2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) | 978-1-6654-9299-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/UEMCON54665.2022.9965679

A Multistage Transfer Learning Approach for Acute Lymphoblastic Leukemia Classification

Renato R. Maaliw III
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
rmaaliw@slsu.edu.ph

Alvin S. Alon
Digital Transformation Center
Batangas State University
Batangas City, Philippines
alvin.alon@g.batstate-u.edu.ph

Ace C. Lagman
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
aclagman@feutech.edu.ph

Manuel B. Garcia
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
mbgarcia@feutech.edu.ph

Julie Ann B. Susa
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
jsusa@slsu.edu.ph

Ryan C. Reyes
Electrical Engineering Dept.
Technological University of the Philippines
Manila, Philippines
ryan_reyes@tup.edu.ph

Ma. Corazon Fernando - Raguro
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
mgfernando@feutech.edu.ph

Alexander A. Hernandez
Graduate Program & External Studies
Technological University of the Philippines
Manila, Philippines
alexander_hernandez@tup.edu.ph

***Abstract*—Automated medical image analysis driven by artificial intelligence can revolutionize modern healthcare in producing swift and precise diagnostics. Due to doctors' varying breadths of training and expertise, traditional leukemia screening methods frequently involve considerable subjectivity. Using a 3-stage transfer learning approach and stacks of convolutional neural networks, we constructed an efficient pathway for automatic leukemia identification and classification through various phases. Experimental findings disclosed that our pipeline powered by InceptionResNetV2 architecture decisively affects the accuracy with 99.60% (normal vs. leukemia) and 94.67% (normal to L3). Moreover, it reduces error rates by 1.65% and 6.05%, respectively. A consistent result via the T-test confirms our proposed framework robustness with a significant positive difference of 4.71% over the standard transfer learning mechanism (p-value = 0.0001 & t = 0.85310). This research could aid and support oncologists in early yet reliable prognoses of acute lymphoblastic leukemia types.**

Keywords—blood cancer, convolutional neural networks, deep learning, image processing, machine learning, oncology

I. Introduction

Leukemia is a life-threatening form of cancer distinguished by the growth of abnormal blood cells causing the immune system's degradation. According to World Health Organization (WHO) 2020 statistics, it has an estimated 2.6% (474,519 – ranked 13th) cases out of the 18.1 million total cancer instances worldwide [1]. Their classification is either acute (rapidly progressing) or chronic (gradually advancing) and by the type of cell in which it manifests: myeloid or lymphoid. Understanding the exact type of its nature at an early stage enables physicians to accurately anticipate a patient's prognosis and identify the most effective treatment for increasing a patient's survival rate significantly.

Acute lymphoblastic leukemia (ALL) is a fatal condition accounting for 25% of pediatric cancers. It is brought by bone marrow's immature lymphocytes where leukemic cells spread swiftly into various organs (spleen, liver, kidney & lymph nodes) and the central nervous system (brain & spinal cord) of the body [2]. Fig. 1 displays ALL's variants according to the French-American-British (FAB) system [3]. The L1 is characterized by its uniform nuclear structure with homogenous chromatin, sparse basophilic cytoplasm, and modest sizes. On the other hand, L2 exhibits irregular nuclear shape and evident splits in large blasts, which is intensely basophilic polymorphic chromatin. Severe stage L3 ranges from medium to large sizes, containing noticeable cytoplasmic vacuoles of round to oval shape, comprising two or three nucleoli. For its diagnosis, hematologists performed blood smears or bone marrow microscopic examinations. This process is manual, complex, costly, and time-consuming for doctors to differentiate the morphological characteristics between normal versus immature leukemic cells due to other surrounding blood components such as erythrocytes (red blood cells or RBC) and platelets. Unfortunately, errors in medical image interpretation are not negligible as it heavily relies on the pathologist's knowledge, training, experience, and expertise [4]. In order to circumvent these constraints, experts recommend using an accurate yet economical intelligent automated system against labor-intensive manual diagnostics for categorizing malignant lymphocytes (white blood cells or WBC) [5].

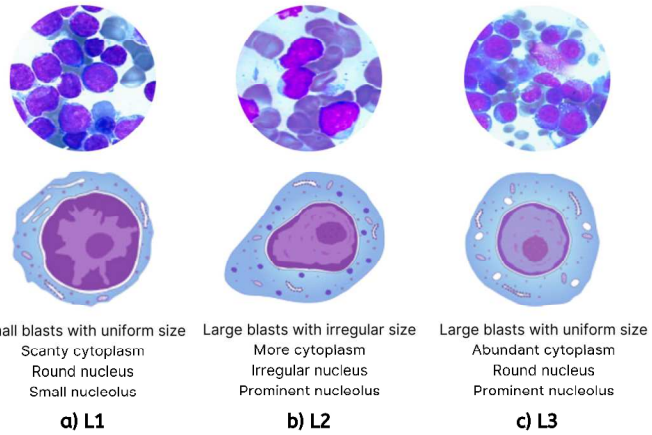

Fig. 1. FAB classification system for ALL (a-c) [6][7].

