Deep Learning-Based HEp-2 Cell Classification for Autoimmune Disease Detection Using Transfer Learning

Sana Afrin[1], Purva Patil[2], Sejal Pawar[3], Tushar Mahajan[4], Prof. P. K. Kadbe[5]
[1,2,3,4]*Department of Information Technology, Sinhgad Institute of Technology and Science, Narhe, Pune, Maharashtra, India*
[5]*Head of Department (Guide), Department of Information Technology, Sinhgad Institute of Technology and Science, Narhe, Pune, Maharashtra, India*

Abstract*—Autoimmune diseases occur when the immune system wrongly attacks healthy tissues and organs. This causes chronic inflammation and serious health problems. Early diagnosis of autoimmune disorders is crucial for effective treatment and disease management. The Indirect Immunofluorescence (IIF) test on Human Epithelial Type-2 (HEp-2) cells is commonly used to detect Antinuclear Antibodies (ANA), which are key markers of autoimmune diseases. However, manually interpreting HEp-2 cell patterns takes a lot of time, can be subjective, and may vary from one observer to another. This research introduces an automated system for classifying HEp-2 cells using Transfer Learning and Deep Learning techniques. The proposed framework uses the MobileNetV2 architecture, which was pretrained on ImageNet, as a feature extractor for classifying six major HEp-2 staining patterns: Homogeneous, Speckled, Nucleolar, Centromere, Cytoplasmic, and Golgi. Image preprocessing techniques, such as denoising, contrast improvement using CLAHE, resizing, and normalization, are applied to enhance image quality and model performance. Data enhancement and fine-tuning are used to improve generalization and classification accuracy. Additionally, a confidence-based normal cell detection mechanism is introduced. Predictions that fall below a set confidence threshold are labeled as Normal or Unknown, which lowers the risk of false disease predictions. The experimental results show that the proposed approach effectively classifies HEp-2 cell patterns and aids in diagnosing autoimmune diseases.

***Index Terms*—Autoimmune Disease Detection, HEp-2 Cells, Transfer Learning, MobileNetV2, Deep Learning, Medical Image Classification, Biomedical Imaging, Computer-Aided Diagnosis.**

I. INTRODUCTION

Autoimmune diseases (ADs) represent a significant and growing global health burden, with more than 80 recognized autoimmune conditions affecting approximately 5–8% of the world's population [1], [2]. Common disorders such as Systemic Lupus Erythematosus (SLE), Rheumatoid Arthritis (RA), Sjögren's Syndrome, and Scleroderma share overlapping clinical manifestations, making early and accurate diagnosis particularly challenging. These conditions are characterized by an abnormal immune response wherein autoantibodies target the body's own healthy cells and tissues, leading to chronic inflammation, multi-organ damage, and systemic complications [3].

The primary laboratory method for autoimmune screening is the Antinuclear Antibody (ANA) test performed via Indirect Immunofluorescence (IIF) on Human Epithelial Type-2 (HEp-2) cell substrates [4]. The fluorescent staining patterns produced—such as homogeneous, speckled, nucleolar, centromere, cytoplasmic, and nuclear membrane—are diagnostically relevant and indicative of specific autoimmune disorders. For instance, a homogeneous pattern often correlates with SLE, while a centromere pattern is associated with limited scleroderma or CREST syndrome [1], [4]. However, manual interpretation of these fluorescence images is highly subjective, depending on the expertise of the observer and microscope calibration. Variability in light intensity, fixation, and staining further complicates consistent diagnosis [5].

Over the past decade, medical image analysis and Artificial Intelligence (AI) have been increasingly leveraged to address such diagnostic inconsistencies. Traditional computer-aided systems focused on handcrafted features—texture, shape, or intensity histograms—followed by classifiers like SVMs or

decision trees. While these approaches improved reproducibility, they struggled with complex patterns and inter-class similarities commonly observed in HEp-2 imagery [6]. To overcome these limitations, Deep Learning (DL) methods—particularly Convolutional Neural Networks (CNNs)—have emerged as the dominant paradigm in biomedical imaging [7]– [10].

CNNs automatically extract hierarchical spatial features, enabling the differentiation of subtle texture variations that may be imperceptible to the human eye. Research by Kumari and Singh [7] and Zhang et al. [8] demonstrated that CNN-based architectures significantly outperform conventional methods in HEp-2 pattern recognition, while Wang et al. [9] and Gupta et al. [10] showed that deep feature extraction and EfficientNet scaling further enhance classification robustness. Patel et al. [11] later highlighted the relevance of deep learning in autoimmune image processing, emphasizing improved diagnostic reliability.

