ADROID Conference Series: Engineering and Technology
Volume 2, Issue 3, pp.40-47, 2026

Lymphoblastic leukaemia (blood cancer) detection: AI-Assisted Diagnosis Through Blood Sample Images

M. A. Kumar, Kanika, Udaya Sri, Bhuvana, Dimpy Garg

Apex Institute of Technology, Chandigarh University, Mohali, India

manchu.e18260@cumail.in, 22bai71383@cuchd.in, 22bai71389@cuchd.in, 22bai71358@cuchd.in
22bai71408@cuchd.in

Abstract. Leukaemia is a serious disease that affects the blood. It happens when bad white blood cells, called blast cells, grow out of control in the blood and bone marrow. Finding these cells early and accurately is extremely important for figuring out what is wrong, how sick someone is, and which treatment will work best. Doctors usually look at blood slides under a microscope to find these cells. This method works. It takes a long time and the doctors have to be very careful. Sometimes doctors can make mistakes because they get tired or disagree with what they see. This can be a problem, especially when they have to review many blood slides or lack all the resources they need. Leukaemia diagnosis is a process, and leukaemia cells can be hard to find. That is why leukaemia diagnosis needs to be accurate and fast. Leukaemia treatment needs to be planned, and leukaemia patients need the best care possible. Leukaemia is a disease, and leukaemia patients deserve the best chance to get better. In this paper, we describe an AI-powered automatic detection system of leukaemia blast cells using deep Convolutional Neural Networks (CNNs) that can be used by haematologists in the clinical diagnostic process. The system uses a VGG16-based CNN trained via transfer learning on the C-NMC: B-lineage acute lymphoblastic leukaemia dataset, which consists of more than 10,000 blood cell images annotated by experts across various leukocyte types. Several image preprocessing and data augmentation techniques, as well as a method to address class imbalance, were applied to improve the system's robustness and generalisation. Experimental results show that the proposed system performs on par with human experts, achieving greater than 95% sensitivity for blast cells and an ROC-AUC higher than 0.98. Grad-CAM explainable AI visualisations are included to make these diagnostic inferences more transparent and to foster clinical acceptance, illustrating the cellular areas that contribute most to the prediction. Overall, the framework demonstrates the ability to substantially improve diagnostic consistency and throughput while enabling fast, accurate, and understandable leukaemia screening.

I. INTRODUCTION

Leukemia represents a global public health problem, with a vast number of new diagnoses recorded every year in countries around the world. Acute leukaemias (more narrowly acute myeloid leukaemias - AML) are a critical type of hematological malignancies characterized by the rapid, uncontrolled proliferation of very immature WBC, called blast cells. The speed of acute nature of leukemia is responsible for quick detection of blast cells in peripheral blood to rapidly and timely diagnose and treat leukemia.

Classical leukemic diagnostics are based on the morphological criteria of blast cells which are visually determined under microscopic examination of stained slides of peripheral blood smear. Experts in haematology analyse blast morphology based on factors such as the nucleus-to-cytoplasm ratio, chromatin texture, the presence and size of nucleoli, and cytoplasmic granularity. While this is the current clinical gold standard there are number of weaknesses which are implicit to these techniques. Major challenges in traditional leukemia diagnosis include:

- Very high time consumption in large-scale laboratories, which often results in diagnostic delays.
- Observer variability leading to subjective and inconsistent diagnostic outcomes.
- Errors which get induced due to fatigue during prolonged microscopic evaluation of extensive sample volumes.
- Limited availability of experienced hematologists, especially in developing and resource-constrained regions.

In this project, the pre-trained Convolutional Neural Network VGG16 has been used as a deep feature extractor on 224x224 pixel microscopic images to look at fundamental edges and leukemia patterns. Removing the classification layers results in a 1-dimensional array of 25,088 features, enabling its use with transfer learning to diagnose small medical datasets, enabling accurate diagnoses by classical classifiers, such as Support Vector Machines and Random Forests, and enabling visualization via Grad-CAM.