According to the literature, various studies ventured into leukemia screening and grading. Previous efforts

Commission on Higher Education (CHED) – Institutional Development and Innovation Grants (IDIG)

concentrated on traditional image processing and machine learning techniques based on feature extractions, segmentations, and classifications. The authors [8 – 10] offered different classifiers such as support vector machines (SVM) and Naïve Bayes in conjunctions with Gray-level-run-length matrix (GLRLM), based on textures, color, and shape attributes. Proponents [11] employed watershed transformation and the Gaussian mixture model (GMM) to recognize the condition in blood micrographs. WBC nucleus and cytoplasm fractionization proposed by [12] utilized GVF snake to identify the illness. Conversely, paper [13] conducted a morphological analysis to identify leukemic cells via orthogonal S-transform feature extraction and linear discriminant analysis (LDA). Researchers [14] presented the localization of lymphoblast cells using color models accompanied by pathological operations for the partitions. Scientists [15] [16] experimented with different decision tree algorithms (random forest, gradient boosting & CART) with satisfactory accuracy. Cell fragmentization techniques introduced by [17] leverage contrast enhancements, histogram equalization, Otsu threshold, and K-nearest neighbor, claiming accuracy of 93%. All of the aforementioned strategies involve feature extractions and segmentations that significantly influence the overall performance. However, data scientists regard handcrafted features as a drawback of conventional machine learning because of adaptability and scalability concerns [18].

In the context of image recognition, deep learning's (DL) mechanism delegate attribute design to the underlying network, making previous solutions obsolete with the emergence of transfer learning (TL). TL is a constantly evolving and remarkably effective tool for medical image processing powered by convolutional neural networks (CNN). The work of [19] utilized modified pre-trained CNNs of *ResNet-101*, *AlexNet* [20], and *GoogleNet* for WBC counting and classifications. Several studies deployed *VGGNet* [21] [22], *InceptionV3* [23], *DenseNet* [24] [25], and *MobileNet* [22] for a four-level discrimination of ALL with average accuracies of 95% to 97%. Another study attained 98% high precision ratings using *ResNetV2* by extracting topological features through augmented data preprocessing and skipped connections to mitigate the vanishing gradient problem [26]. Article [27] established a practical yet lightweight architecture involving depth-wise separable convolutions under a *MobileNetV2* that capitalized on width (layer reduction) and resolution multiplier (computational cost optimization). Diverse scholars suggested ensemble learning frameworks to enhance performance further and avoid overfitting in exchange for execution times. Integration of *VGG19* and *NASNetLarge* achieves an accuracy of 96.58% exceeding individual models [28].

We observed that the prior system relied solely on manual, standard ML, and DL modalities. Furthermore, existing implementations of ALL's characterization suffer due to the small dataset, compromising generalization capabilities. Leukemia prognosis is a sensitive matter that can lead to life or death situations needing an almost perfect evaluation. Our contribution aims to advance medical image analysis by developing comprehensive procedures for extracting structural information by considering the long-distance dependencies of hematological microscopic images. Improving the identification and four-level classification centered on distinct processes and multistage deep learning networks can benefit medical practitioners in the expeditious and reliable diagnosis of ALL with little physical interventions.

II. Methodology

The subsequent section comprehensively explains ALL's automatic detection and classification. Key portions consist of data acquisition, preprocessing, augmentation, image segmentation, convolutional neural network architectures implementation, hyperparameter tuning, and evaluation metrics.

A. Data Collection

We collected 100 (ALL-IDB1 – for detection & classification) and 240 (ALL-IDB2 – for segmentation) 24-bit colored JPEG images from a private repository [29], where the first set contains 39,000 labeled lymphocytes by highly trained oncologists with 2592 x 1944 resolutions. The second batch of data consists of clipped regions of interest from ALL-IDB1 representing normal and leukemic cells with dimensions of 257 x 257 pixels, where 50% denote lymphoblasts. Fig. 2 and Fig. 3 show excerpts of the images.

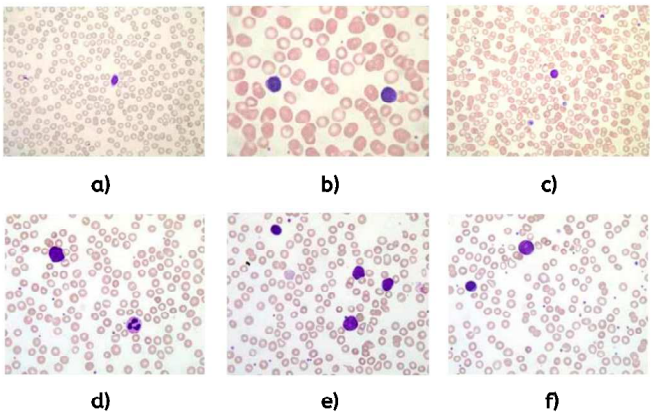

Fig. 2. Non-ALL (a-c) and ALL (d-f) miscroscopy of patients.

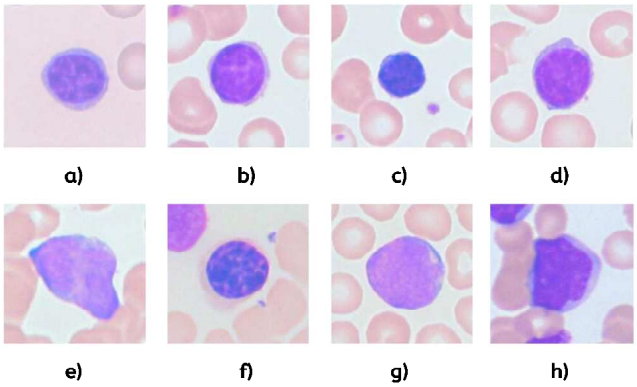

Fig. 3. Healthy (a-d) and lymphoblastic (e-h) cells.

B. Data Preprocessing

Most microscopic images show all the blood components (RBC & platelets). However, we are only explicitly concerned with immature lymphocytes for ALL's type. We experimented with k-means clustering, watershed, and morphological operations to define the region of interest (ROI), but the RGB (red-green-blue) channels in monochrome filter proved simple yet efficient. To precisely reflect the blasts, we applied specific values of red = +88, green = +41, and blue = +20. The refinement of the ROI is handled by threshold (t) operation, with t = 98, accomplished to discard even more unnecessary information to discriminate lymphoblast. This initial stage was essential for segmentation purposes on the next phase. Fig. 4 depicts the process.

