuracy of 0.9896 on the training data, complemented by a remarkably low loss value of 0.031. These metrics are compelling evidence of the model's inherent capabilities and ability to detect and

classify ALL with remarkable accuracy efficiently.

However, this study exclusively focuses on the classification of ALL and does not encompass the careful consideration of other factors that may influence prognosis or treatment outcomes. While the results obtained from this research provide significant insights into the potential of the CNN InceptionV3 model in the context of ALL diagnoses, it is imperative to conduct further investigations and validation studies on a larger and more diverse population. Such endeavors should consider additional prognostic factors, including patient demographics, genetic markers, and treatment response, to comprehensively understand the model's utility and applicability in real-world clinical settings.

References

- [1] K. D. Nugroho, "Studi Fenomenologi : Pengabaian Gejala Awal Kanker mengakibatkan Respon Negatif bagi Klien dan Keluarga," *Jurnal Keperawatan Malang*, vol. 5, no. 1, pp. 46–54, Jun. 2020, doi: 10.36916/jkm.v5i1.105.
- [2] E. Ward, C. Desantis, ; Anthony Robbins, B. Kohler, and A. Jemal, "Childhood and adolescent cancer statistics, 2014," *CA Cancer J Clin*, vol. 64, no. 2, pp. 83–103, Mar. 2014, doi: 10.3322/CAAC.21219.
- [3] V. S. Pojoh, M. F. J. Mantik, and J. I. Ch. Manoppo, "Hubungan Indeks Massa Tubuh dan Tercapainya Remisi pada Anak Penderita Leukemia Limfoblastik Akut," *e-Clinic*, vol. 8, no. 1, Dec. 2019, doi: 10.35790/ec1.8.1.2020.27186.
- [4] Surveillance Epidemiology and End Result, "Leukemia — Cancer Stat Facts," 2022. <https://seer.cancer.gov/statfacts/html/leuks.html> (accessed Oct. 08, 2022).
- [5] D. Gunawan and H. Setiawan, "Convolutional Neural Network dalam Analisis Citra Medis," 2022.
- [6] A. Sarkar, T. Das, G. Das, and Z. Ghosh, "MicroRNA mediated gene regulatory circuits leads to machine learning based preliminary detection of acute myeloid leukemia," *Comput Biol Chem*, vol. 104, p. 107859, Jun. 2023, doi: 10.1016/j.compbiolchem.2023.107859.
- [7] E. M. Nagiub Abdelsalam, K. F. Hussain, N. M. Omar, and Q. T. Ali, "Computer Aided Leukemia Detection using Microscopic Blood Image Based Machine Learning 'Convolutional Neural Network,'" *Clin Lymphoma Myeloma Leuk*, vol. 18, p. S297, Sep. 2018, doi: 10.1016/j.clml.2018.07.246.
- [8] E. Nagiub, K. F. Hussain, N. Omar, and Q. Taher, "ALL-018 Application of Machine Learning in the Diagnosis of Acute Lymphoblastic Leukemia," *Clin Lymphoma Myeloma Leuk*, vol. 22, p. S190, Oct. 2022, doi: 10.1016/S2152-2650(22)01174-0.
- [9] A. Abhishek, R. K. Jha, R. Sinha, and K. Jha, "Automated detection and classification of leukemia on a subject-independent test dataset using deep transfer learning supported by Grad-CAM visualization," *Biomed Signal Process Control*, vol. 83, p. 104722, May 2023, doi: 10.1016/j.bspc.2023.104722.
- [10] K. K. Anilkumar, V. J. Manoj, and T. M. Sagi, "Automated detection of leukemia by pretrained deep neural networks and transfer learning: A comparison," *Med Eng Phys*, vol. 98, pp. 8–19, Dec. 2021, doi: 10.1016/j.medengphy.2021.10.006.
- [11] H. Xu, J. Jia, H.-H. Jeong, and Z. Zhao, "Deep learning for detecting and elucidating human T-cell leukemia virus type 1 integration in the human genome," *Patterns*, vol. 4, no. 2, p. 100674, Feb. 2023, doi: 10.1016/j.patter.2022.100674.
- [12] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network," *JUSTINDO (Jurnal Sistem & Teknologi Informasi Indonesia)*, 2018.
- [13] M. Zakir Ullah *et al.*, "An Attention-Based Convolutional Neural Network for Acute Lymphoblastic Leukemia Classification," *Applied Sciences*, vol. 11, no. 22, p. 10662, Nov. 2021, doi: 10.3390/app112210662.
- [14] K. K. Anilkumar, V. J. Manoj, and T. M. Sagi, "Automated Detection of B Cell and T Cell Acute Lymphoblastic Leukaemia Using Deep Learning," *IRBM*, vol. 43, no. 5, pp. 405–413, Oct. 2022, doi: 10.1016/J.IRBM.2021.05.005.
- [15] K. Joshi, V. Tripathi, C. Bose, and C. Bhardwaj, "Robust Sports Image Classification Using InceptionV3 and Neural Networks," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 2374–2381. doi: 10.1016/j.procs.2020.03.290.
- [16] T. Shawly and A. A. Alsheikhy, "Biomedical Diagnosis of Leukemia Using a Deep Learner Classifier," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/1568375.
- [17] M. Akter Hossain, M. Islam Sabik, I. Muntasir, A. K. M. Muzahidul Islam, S. Islam, and A. Ahmed, "Leukemia Detection Mechanism through Microscopic Image and ML Techniques," in *2020 IEEE REGION 10 CONFERENCE (TENCON)*, IEEE, Nov. 2020, pp. 61–66. doi: 10.1109/TENCON50793.2020.9293925.
- [18] K. Clark *et al.*, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *J Digit Imaging*, vol. 26, no. 6, pp. 1045–1057, Dec. 2013, doi: 10.1007/s10278-013-9622-7.
- [19] A. Gupta and R. Gupta, "ALL Challenge dataset of ISBI 2019 [Data set]," *The Cancer Imaging Archive*, 2019, doi: <https://doi.org/10.7937/tcia.2019.dc64i46r>.
- [20] P. Kim, *MATLAB Deep Learning*. Berkeley, CA: Apress, 2017. doi: 10.1007/978-1-4842-2845-6.
- [21] N. R. Fauziyya, "Metoda Convolutional Neural Network (CNN) untuk Pendeteksi Tangga pada Alat Pemandu Arah bagi Penyandang Tunanetra," *Telekontran: Jurnal Ilmiah Telekomunikasi, Kendali dan Elektronika Terapan*, vol. 8, no. 2, pp. 145–153, Apr. 2021, doi: 10.34010/telekontran.v8i2.4709.
- [22] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," *Procedia Comput Sci*, vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.
- [23] N. K. C. PRATIWI, N. IBRAHIM, Y. N. FU'ADAH, and S. RIZAL, "Deteksi Parasit Plasmodium pada Citra Mikroskopis Hapusan Darah dengan Metode Deep Learning," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 9, no. 2, p. 306, Apr. 2021, doi: 10.26760/elkomika.v9i2.306.
- [24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.
- [25] M. Ahmed, N. Afreen, M. Ahmed, M. Sameer, and J. Ahamed, "An inception V3 approach for malware classification using machine learning and transfer learning," *International Journal of Intelligent Networks*, vol. 4, pp. 11–18, 2023, doi: 10.1016/j.ijin.2022.11.005.