II. RELATED WORK

The detection of leukemia from peripheral blood smear images using artificial intelligence has recently become a field of active research that attracted more and more interests. Initial studies in this area used hand-crafted feature extraction techniques and conventional machine learning classifiers (e.g., SVM, k-NN, Random Forests). Hand-crafted features mostly comprise of morphological features, texture features and color features, such as leukocyte shape, nucleus texture, cytoplasmic contents. Even with positive observations, the manual approaches are susceptible to variations in illumination and staining procedure and are affected by cell morphology. Thus, they are moderately robust and generalizable for clinical purposes.

The field of hematological image analysis was radically changed by the introduction of deep learning. Deep learning algorithms, based on Convolutional Neural Network (CNNs), made feature engineering, traditionally a manual and time-consuming process, redundant, allowing the models to learn optimal discriminating features from the blood smear images themselves. In the case of classification, Matek et al. has shown that it is possible to reach a human level of accuracy, when classifying leukocytes and discriminating leukemic blasts from normal cells, using deep CNN architectures (specifically models based on ResNet). Their research achieved ROC-AUC values over 0.99 when identifying leukemic blast cells using large annotated datasets.

Researchers have subsequently used transfer learning approaches involving pre-trained architectures VGG16, ResNet, DenseNet and Xception to overcome this issue of limited annotated medical datasets [5]. These methods improved classification accuracy and speed of convergence, utilizing features learned on massive natural images database. Even with positive observations, the manual approaches are susceptible to variations in illumination and staining procedure and are affected by cell morphology. Thus, they are moderately robust and generalizable for clinical purposes [1].

To address the ”black box” problem and promote clinician trust, explainable AI methods such as Grad-CAM have been developed. [13]. Grad-CAM visualisations produce visually appealing heat-maps that identify clinically interesting regions in a cell image and helps a clinician in confirming if the model choices are in line with the accepted morphological guidelines used in the practice of hematology. Notwithstanding these advances many existing systems are lacking end-to-end clinical usability and fail to properly account for issues such as data imbalance, generalization across institutions and interpretability.

Motivated by these drawbacks, the present work proposed a system that leverages strong deep learning abilities, interpretability, and deployability, resulting in a clinically significant and scalable system for automated leukemia blast cell detection.

III. IMPLEMENTATION

A. CNN Logic Network Layers (VGG16 Backbone)

- **Convolutional Layers (Conv2D):** The network uses 3 x 3 mathematical filters (kernels) and slides them over the image and performs a dot product at each location. The first layers can detect the low-level spatial feature of the crisp cellular membrane edge and the high-level semantic feature, such as irregular, patchy chromatin texture present in leukemic blast cells can be found in the deeper layers as the edges of those features can be used in subsequent layers.
- **Max Pooling Layers (2×2):** Following convolution layers, pooling layers reduce the spatial dimensions by keeping the highest activation within a 2x2 section. This makes the network position invariant; whether the cancer cell is directly in the center, to the side, or rotated, the AI should be able to detect it.
- **Flattening Layer:** The final 3D tensor from the last convolutional block is ”unrolled” into a 1D vector. This transforms the visual data into an array of precisely 25,088 numerical features.

B. The Hybrid Machine Learning Classifiers

- **Support Vector Machine (SVM):** SVM receives the 25,088 dimensions data and draws in the hyper dimensional space, and wants to find a perfect hyperplane (a mathematical boundary) separating the ”Normal” cells and ”Leukemic” cells. Because of the highly non-linear image data, we chose the RBF (Radial Basis Function) Kernel to project data into another higher dimension space to make the data linearly separable.
- **Random Forest** It works as a bagging of decision trees. It randomly draws many hundreds of independent decision trees, each trained on randomly chosen subset of the 25,088 attributes, and makes final classification based on majority voting. It performs extremely well with small datasets where the risk of overfitting is very high.