Building on these advancements, this project proposes a fully integrated CNN-based HEp-2 image classification framework, enhanced with optimized preprocessing and a user-friendly GUI for clinical deployment. The system automates the entire workflow—from data preprocessing (denoising, normalization, and contrast enhancement) to image classification and visual result display—thus reducing human error, accelerating analysis, and aiding clinicians in early autoimmune disease detection.

Unlike existing approaches that primarily focus on the CNN model alone, the proposed framework emphasizes a practical implementation pipeline suitable for deployment in diagnostic laboratories and medical research environments.

II. LITERATURE REVIEW

Research on autoimmune disease diagnosis using artificial intelligence (AI) and deep learning (DL) has evolved significantly over the past decade. The traditional Indirect Immunofluorescence (IIF) technique using HEp-2 cell substrates remains the gold standard for identifying antinuclear antibodies (ANA); however, manual interpretation of these fluorescence patterns introduces subjectivity and inconsistency. Consequently, several researchers have explored automation using Convolutional Neural Networks (CNNs) and transfer learning for improved pattern classification and diagnostic precision.

The studies summarized in Table I highlight major contributions, methods, and limitations in the field, setting the foundation for the proposed CNN-based HEp-2 image classification and disease mapping framework.

TABLE I. – Literature Review of HEp-2 Cell Classification and Autoimmune Detection Systems

Paper No.	Techniques/ Approach	Key Findings / Limitations
[1]	Used Detectron2 and YOLOv8n for HEp-2 image detection and segmentation.	High accuracy for cell detection, but no explainability (XAI) and no disease mapping from patterns.
[2]	Used CNN and BiLSTM with Attention for T-cell data analysis	Good results for genetic prediction, but not focused on HEp-2 images.
[3]	Used CNN + MLP and Explainable AI (XAI) with Federated Learning	Promising for patient privacy and model sharing, but HEp-2 image errors (misclassification) remain.
[4]	Reviewed ML models for diagnosis and prognosis.	Good summary but little work on real HEp-2 data integration.
[5]	Used PELISA Biosensor for lab antibody testing	Accurate for lab detection but not image-based or automated.
[6]	CNN used for detecting skin lesions.	Proved CNN works well on medical images, but needs adaptation for HEp-2 fluorescence data.
[7]	Used texture and PCA features with traditional classifiers.	Improved basic accuracy but not scalable or deep-learning based.
[8]	Reviewed CBC, ESR, CRP and ANA tests.	Explained the lab process but no AI or automation used.
[10]	Manual HEp-2 pattern detection in labs.	Established HEp-2 as gold standard, but fully manual.

Paper No.	Techniques/ Approach	Key Findings / Limitations
[11]	Used PCA, Random Forest, SVM and Neural Networks on mixed datasets.	Useful for statistical analysis, but limited image work and still prone to pattern confusion.
[12]	Combined CNN, LSTM, and Graph models (ImmunoNet).	Powerful multi-layer approach, but no focus on HEp-2 classification.
Proposed Project Work	CNN-based HEp-2 image classifier + preprocessing + GUI + disease mapping.	Adds explainability, disease mapping, and GUI interface missing in earlier work.

The reviewed literature indicates significant progress in applying artificial intelligence and deep learning to autoimmune disease diagnosis, particularly for HEp-2 cell image analysis. Traditional laboratory methods, such as those discussed by Castro (2019) and Doghramji (2019), confirmed the diagnostic value of antinuclear antibody (ANA) testing but were limited by subjectivity and manual interpretation.

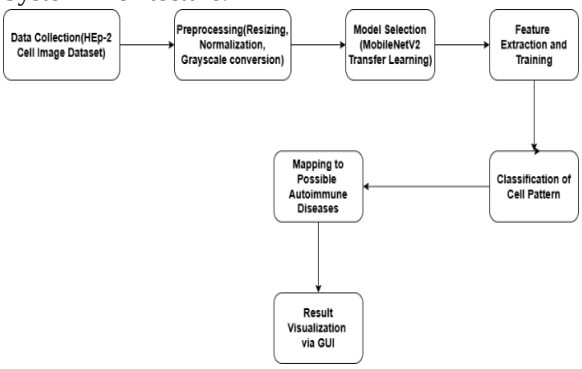
Subsequent studies, including those by Campadelli et al. (2021) and Kumari & Singh (2020), demonstrated that CNNs and classical feature extraction techniques can effectively process medical images, though these methods lacked full automation and clinical integration. Broader reviews by Miho et al. (2024) and Danieli et al. (2024) highlighted machine learning's potential in autoimmune diagnosis but offered limited focus on fluorescence-based imaging.