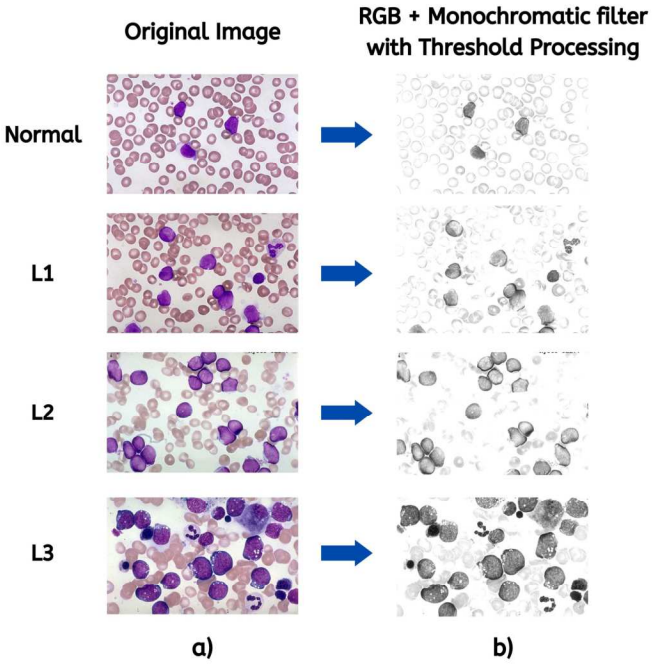

Fig. 4. Original images (a) are preprocessed using RGB channels with monochrome filter (b) to extract ROI and discard other information.

C. Image Segmentation

The volume and localization of WBC in blood slides are helpful for the diagnosis of all forms of leukemia. Segmentation offers a practical basis for the advancement of image processing techniques. Nonetheless, evaluating microscopic hematological photographs remains extremely difficult owing to intricacy issues of backdrop due to the overlapping blood-related structures. Moreover, histological staining conditions add to the variances [30]. To solve this dilemma, we designed a modified Residual U-Net with a 3 x 3 stride to buffer its default CNN's convolution layers, a rectified learning unit (ReLU) activation function, and incorporated standardization to encoder and decoder sections of the network (Fig. 5). Using a 2 x 2 max-pooling, we then reduce the size of the feature map to make learned features robust and diminish the effects of noise (Fig. 6(a)). The final step is to transform the partitioned WBC via RGB's blue-channel (Fig. 6(b)) to isolate the lymphoblast and eliminate other components with a black backdrop (Fig. 6(c)).

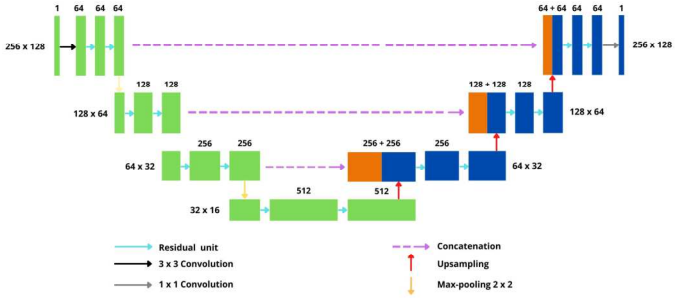

Fig. 5. Modified Residual U-Net for WBC segmentation with encoder (left), decoder (right) sections, and concatenation mechanism.

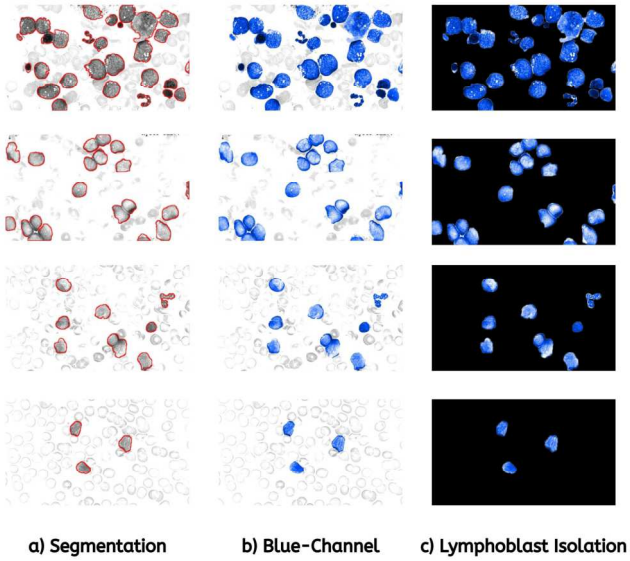

Fig. 6. Segmentation (a), blue-channel transformation (b), and lymphoblast isolation using black backdrop to eliminate other blood components (c).

D. Data Augmentation

The inadequacies of medical image training impede pattern recognition betterment, as it directly affects the performance of DL. As a solution, we performed dataset augmentation by producing additional unique data synthetically. Image transformations such as flipping (vertical & horizontal), rotating (random angle of 0 to 180 degrees), shifting, and scaling produced an additional 1,360 (ALL-IDB1 = 400 & ALL-IDB2 = 960) images for a total of 1,700 well-balanced data based on ALL's categorization. By leveraging this procedure, data scientists can avoid the painstaking and costly data collection in preparation for building robust prediction models impervious to overfitting. Lastly, we divided the dataset into 80/20 splits for training and testing, using 10-fold cross-validation.