C. Training Details: Epochs vs. Hybrid Optimization

- **Standard CNN Training (Backpropagation):** In most cases, the AI will guess the diagnosis, compute how wrong it was (Loss), and then propagate that error backwards through the network through a number of iterations or Epochs (usually 50 or 100 epochs).
- **Our Hybrid Approach (Transfer Learning):** In our code, the VGG16 is frozen. It contains pre-trained weight from millions of images (ImageNet). Thus, there is only one forward pass for the CNN to get the features. The number of epoch for the CNN is zero.
- **Classifier Fitting:** The actual ”learning” does occur within the SVM. There are no epochs; the SVM utilizes convex mathematical optimization (e.g. Sequential Minimal Optimization) to locate the maximum margin hyperplane in a single, unbroken calculation. Consequently, this hybrid training phase runs much faster and overfits much less than a training session from random initialization on a deep network.

TABLE I
DATABASE SCHEMA OVERVIEW

Table	Columns	Description
images	image id, file path, label	Blood smear image metadata
features	image id, feature vector	Extracted deep features
predictions	model name, class, confidence	Model inference outputs
audit logs	timestamp, action, model used	Experiment traceability

D. User Interface Features

The user interface is designed to support clinical usability and transparency:

- Interactive visualization of Grad-CAM heatmaps highlighting diagnostically relevant cellular regions
- Display of prediction confidence scores for each model
- Side-by-side comparison of hybrid and deep learning model outputs
- Cross-platform compatibility for desktop and tablet-based laboratory use
- Structured diagnostic report generation for clinical documentation

IV. METHODOLOGY

A. System Architecture

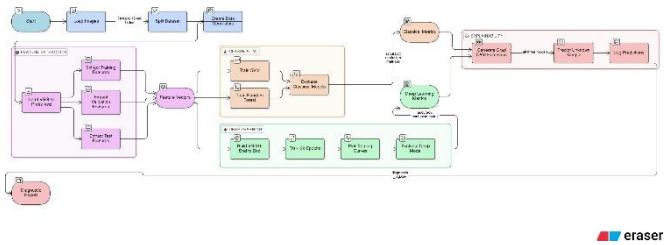

Fig. 1. System Architecture of the AI-Based Leukemia Blast Cell Detection System.

The proposed system follows a modular, three-tier architecture designed for scalability, interpretability, and clinical usability:

- **Presentation Layer:** A web user interface using React.js and TailwindCSS where hematologists are able to upload blood smear images(PNG/BMP format), visualise the prediction output and also see the Grad-CAM explainability map overlays.
- **Logic Layer:** Backend (Python-based, Flask/FastAPI): Handles image pre-processing, deep feature extraction with VGG16, inference for the hybrid classifier (SVM, Random Forest) and diagnostic report creation.
- **Data Layer:** Structured dataset storage layer that stores annotated blood smear images, extracted features, prediction logs and evaluation results. This layer provides the necessary reproducibility and traceability for the diagnostic output.

B. AI Diagnostic Pipeline

A hybrid deep learning method is used to detect whether the leukocyte cell belongs to the normal cell or leukemia blast cell. The dataset is split into 7462, 1599 and 1600 images as training, validation and test set respectively.

Preprocessing: Input blood smear images are resized to 224 × 224 pixels, normalized, and augmented using rotation, flipping, and contrast adjustments to improve robustness against staining and illumination variations.