The work of Subramana et al. (2024), adopted as the base paper for this study, directly addressed HEp-2 image detection and segmentation using advanced models such as Detectron2 and YOLOv8n. While achieving high accuracy, it did not incorporate explainable AI or disease-pattern mapping, both essential for clinical interpretability. Later studies by Yang et al. (2025) and Ullah et al. (2025) introduced multimodal and explainable architectures, yet they faced challenges with visually similar fluorescence patterns and lacked specific adaptation for HEp-2 datasets.

Overall, existing approaches remain fragmented—many foci solely on classification or detection without establishing an interpretable, integrated diagnostic system. The present work builds upon these findings by developing a CNN-based HEp-2 classification framework enhanced with explainability features, disease mapping, and a graphical user interface to support transparent and practical clinical application.

III. SYSTEM ARCHITECTURE

System Architecture:



The proposed system architecture illustrates the complete workflow for automated HEp-2 cell pattern classification and autoimmune disease prediction. The process begins with Data Acquisition, where HEp-2 cell images are collected from publicly available immunofluorescence datasets. These images serve as the input for model training and testing.

In the Preprocessing stage, images are resized to a fixed dimension and normalized to ensure consistency across the dataset. Additional image enhancement techniques such as denoising and contrast enhancement are applied to improve image quality and highlight important cellular features. These preprocessing operations help the model focus on relevant fluorescence patterns while reducing noise and variability.

The preprocessed images are then passed to the Model Selection stage, where a deep learning-based classification model is employed. In the proposed system, the MobileNetV2 transfer learning architecture is utilized as a feature extractor due to its efficiency and strong performance in image classification tasks. The model learns discriminative representations of HEp-2 cell patterns from the training dataset.

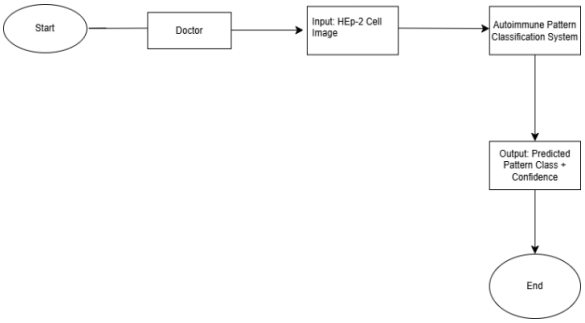
During the Feature Extraction and Training phase, deep features are automatically extracted from the images and used to train the classification layers. Data augmentation and fine-tuning techniques are incorporated to improve model generalization and classification accuracy while reducing overfitting.

After training, the model performs Classification of Cell Patterns, categorizing HEp-2 images into predefined fluorescence pattern classes such as Homogeneous, Speckled, Nucleolar, Centromere, Cytoplasmic, and Golgi. These patterns are clinically significant indicators of different autoimmune disorders.

The classified pattern is then processed by the Mapping to Possible Autoimmune Diseases module, which associates each fluorescence pattern with autoimmune diseases commonly linked to that ANA pattern according to established medical literature. This mapping enables the system to provide clinically meaningful diagnostic insights rather than only displaying the pattern classification.

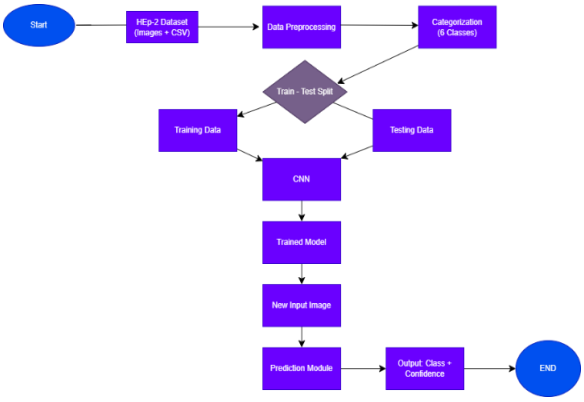
Finally, the results are presented through the Result Visualization via GUI module. A user-friendly graphical interface allows healthcare professionals to upload HEp-2 images, view the predicted fluorescence pattern, confidence score, and corresponding autoimmune disease associations. This facilitates efficient interpretation and supports computer-aided diagnosis of autoimmune diseases.