E. Multistage Transfer Learning

TL enables deep learning algorithms to construct an effective model by transferring information from one operation to another despite lacking sufficient training data. While medical photos are not included in pre-trained

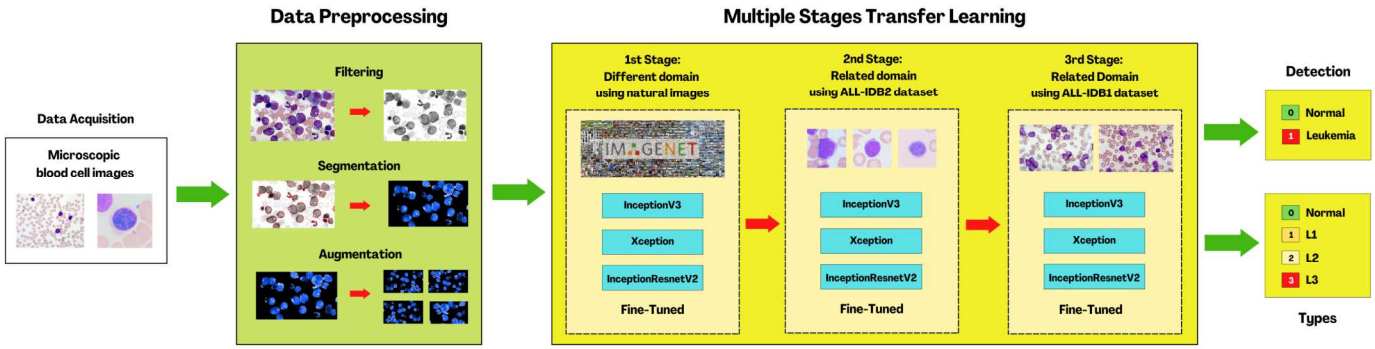

Fig. 7. Proposed pipeline for multistage transfer learning (MTL) approach for acute lymphoblastic leukemia (ALL) detection and classification.

frameworks such as *ImageNet's* natural images [31], they are nevertheless widely used in different fields. Using microscopic images that can be collected and augmented in large quantities, we firmly believe that the most effective way to improve computer vision analysis is to learn both from natural and domain-specific datasets through the multistage transfer learning (MTL) approach. In detail, a single-step TL works through a given source domain (D_s), target domain (D_t), and a learning problem (L_p). It increases the knowledge acquisition of a target function (f_t) using D_s and D_t. We strengthen the pipeline by performing a three-stage TL. Fig. 7 shows the proposed architecture consisting of several phases. In the first phase, a preliminary TL from *ImageNet* converts the natural (raw) images' similar features to those found in blood smear scans. This reduces the cross-entropy function using the initial weights (W_0) from the pre-trained model to yield (W_1) described in Equation 1 [32]:

$$J(\langle W_1, b|W_0\rangle) = \frac{-1}{mn}\sum_{i=1}\sum_{j=1} y^{ij}\, log\big(P\langle y^{ij}|x^{ij}, W_0, W_1, b\rangle\big) \quad (1)$$

where x^i is the iterative input, y^i is the concurrent label, $\langle y^{ij}|x^{ij}, W_0, W_1, b\rangle$ is the *softmax's* probability, and b is the bias. Next, we leverage the previous stage's TL as a baseline (W_1) and provide weights (W_2) for classifying individual lymphoblast (Fig. 3 (a-h)) with m training samples depicted in Equation 2 [32]:

$$J(\langle W_2, b|W_1\rangle) = \frac{-1}{m}\sum_{i=1} y^i\, log\big(P\langle y^i|x^i, W_1, W_2, b\rangle\big) + (1-y^i)\log(1-P\langle y^i|x^i, W_1, b\rangle) \quad (2)$$

where $P\langle y^i|x^i, W_1, W_2, b\rangle$ as the *sigmoid* unit's output probability. The final phase is to incorporate the learnt information from single lymphoblast (W_2) to multiple instances of the cells (W_3) in order to efficiently generalize the categorization (Fig. 2 (a-f)).

F. Deep Convolutional Neural Network Models

We implemented the top three DL architectures based on preliminary assessments such as *InceptionV3* [33], *Xception* [34], and *InceptionResNetV2* [35]. In our experiment, we utilized a global average pooling, a single dense layer with *softmax*, and weight fine-tuning except for the last layer. Optimization of the learning rate starts with a 0.001 exponential decay. Most models showed exceptional performance compared with cases that included both regularization and dropout. Still, no issues point to overfitting due to the substantially augmented dataset. After the transition from single to multiple cells in the different stages of TL, network layers were eliminated and replaced by multiple dense layers, with dropouts, and finally an activation function (sigmoid) to finalize the CNN's design. Fig. 8 shows the neural networks implementation for each phase.

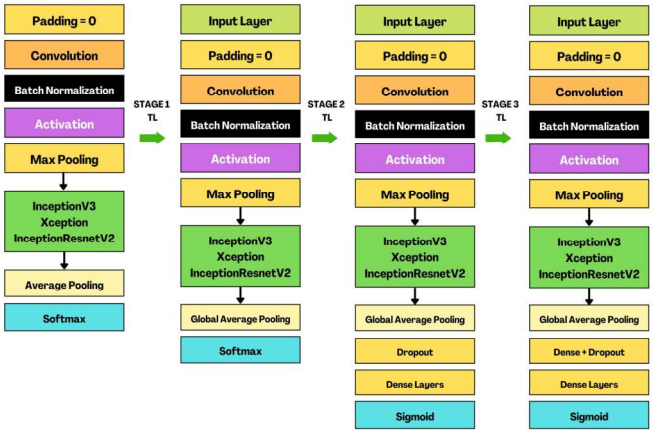

Fig. 8. CNN's architectural configurations for each stages of transfer learning based on pre-trained natural images. Stage 1 (ImageNet) followed by domain-specific ALL datasets (Stage 2 & Stage 3).

G. Hyperparameter Fine-Tuning

Optimizing hyperparameters are essential for machine learning to operate efficiently. Unlike model parameters, these settings are established before training. It is one of the most challenging and ignored aspects of creating DL networks. Due to manual configurations' intricacy and time-consuming expense, we utilized a sequential-based optimization approach with the average adjusted values for CNN models presented in Table I.