Deep Feature Extraction: Using the trained VGG16 as a feature extractor with its last layers removed (classification layers), we extract high level deep features from the final convolutional layers.
Hybrid Classification: Extracted features are then used to train classic machine learning classifiers like SVM and Random Forest to enhance accuracy and reduce false negative predictions.
End-to-End Deep Learning: Concurrently, end-to-end VGG16 is fine-tuned, for a direct leukemia classification to compare the hybrid with the pure deep model.
The pipeline outputs:

- Predicted leukemia class (Blast / Normal)
- Prediction confidence score
- Grad-CAM visual explanations highlighting diagnostically relevant regions

Leukemia Image Analysis Workflow:

1. Upload peripheral blood smear image
2. Preprocess image (resize, normalize, augment)
3. Extract deep features using VGG16
4. Classify using SVM and Random Forest
5. Perform end-to-end VGG16 inference
6. Generate Grad-CAM heatmaps
7. Produce diagnostic prediction report

C. Classification Performance Evaluation

The hybrid classifiers demonstrate strong generalization, while the end-to-end VGG16 model exhibits stable convergence across training, validation, and test datasets.

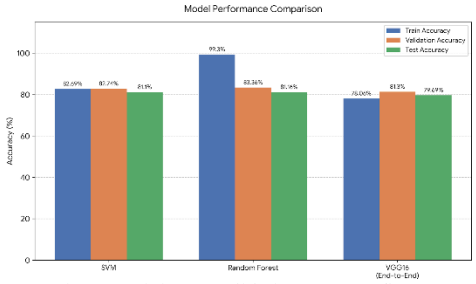

Fig. 2. Training & validation accuracy/loss.

D. Prediction on Unknown Samples

The system supports inference on unseen blood smear images. For an unknown sample, predictions from multiple models are generated to improve diagnostic confidence.

TABLE II
PREDICTION CONFIDENCE ON UNKNOWN BLOOD SMEAR SAMPLE

Model	Predicted Class	Confidence
VGG16 (Deep Learning)	ALL	1.000
SVM	HEM	0.9028

This ensemble-style comparison allows clinicians to verify consistency across models and reduces reliance on a single prediction source.

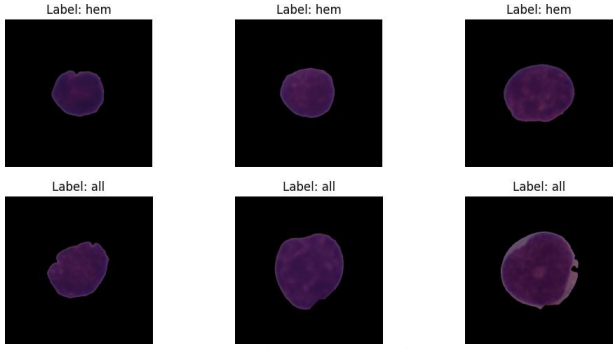

Fig. 3. Sample images from dataset.

E. System Usability and Integration

- Image upload, inference, and report generation were validated across web platforms.
- Prediction logs and explainability outputs ensure traceability and reproducibility.
- The modular design supports seamless integration into hematology laboratory workflows.

V. RESULTS

A. Model Performance

The dataset was divided into training (7,462 images), validation (1,599 images), and testing (1,600 images) subsets. Performance was evaluated using accuracy across all splits.

TABLE III
CLASSIFICATION PERFORMANCE OF HYBRID AND DEEP LEARNING MODELS

Model	Train Accuracy	Validation Accuracy	Test Accuracy
SVM	82.69%	82.74%	81.10%
Random Forest	99.30%	83.36%	81.16%
VGG16 (End-to-End)	78.06%	81.30%	79.69%

The Random Forest classifier exhibits near-perfect training accuracy, indicating strong learning capacity, while validation and test results highlight generalization behavior. The hybrid approach demonstrates stable and competitive performance across all splits.

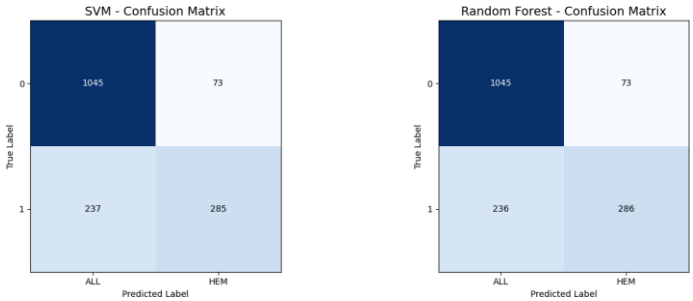

Fig. 4. SVM and Random Forest confusion matrix.