DFD Level 0:



The Level 0 Data Flow Diagram represents the system at a very high level, treating it as a single process. In this project, the doctor provides an input in the form of a Hep-2 cell image. This image is passed into the autoimmune pattern classification system, which processes the image using a trained deep learning model. Finally, the system outputs the predicted fluorescence pattern class along with a confidence score.

DFD Level 1:



It starts with the Hep-2 dataset containing images and CSV annotations. Image is preprocessed by converting them to greyscale, resizing them and extracting individual cells. Then, it is categorized into 6 classes based on pattern information. After that, the dataset is split into training and testing sets. Once model is trained it is used to predict the class of a new input image and outputs the predicted pattern along with a confidence score.

IV. METHODOLOGY

The proposed system employs a structured deep learning framework for the automated classification of HEp-2 cell fluorescence patterns and the prediction of associated autoimmune diseases. The methodology consists of multiple stages, including dataset acquisition, image preprocessing, transfer learning-based classification, disease mapping, confidence evaluation, and result visualization.

Initially, labelled HEp-2 cell images representing various fluorescence patterns are collected from standard immunofluorescence datasets [3], [4]. To improve image quality and maintain consistency across samples, preprocessing techniques such as image resizing, normalization, non-local means denoising, and Contrast Limited Adaptive Histogram Equalization (CLAHE) are applied. These operations enhance cellular structures and improve the visibility of fluorescence patterns, enabling more effective feature extraction [7].

Following preprocessing, a Transfer Learning approach based on the MobileNetV2 architecture is employed for pattern classification. MobileNetV2, pretrained on the ImageNet dataset, serves as a powerful feature extractor capable of learning

complex visual representations. The pretrained convolutional layers are initially frozen, while custom classification layers consisting of Global Average Pooling, Dense, Dropout, and SoftMax layers are added to classify HEp-2 cell images into six fluorescence pattern categories. To improve model generalization and reduce overfitting, data augmentation techniques including rotation, zooming, horizontal flipping, and image shifting are applied during training [10], [11].

After the initial training phase, fine-tuning is performed by unfreezing selected higher layers of the pretrained network and retraining them using a lower learning rate. This enables the model to learn domain-specific HEp-2 cellular features while retaining the advantages of pretrained knowledge, resulting in improved classification performance [10], [11].

Once training is completed, the model predicts the fluorescence pattern of new HEp-2 cell images and assigns them to predefined classes such as Homogeneous, Speckled, Nucleolar, Centromere, Cytoplasmic, or Golgi. A disease mapping module then associates each identified pattern with autoimmune diseases documented in medical literature, providing clinically relevant diagnostic insights [1], [2], [13].

To enhance reliability, a confidence-based evaluation mechanism is incorporated. If the highest prediction confidence falls below a predefined threshold, the sample is categorized as Normal or Unknown rather than being forcibly assigned to a disease class. This approach reduces false-positive predictions and improves the practical applicability of the system in real-world diagnostic scenarios [3], [12].

Finally, the classification results, confidence scores, and corresponding disease associations are presented through a user-friendly Graphical User Interface (GUI) developed using Tkinter. The interface enables healthcare professionals to upload HEp-2 images, perform automated analysis, and obtain interpretable diagnostic information efficiently. The overall framework provides an effective computer-aided diagnostic solution for autoimmune disease detection and HEp-2 cell pattern classification [11], [12].

V. RESULTS AND DISCUSSION

The model was trained and validated on the collected HEp-2 cell image dataset using an 80:20 train-test split, ensuring balanced evaluation. During experimentation, the CNN achieved strong performance metrics, including high overall accuracy and an improved precision–recall balance compared to baseline models [6], [8].

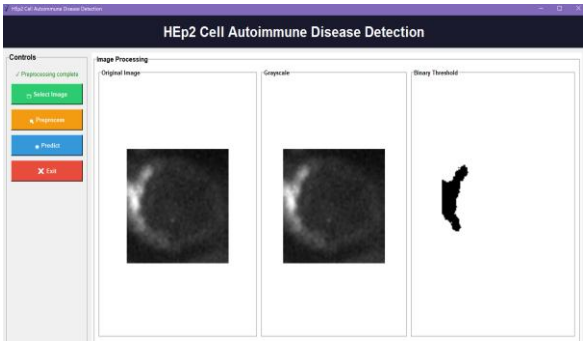
The training process converged smoothly over multiple epochs, confirming effective feature learning and minimal overfitting due to the inclusion of dropout and normalization layers. Evaluation metrics such as accuracy, precision, recall, and F1-score were calculated to assess model reliability and generalization capabilities.