TABLE I. NEURAL NETWORKS' FINE-TUNED CONFIGURATIONS

Architecture	Configuration	Value
InceptionV3	Epoch	160
	Environment	GPU
	Optimizer	ADAM
Xception	Learning rate	0.001
	Batch size	48
InceptionResnetV2	Loss	Multiclass cross-entropy
	Shuffling	Per epoch

H. Performance Evaluation Metrics

Accuracy alone is entirely inadequate for assessing the cogency of our MTL model. For a full-blown evaluation, we quantitatively experimented with different deep learning architectures in terms of F1-Score, precision, recall, confusion matrices, receiver operating characteristics (ROC), and area under the curve (AUC) scores. Furthermore, we also conducted a T-test [36] to quantify the substantial advancement of our methods in classifying ALL by comparing two independent groups (classical versus multistage TL). The following equations provide an overview of the metrics based on true positives (TP), true negatives (TN), false positives (FP), false negatives (FN), and significance tests (*confidence* = 95%, *p-value* at 0.05). A result below the predetermined p-value implies a measurable difference, whereas a value above indicates otherwise.

$$Accuracy\ (AC) = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision\ (PR) = \frac{TP}{TP + FP} \tag{4}$$

$$Recall\ (RE) = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - Score\ (F1) = 2\,x\frac{PR\ x\ RE}{PR + RE} \tag{6}$$

III. Results

For this experiment, we executed the algorithms through a high-capacity machine powered by a Core-i9 processor with 5GHz & 32MB L1 and L2 cache, a 64GB RAM with an ASUSRTX3070 DDR6 graphics card. Data preprocessing, augmentation, and DL modeling were implemented via Python, Keras, and Tensorflow libraries. The subsequent sections detail the outcomes.

A. Learning Stabililization Convergence Plots

Fig. 9 showcases the convergence loss plots of distinct neural network variants concerning the number of epochs. Due to the implementation of additional TL stages, understandably, all evaluated networks took longer to achieve equilibrium than the single and double-stage approaches. Although it entails extra processing overheads, the MTL method shows significant gains in error loss, especially with fine-tuned hyperparameters.

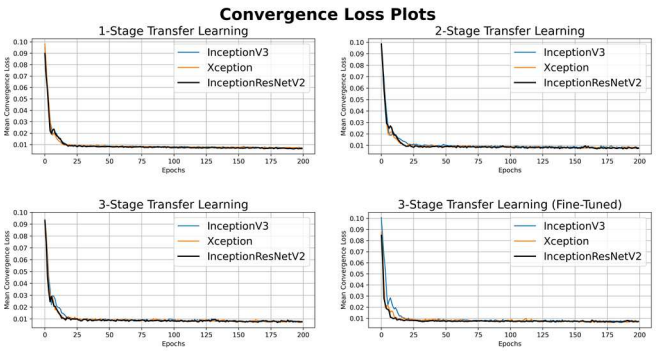

Fig. 9. Mean convergence loss plots based on different TL configurations.

B. Leukemia Detection Performance

The results shown in Table II suggest that the *InceptionResnetV2* is preferable to other neural network's architectural design in binary classification, with an overall accuracy of 99.60%. Moreover, it had a predictive forecasting power higher than other models with 99.10% (F1-Score), 98.80% (recall), and 99.60% (precision). Although it came in last place, the *InceptionV3* had a respectable 97.80% accuracy.

TABLE II. 10-fold Cross-validated Detection Performance for All (normal vs. leukemia)

Model	Evaluation Metrics			
	Accuracy	**Precision**	**Recall**	**F1-Score**
InceptionV3	0.978	0.985	0.963	0.973
Xception	0.981	0.986	0.967	0.976
InceptionResnetV2	**0.996**	**0.996**	**0.988**	**0.991**

C. Leukemia Classification Performance

With an accuracy of 94.67%, the numbers presented in Table III proved that the *InceptionResNetV2* was better. It outperformed the generalization ability of other networks by a margin of 94.68% (F1-Score), 94.66% (recall), and 94.67% (precision). The *InceptionV3* provides the lowest performance with an accuracy of 87.33% comparable to the identification results of leukemia on Table II.

TABLE III. 10-fold Cross-validated Classification Performance For All (normal to L3)

Model	Evaluation Metrics			
	Accuracy	**Precision**	**Recall**	**F1-Score**
InceptionV3	0.8733	0.8766	0.8736	0.8741
Xception	0.9000	0.9024	0.9000	0.9005
InceptionResnetV2	**0.9467**	**0.9467**	**0.9466**	**0.9468**

Furthermore, we also constructed ROC and AUC graphs for the best-performing architecture to evaluate its prediction quality, including its generalization ability. Numbers imply that the *InceptionResNetV2* can determine the severity of leukemia as illustrated in Fig. 10, with scores of 0.98 (normal), 0.97 (L1), 0.92 (L2), and 0.97 (L3). These values indicate significant ratings that go above randomness (red diagonal dashed line). In related measurements, micro and macro averages of 0.97 and 0.96 reinforced its robustness.

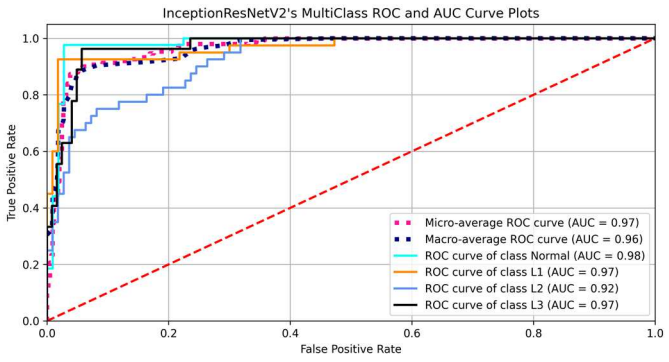

Fig. 10. Prediction quality of *InceptionResNetV2* on leukemia categorization.