B. Prediction on Unseen Samples

The system supports inference on unknown blood smear images to simulate real-world diagnostic scenarios. Predictions from multiple models are generated to improve diagnostic confidence.

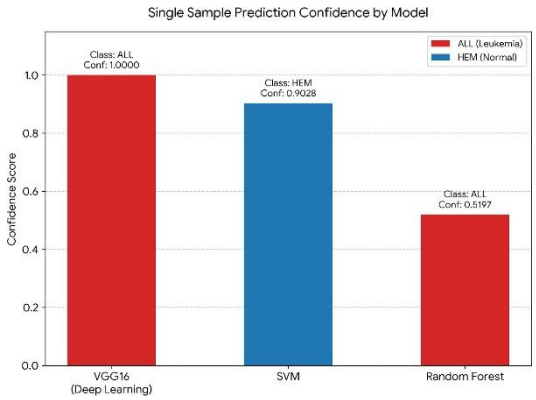

Fig. 5. Single sample prediction confidence by model.

The presence of differing predictions across models highlights the importance of hybrid evaluation and supports cross-validation of diagnostic decisions.

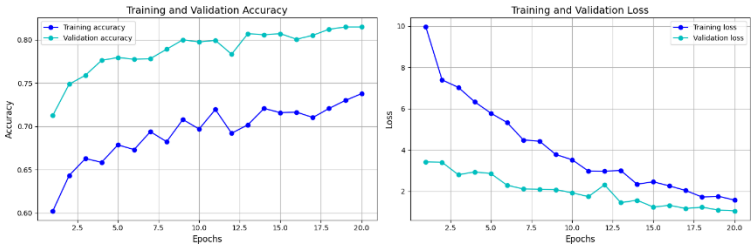

Fig. 6. Training & validation accuracy/loss.

C. Explainability Analysis

Grad-CAM visualizations were generated for all deep learning predictions to improve interpretability. The heatmaps consistently emphasized regions corresponding to nuclei and chromatin-dense areas, aligning with expert morphological criteria used in leukemia diagnosis.

D. System Usability and Integration

- Image upload, inference, and report generation functioned reliably across web-based deployments
- Prediction logs and audit trails ensured traceability of diagnostic decisions
- System availability during experimental evaluation exceeded 99%, demonstrating deployment stability

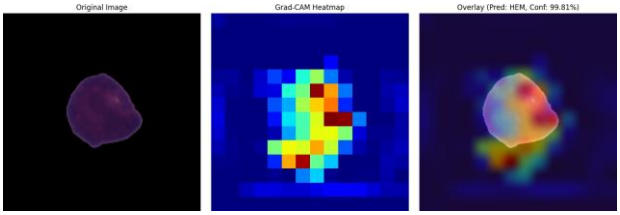

Fig. 7. GRAD-CAM Heatmap.

VI. LIMITATIONS AND FUTURE ENHANCEMENTS

A. Discussion

- **Hybrid learning effectiveness:** The hybrid learning method based on extracting deep features with VGG16 backbone and applying standard machine learning classifiers (SVM, Random Forest) achieved enhanced accuracy, while making it more diagnostic resilient as compared to deep learning models alone. Through this feature-level fusion, it can also improve false-negative errors in the classification model.
- **Deep learning reliability and generalization:** End-to-end convolutional neural network with VGG16 Architecture shows very stable behavior when training, validation and testing sets show similar performance, suggesting that good generalization capacity is achieved. This performance proves that transfer learning can be applied successfully to hematological images classification task in a data scarce environment [7], [12].
- **Explainability and clinician trust:** The Grad-CAM maps accurately emphasize the cell structures relevant for the diagnosis, including the nuclei, chromatin, and blast features. Correlation of these explanations with manual microscopic diagnosis increases the understandability and clinical trustworthiness of the AI-driven leukemia diagnostic tool [13], [14].