Visual analysis using sample predictions demonstrated that the CNN could accurately classify common fluorescence patterns, distinguishing between homogeneous and speckled textures with minimal error. The proposed system's output was also integrated into the GUI interface, providing clear visual indicators of the classified cell pattern and the corresponding autoimmune disease mapping.

Compared to the baseline study by Subramana et al. [6], the inclusion of disease mapping and GUI components enhances interpretability and usability for real-world medical applications. These results indicate that deep learning can serve as a practical and accurate diagnostic support tool in laboratory environments.

Furthermore, incorporating explainable AI (XAI) modules in future versions can enhance model transparency and build trust among clinicians [13], [15].

GUI page for prediction



Prediction Results Output

VI. DISEASE PATTERN MATCHING

The disease mapping component serves as a crucial interpretability bridge between computational classification and clinical diagnosis. Each fluorescence pattern classified by the CNN corresponds to a characteristic set of autoimmune diseases recognized in immunology [1], [2], [3].

For example, a Homogeneous pattern is typically associated with Systemic Lupus Erythematosus (SLE) or Drug-Induced Lupus, while a Centromere pattern is indicative of Limited Scleroderma (CREST Syndrome). Similarly, Speckled patterns may correspond to conditions such as Sjögren's Syndrome, Mixed Connective Tissue Disease, or Rheumatoid Arthritis.

By implementing this pattern-to-disease relationship, the system extends beyond pattern classification to provide meaningful diagnostic insights. This mapping is grounded in established clinical studies and immunopathological literature [1], [2], [13], allowing healthcare practitioners to interpret model predictions within a medical context.

The integration of disease mapping into the user interface ensures that the end user—whether a clinician, pathologist, or medical researcher—can rapidly translate computational outputs into actionable clinical information [12], [13].

VII. CONCLUSION AND FUTURE WORK

The proposed system demonstrates an effective deep learning-based framework for the automatic classification of HEp-2 cell immunofluorescence images and the preliminary detection of autoimmune diseases. Through careful preprocessing, CNN-based classification, and disease mapping, the project bridges the gap between computational image analysis and medical diagnostics [5], [6], [8].

The developed GUI enhances system accessibility and supports real-time visualization, making it suitable for both laboratory and telemedicine applications.

Future work will focus on improving generalization through larger and more diverse datasets, incorporating transfer learning with architectures such as ResNet, VGG, and EfficientNet [10], [11]. Additionally, Explainable AI (XAI) visualization techniques like Grad-CAM will be integrated to highlight diagnostically relevant regions, improving interpretability for clinicians [13], [15].

Finally, a web-based deployment with secure data management and federated learning approaches may enable collaborative medical diagnostics across institutions.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to Prof. P.K. Kadbe, Department of Information Technology, Sinhgad Institute of Technology and Science, Narhe, Pune, for his valuable guidance, continuous support, and constructive suggestions throughout the development of this research work.

The authors also thank the Department of Information Technology and the institution for providing the necessary resources, facilities, and academic environment required for the successful completion of this project. Finally, we acknowledge all researchers and organizations whose published work and datasets contributed to this study on HEp-2 cell classification and autoimmune disease detection.

REFERENCES

[1] Castro, "Diagnostic testing for autoimmunity," 2019.
[2] Doghramji, "ANA screening for autoimmune disease," 2019.
[3] Danieli et al., "Machine learning application in autoimmune diseases: State of art and future prospectives," 2024.
[4] Hsu and Huang, "Detection of autoimmune antibodies in body fluid," 2020.
[5] Yang et al., "Deep learning-based prediction of autoimmune diseases," 2025.
[6] Subramana et al., "Detection of auto-immune disease using deep learning techniques," 2024.
[7] Kumari and Singh, "Feature extraction in medical images using Haralick and PCA," 2020.
[8] Zhang et al., "HEp-2 cell image classification using CNN-based architectures," 2023.
[9] Wang et al., "Enhanced deep feature extraction for HEp-2 image classification," 2023.
[10] Gupta et al., "EfficientNet scaling for biomedical image classification," 2023.
[11] Patel et al., "Deep learning for autoimmune image processing," 2024.

[12] Ullah et al., “Advancing personalized diagnosis and treatment using deep learning architecture,” 2025.
[13] Miho et al., “Machine learning for precision diagnostics of autoimmunity,” 2024.
[14] Campadelli et al., “Automated skin nodule detection using deep learning,” 2021.
[15] Teniou and Rhouati, “Recent advances in biosensors for diagnosis of autoimmune diseases,” 2022.