We then compared the confusion matrices in Table IV to deliver in-depth perspicuity into each network's classification capabilities. It is clear from the results that the majority of errors (vice-versa) originated from combinations of Normal-L1 (6.67%), L2-L3 (6.18%), and L1-L2 (5.77%). There are no instances of misdiagnosis for Normal-L2, Normal-L3, and L1-L3. These marks unequivocally demonstrate the generalization strength of each model, particularly the *InceptionResNetV2* in distinguishing between leukemia stages.

TABLE IV. CONFUSION MATRICES FOR ALL'S CLASSIFICATION (TEST DATA)

InceptionResnetV2				
	Normal	**L1**	**L2**	**L3**
Normal	36 (97.29%)	1 (2.70%)	0	0
L1	1 (2.63%)	35 (92.10%)	2 (5.26%)	0
L2	0	2 (5.40%)	34 (91.89%)	1 (2.70%)
L3	0	0	1 (2.63%)	37 (97.36%)
Xception				
	Normal	**L1**	**L2**	**L3**
Normal	34 (91.89%)	3 (8.10%)	0	0
L1	3 (7.89%)	33 (86.84%)	2 (5.26%)	0
L2	0	2 (5.40%)	34 (91.89%)	1 (2.70%)
L3	0	0	4 (10.52%)	34 (89.47%)
InceptionV3				
	Normal	**L1**	**L2**	**L3**
Normal	33 (89.18%)	4 (10.81%)	0	0
L1	3 (7.89%)	32 (84.21%)	3 (7.89%)	0
L2	0	2 (5.40%)	33 (89.18%)	2 (5.40%)
L3	0	0	5 (13.15%)	33 (86.84%)

D. *InceptionResNetV2's Training/Validation Loss and Accuracy*

Fig. 11(a) highlights logarithmic values of 0.58 and 0.68 as initial training and validation losses. Based on the plots, a point of growing convergence was observed from epochs 65 to 108 until it reached a point of stabilization at the 160th. On the other side, Fig. 11(b) illustrates a progressive convergence at epochs 70 to 120, reaching the most outstanding training (97.78%) and validation (97.89%) accuracies at the 154th epoch. Both figures proved that the best model did not diverge (overfit or underfit) in classifying leukemia.

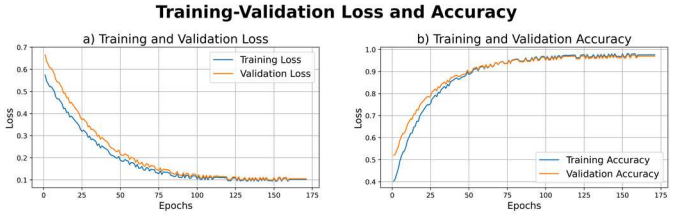

Fig. 11. The InceptionResNetV2 model showed no signs of divergence based from the training-validation loss and accuracy plots.

E. *Comparative Benchmark Performance between Classical versus MultiStage Transfer Learning*

To completely appreciate the improvement of our recommended strategy, we conducted ten repeated random test for each trials to evaluate the conventional and the multistage transfer learning methods mean accuracies. The expectations of the two schemes are compared in Table V. The T-test indicates that there is substantial difference between the two sets with p-value = 0.0001, t = 0.85310, and degree of freedom of 18. With a 4.71% average increase in accuracy, this reflects that our method is significantly superior to the standard transfer learning approach.

TABLE V. ACCURACY COMPARISON BETWEEN DIFFERENT TRANSFER LEARNING APPROACHES

Trials	**Classical Transfer Learning**	**Multistage Transfer Learning**
	Accuracy (%)	**Accuracy (%)**
1	89.17	93.45
2	90.28	95.23
3	87.31	92.82
4	90.12	94.73
5	89.35	92.38
6	88.72	93.85
7	87.39	95.03
8	92.43	93.73
9	89.32	94.39
10	88.85	94.52

IV. DISCUSSIONS

Our experimentation verified that multiple stages of transfer learning using stacks of pre-trained CNNs elevated the performance of leukemia identification and classification through various procedures. Moreover, the progression of training using natural images with domain-specific medical datasets benefited the knowledge abstraction of lymphoblast characteristics. The empirical results indicated that the *InceptionResnetV2* surpasses both the *Xception* and *InceptionV3* models in recognizing the internal and external cellular blood structures with accuracies of 99.60% (detection) and 94.67% (categorization). It also downsizes error rates by a mean of 1.65% and 6.05% - a noteworthy gain in image processing. We also documented that exhaustive procedures such as channel filtering, segmentation, augmentation, color transformations, MTL, convergence graphs, and training-validation tests stood as a preamble to our pipeline's marked precision. Despite the time-intensive and computational nature of each stages optimization, the benefits exceeded the downsides. Confirmatory T-test calculations revealed a consistent improvement of almost 5% significant positive difference for MTL over CTL. This work remarked that complex neural network architectures did not necessarily usher in high metric scores as each has distinct advantages purely dependent on the targeted medical image domains. The research published by [5] [8] [19] [32] [37 – 46], and [21 – 23] in machine learning-based medical image processing is on equal ground with our obtained findings. Lastly, the hindrances of noisy blood scans leading to low picture quality caused prediction discrepancies. This paper did not consider any solutions for mitigating the effects of the aforementioned problems, such as graphics enhancement and reconstructions.