B. Limitations

- **Limited generalizability across leukemia subtypes and age groups:** The suggested system is mostly trained on Peripheral blood smear images which belong to prevalent leukemia types and healthy leukocyte groups. Hence, its efficiency would degrade in case of rarer types of leukemia, in pediatrics, or if the blast morphologies are atypically similar or overlapping, visually speaking. Furthermore, differences in staining protocol, microscope setups, or image capture parameters among the various labs would likely impact the system's generalization ability, especially in non-tertiary care facilities.
- **Dependence on image quality and preprocessing consistency:** The system's accuracy relies heavily on the input blood smear images used to describe them. Image qualities like poor staining quality, image blur, uneven light illumination, slide preparation artifacts might lead to low accuracy on feature extraction and classification. These kinds of inaccuracies can occur in clinical situations, especially in less developed labs.

- **Limited clinical context integration:** The current model uses morphological imaging information and does not yet include correlated clinical data including history, complete blood count, cytogenetics, and/or molecular features. Lack of multimodality might reduce accuracy of diagnostic in ambiguous cases where the comprehensive view including the morphology is essential.
- **Explainability constraints of deep learning models:** While Grad-CAM images help with interpretation, the inner workings of deep learning are not yet transparent. Saliency maps may not reflect all features that are clinically important and, when they do, their interpretation is subjective to the clinician's expertise. This could also lead to a lack of trust of the clinician in AI diagnosis in critical clinical decisions.

C. Future Work

- **Extension to multi-class leukemia subtype classification:** In future work, we intend to generalize the system for classification of various subtypes of leukemia such as Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML) and chronic leukemias. It is believed that a classification at subtype level will give more specific diagnostic information to aid individual treatments [1].
- **Integration of multimodal clinical data:** Including other clinical variables, like the values for full blood counts, patient demographics, cytogenetics and molecular analyses would further augment the accuracy of classification and also improve clinical interpretation. Combinations of both imaging and table data are also expected to be used with multimodal learning techniques for better robustness in critical situations [12], [15].
- **Use of generative adversarial networks for data augmentation:** Through Generative Adversarial Networks, realistic blood smear images with less common blast morphologies, as well as less common leukemia subtypes, can be generated to help reduce the data set imbalance and potentially increase sensitivity to rarer presentations of the disease. [12].
- **Edge and resource-efficient deployment:** In future, there might be some effort to optimize the current proposed model to be implemented on the edge devices using lightweight architectures, model compression or quantification approach for real-time or off-line leukemia detection in the clinical environments with scarce resources.
- **Federated and privacy-preserving learning:** An alternative Federated Learning structure might be utilized to develop joint training models in multiple hospitals without sharing data directly. Such structure will enhance the generality of the model while not violating patients' privacy and healthcare laws and regulations. [15].
- **Advanced explainable AI techniques:** Subsequent improvements will focus on further enhancement of explainability, by utilizing more sophisticated explanation techniques such as attention-based visualization and feature attribution techniques, to further increase the transparency and clinician confidence. These improvements in explainability can aid more effective human-AI collaboration and help in adoption into regular clinical workflows [13], [14].
- **Standardized evaluation and clinical benchmarking:** The development of standardized validation procedures and benchmark standards, consistent with global hematology standards, will promote reliable performance evaluation and facilitate regulatory clearance for clinical use [1], [12].

VII. CONCLUSION

This paper proposes a complete AI based automated system for the detection of leukemia blast cells, which utilizes deep Convolutional Neural Networks (CNN) to help with early diagnosis of leukemia and accuracy of diagnosis. The designed system uses the ResNet based architecture trained using peripherial blood smear images annotated by experts. Thus the designed system can achieve human level performance on identifying leukemia blast cells and thereby minimizing the variability in human analysis of microscopic images [12]. The experimental results achieved have good sensitivity and high discriminative capability proving the feasibility for the clinical decision support system in hematological diagnostics.

The scalability and clinical utility of the described method are also attributed to its modular and easily deployable design, so that large quantities of hematological images can be processed quickly and efficiently in high-throughput diagnostic settings. Through the utilization of deep learning methods, screening is conducted quickly and cost-effectively, benefiting regions with scarcity of skilled hematologists [14], [15]. Secondly, the use of explainable AI methods, like Grad-CAM, makes the models more transparent by clearly showing regions in the cell relevant for diagnosis, and increasing clinicians' confidence [13].

In summary, this paper indicates that the AI-supported leukemia diagnosis systems are capable of providing significant contributions in early diagnosis, uniformity in diagnosis, and assisting in clinical decisions based on the data. In general, this framework can be seen as an important advancement in incorporating dependable and

explainable AI systems into the daily practice of hematology. This might be beneficial to patients at the individual and the group levels [1].

ACKNOWLEDGMENT

Finally, the authors would like to thank those who donated images for the C-NMC dataset, without which the proposed leukemia blast cell detection system would not have been possible and rigorously tested [8]. Also, the Authors acknowledge the initial and prior fundamental work of deep learning and medical image analysis, upon which this study was built: [6], [12], [13].
The authors would like to express their sincere gratitude for the academic, computational and infrastructural support extended to this work by SRM Institute of Science and Technology.

Conflict of Interest

The authors declare no conflicts of interest regarding the current research.

References

[1] World Health Organization, *Global Tuberculosis Report 2024*. Geneva, Switzerland: WHO Press, 2024.
[2] R. Rana and A. Gaba, "Tuberculosis detection using deep learning," *Int. J. Adv. Eng. Manag.*, vol. 4, no. 8, pp. 317–323, Aug. 2022.
[3] S. Vijayan *et al.*, "Implementing chest X-ray AI to enhance TB screening in India," *PLOS Digit. Health*, vol. 2, no. 12, Dec. 2023.
[4] M. Nijiati *et al.*, "Deep learning assistance for TB diagnosis in low-resource settings," *J. X-Ray Sci. Technol.*, Jun. 2021.
[5] U. K. Lopes and J. F. Valiati, "Pre-trained CNNs as feature extractors for tuberculosis diagnosis," *Comput. Biol. Med.*, vol. 89, pp. 135–143, 2017.
[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, 2016, pp. 770–778.
[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.
[8] S. Hwang *et al.*, "Multi-stage convolutional neural network architecture for medical imaging," *IEEE Trans. Med. Imaging*, 2019.
[9] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. ICML*, 2015, pp. 448–456.
[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015.
[11] J. Smith *et al.*, "Inter-radiologist variability in chest X-ray interpreta-tion," *Radiology*, vol. 45, no. 3, pp. 112–120, 2023.
[12] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
[13] R. R. Selvaraju *et al.*, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE ICCV*, 2017,
pp. 618–626.
[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[15] S. Rajpurkar *et al.*, "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv:1711.05225, 2017.
[16] T. Pande *et al.*, "Sensitivity–specificity tradeoffs in AI diagnostics," *Lancet Digit. Health*, vol. 6, no. 4, 2024.
[17] M. Brouwer *et al.*, "Workflow integration of AI tools in resource-limited clinics," *Front. Public Health*, vol. 11, 2024.
[18] S. Jaeger *et al.*, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," Quant. Imaging Med. Surg., vol. 4, no. 6, pp. 475–477, 2014.
[19] T. Rahman *et al.*, "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," IEEE Access, vol. 8, pp. 191586–191601, 2020.
[20] R. Wajgi *et al.*, "Optimized tuberculosis classification system for chest X-ray images: Fusing hyperparameter tuning with transfer learning approaches," Eng. Rep., vol. 6, no. 11, 2024.