V. Conclusions and Future Work

ALL is a blood cancer that advances quickly in the absence of interventions. If not analyzed correctly in the earlier stages, it can result in multiple organ failures and death. Prognoses that are dependable, reliable, and timely are crucial for oncologists when formulating treatment plans for their patients. The existing systems for its evaluation are rigorous, laborious, and primarily prone to subjective disparities due to personal knowledge and experiences of doctors. In addition, most medical experts agreed and attested to the difficulties of using only the naked eye for microscopic determination. We created an end-to-end pipeline using artificial intelligence to detect and classify a fatal disease using a multistage transfer learning approach composed of different convolutional neural networks. Our trials showed exceptional accuracy improvements against the classical transfer learning model. As an essential contribution to the advancement of science, specifically in medical image analysis, we formulated a streamlined system for automated recognition of ALL. There is optimism that this research will lead to a better understanding of all forms of cancers and their corresponding medications. In future works, the authors intend to ramp up the efficiency by including an array of image enhancing techniques and benchmarking other CNNs.

Acknowledgement

The corresponding author would like to express his gratitude to Prof. Fabio Scotti (Università Degli Studi di Milano) for the acute lymphoblastic leukemia datasets, the Commission on Higher Education, and the Southern Luzon State University for supporting this research.

References

[1] H. Inaba and C. G. Mullighan, "Pediatric acute lymphoblastic leukemia," Haematologica, 105(11), 2020.

[2] C. H. Pui, "Precision medicine in acute lymphoblastic leukemia," Frontiers of Medicine, 14(6), pp. 689-700, 2020.

[3] K. K. Jha, P. Das and H. S. Dutta, "FAB classification based leukemia identification and prediction using machine learning," IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1-6, 2020.

[4] C. M. Albert, J. L. Davis, N. Federman, M. Casanova and T. W. Laetsch, "TRK fusion cancers in children: A clinical review and recommendations for screening," Journal of Clinical Oncology, 37(6), pp. 513-524, 2019.

[5] A. T. Sahlol, P. Kollmannsberger and A. A. Ewees, "Efficient classification of white blood cell leukemia with improved swarm optimization of deep features," Scientific Reports, 10(1), pp. 1-11, 2020.

[6] Nagoya University School of Medicine – Department of Medicine and Nagasaki University School of Medicine – Department of Hematology, Available at: http://www3.med.unipmn.it/did/will/atlashem/node54.htm.

[7] Y. Cui, M. Zhou, P. Zou, X. Liao and J. Xiao, "Mature B cell acute lymphoblastic leukaemia with KMT2A-MLLT3 transcripts in children: three case reports and literature reviews," Orphanet Journal of Rare Diseases, 16(1), pp. 1-10, 2021.

[8] B. V. S. Krishna, J. J. Godwin, S. T. Shree, B. Sreenidhi and T. Abinaya, "Detection of leukemia and its types using combination of support vector machine and K-nearest neighbors algorithm," Next Generation of Internet of Things, pp. 435-444, 2021.

[9] B. K. Das and H. S. Dutta, "GFNB: Gini index–based fuzzy naive bayes and blast cell segmentation for leukemia detection using multi-cell blood smear images," Medical, Biological Engineering & Computing, 58(11), pp. 2789-2803, 2020.

[10] P. K. Das, P. Jadoun and S. Meher, "Detection and classification of acute lymphocytic leukemia," IEEE-HYDCON, pp. 1-5, 2020.

[11] P. Mirmohammadi, M. Ameri and A. Shalbaf, "Recognition of acute lymphoblastic leukemia and lymphocytes cell subtypes in microscopic images using random forest classifier," Physical and Engineering Sciences in Medicine, 44(2), pp. 433-441, 2021.

[12] K. G. Dhal, J. Gálvez, S. Ray, A. Das and S. Das, "Acute lymphoblastic leukemia image segmentation driven by stochastic fractal search," Multimedia Tools and Applications, 79(17), pp. 12227-12255, 2020.

[13] P. K. Das, V. A. Diya, S. Meher, R. Panda and A. Abraham, "A systematic review on recent advancements in deep and machine learning based detection and classification of acute lymphoblastic leukemia," IEEE Access, pp. 81741-81763, 2022.

[14] R. Sigit, M. M. Bachtiar and M. I. Fikri, "Identification of leukemia diseases based on microscopic human blood cells using image processing," IEEE International Conference on Applied Engineering (ICAE), pp. 1-5, 2018.

[15] S. Dasariraju, M. Huo and S. McCalla, "Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm," Bioengineering, 7(4), 2020.

[16] M. A. Deif, R. E. Hammam and A. Solyman, "Gradient boosting machine based on PSO for prediction of leukemia after a breast cancer diagnosis," International Journal on Advanced Science Engineering Information Technology, 11(2), pp. 508-515, 2021.

[17] N. M. Deshpande, S. Gite, B. Pradhan, K. Kotecha and A. Alamri, "Improved Otsu and Kapur approach for white blood cells segmentation based on LebTLBO optimization for the detection of leukemia," Math Biosci Eng, 19(2), pp. 1970-2001, 2022.

[18] T. Pansombut, S. Wikaisuksakul, K. Khongkraphan and A. Phon-On, "Convolutional neural networks for recognition of lymphoblast cell images," Computational Intelligence and Neuroscience, 2019.

[19] M. Macawile, V. Quiñones, A. Ballado, J. Cruz and M. Caya, "White blood cell classification and counting using convolutional neural network," 3rd International Conference on Control and Robotics Engineering (ICCRE), pp.259-263, 2018.

[20] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks," Technology in Cancer Research & Treatment, 17(1), 2018.

[21] M. O. Aftab, M. J. Awan, S. Khalid, R. Javed and H. Shabir, "Executing spark BigDL for leukemia detection from microscopic images using transfer learning," IEEE 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), pp. 216-220, 2021.

[22] F. S. K. Sakthiraj, "Autonomous leukemia detection scheme based on hybrid convolutional neural network model using learning algorithm," Wireless Personal Communications, pp. 1-16, 2021.

[23] S. Ramaneswaran, K. Srinivasan, P. M. Vincent and C. Y. Chang, "Hybrid inception v3 XGBoost model for acute lymphoblastic leukemia classification," Computational and Mathematical Methods in Medicine, 2021.

[24] B. Arivuselvam and S. Sudha, "Leukemia classification using the deep learning method of CNN," Journal of X-Ray Science and Technology, 30(3), pp. 567-585, 2022.

[25] K. K. Anilkumar, V. J. Manoj and T. M. Sagi, "Automated detection of leukemia by pretrained deep neural networks and transfer learning: A comparison," Medical Engineering & Physics, 98(1), pp. 8-19, 2021.

[26] X. Xie, Y. Li, M. Zhang, Y. Wu and L. Shen, "Multi-streams and multi-features for cell classification," ISBI C-NMC Challenge: Classification in Cancer Cell Imaging, pp. 95-102, 2019.

[27] J. Amin, M. Sharif, M. A. Anjum, M. Yasmin, K. I. Khattak, S. Kadry, and S. Seo, "An integrated design based on dual thresholding and features optimization for white blood cells detection," IEEE Access, 9(1), pp. 151421-151433, 2021.

[28] P. H. Kasani, S. W. Park and J. W. Jang, “An aggregated-based deep learning method for leukemic B-lymphoblast classification,” Diagnostics, 10(12), 1064, 2020.

[29] F. Scotti, “Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images,” IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA), pp. 96-101, 2005.

[30] Y. Rivenson, H. Wang, Z. Wei, K. de Haan, Y. Zhang, et al., “Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning,” Nature Biomedical Engineering, 3(6), pp. 466-477, 2019.

[31] M. A. Morid, A. Borjali and G. Del Fiol, “A scoping review of transfer learning research on medical image analysis using ImageNet,” Computers in Biology and Medicine, 128, 104115, 2021.

[32] G. Ayana, J. Park, J. Jeong and S. Choe, “A novel multistage transfer learning for ultrasound breast cancer image classification,” Diagnostics, 12(135), pp. 1-14, 2022.

[33] N. Dong, L. Zhao, C. H. Wu and J. F. Chang, “Inception v3 based cervical cell classification combined with artificially extracted features,” Applied Soft Computing, 93, 106311, 2020.

[34] H. Benbrahim and A. Behloul, “Fine-tuned Xception for image classification on tiny Imagenet,” IEEE International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP), pp. 1-4, 2021.

[35] S. Li, W. Yang, A. Zhang, H. Liu, J. Huang, C. Li and J. Hu, “A novel method of bearing fault diagnosis in time-frequency graphs using InceptionResnet and deformable convolution networks,” IEEE Access, 8(1), pp. 92743-92753, 2020.

[36] B. Gerald, “A brief review of independent, dependent and one sample t-test,” International Journal of Applied Mathematics and Theoretical Physics, 4(2), pp. 50-54, 2018.

[37] S. Rezayi, N. Mohammadzadeh, H. Bouraghi, S. Saeedi and A. Mohammadpour, “Timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods,” Computational Intelligence and Neuroscience, 2021.

[38] S. Chand and V. P. Vishwakarma, “A novel deep learning framework (DLF) for classification of acute lymphoblastic leukemia,” Multimedia Tools and Applications, p. 1-20, 2022.

[39] G. V. S. Sashank, C. Jain and N. Venkateswaran, “Detection of acute lymphoblastic leukemia by utilizing deep learning methods,” Machine Vision and Augmented Intelligence—Theory and Applications, pp. 453-467, 2021.

[40] Z. Jiang, Z. Dong, L. Wang and W. Jiang, “Method for diagnosis of acute lymphoblastic leukemia based on ViT-CNN ensemble model,” Computational Intelligence and Neuroscience, 2021.

[41] P. K. Das and S. Meher, “Transfer learning-based automatic detection of acute lymphocytic leukemia,” IEEE National Conference on Communications (NCC), pp. 1-6, 2021.

[42] A. Genovese, M. S. Hosseini, V. Piuri, K. N. Plataniotis and F. Scotti, “Histopathological transfer learning for acute lymphoblastic leukemia detection,” IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), pp. 1-6, 2021.

[43] R. R. Maaliw, J. A. B. Susa, A. S. Alon, A. C. Lagman, S. C. Ambat, M. B. Garcia, K. C. Piad and M. C. Fernando-Raguro, “A deep learning approach for automatic scoliosis cobb angle identification,” IEEE World Artificial Intelligence and Internet of Things Congress (AIIoT), pp. 111-117, 2022.

[44] R. R. Maaliw, A. S. Alon, A. C. Lagman, M. B. Garcia, M. V. Abante, R. C. Belleza, J. B. Tan and R. A. Maaño, “Cataract detection and grading using ensemble neural networks and transfer learning,” IEEE International Conference and Workshop on Computing and Communication (IEMCON), pp. 74-81, 2022.

[45] R. Khandekar, P. Shastry, S. Jaishankar, O. Faust and N. Sampathila, “Automated blast cell detection for acute lymphoblastic leukemia diagnosis,” Biomedical Signal Processing and Control, 68, 102690, 2021.

[46] R. R. Maaliw, M. A. Ballera, Z. P. Mabunga, A. T. Mahusay, D. A. Dejelo and M. P. Seño, “An ensemble machine learning approach for time series forecasting of COVID-19 cases”, IEEE 12th Annual Information Techology, Electronics and Mobile Communication Conference (IEMCON), pp. 633-640, 2021.