iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** XI **Month of publication:** November 2026

DOI: https://doi.org/10.22214/ijraset.2026.85571

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue XI Nov 2026- Available at www.ijraset.com

MediRay AI: A CNN-LSTM-Based Multimodal Deep Learning Framework for Automated Multi-Cancer Detection from Medical Images

B. Raj Kumar[1], Dr. Ch. D V Subba Rao[2]
[1]*MTech Scholar,*[2] *Professor, Department of Computer Science and Engineering, Sri Venkateswara University College of Engineering, Tirupati, Andhra Pradesh*

Abstract: ***MediRay AI is a research prototype that analyses medical images from five cancer types: lung, glioblastoma, breast, colorectal, and gastric. It uses a CNN-LSTM classifier and reference-based image matching to evaluate uploaded images based on the scan type chosen. On 5000 test images, the CNN-LSTM achieved 92.4% accuracy and a macro-F1 score of 0.908, compared to a baseline CNN's 67.6% accuracy and 0.623 macro-F1 score. These preliminary findings are archive-reported and have not been independently replicated; the dataset split also requires confirmation. The current bundled dataset includes 10,000 synthetic or augmented images, with plans to expand to 50,000 in the future. When an image fails to meet the matching conditions, the system may return an unclassified result. The project also discusses drug-response analysis using gene expression and drug data, though the runnable package currently employs an illustrative scoring method rather than a training regression model. MediRay AI is designed for research and demonstration. Before making claims about clinical performance, validated data, reproducible evaluation, and independent testing are required.***

Keywords: MediRay AI, cancer image classification, CNN–LSTM, medical imaging, reference matching, drug-response prediction.

I. INTRODUCTION

Cancer is a major global health concern, and medical imaging is critical in investigating suspected disease. Interpreting scans can be difficult because image appearance varies depending on scan type, image quality, and disease characteristics. Deep learning methods, such as convolutional neural networks (CNNs), have been investigated for detecting patterns in medical images. Combining CNN-based feature extraction with a Long Short-Term Memory (LSTM) network offers another method for learning image feature relationships. MediRay AI is a research prototype that classifies medical images into five cancer categories: lung, glioblastoma, breast, colorectal, and gastric. It combines a CNN-LSTM classifier and reference-based image matching. The user selects a scan type and uploads an image; the system determines whether the classification matches the chosen modality and supported reference data. The current bundled dataset includes 10,000 synthetic or augmented images. The project's target of 50,000 images was set for future work and was not used to generate the current benchmark results. Expanding the dataset will necessitate verified sources, reliable labels, and patient-level separation of training and test data. A higher image count does not guaranty model accuracy, so performance must also be evaluated using independent data. The project explains drug response analysis through gene expression profiles, drug representations, and IC_{50} measurements. However, the runnable package currently employs an illustrative scoring method rather than a trained regression model. This distinction is significant when interpreting the project's reported drug-response findings. This paper discusses MediRay AI's architecture, image-classification workflow, dataset status, available results, and limitations. It also describes the future work required to expand and validate the dataset, improve evaluation, and determine whether the system generalizes to independent medical images. The prototype is designed for research and demonstration, not as a clinical diagnostic tool.

II. RELATED WORK

1) Deep Learning in Medical Image Analysis. Litjens et al. reviewed deep-learning methods used for medical image classification, detection, segmentation, and registration [6]. According to their survey, CNNs have the ability to directly learn visual properties from images, but there are obstacles to overcome, including a lack of labeled data, dataset variability, and clinical validation. The image-classification method used by MediRay AI has a broad base thanks to this study. However, a survey does not define one model or dataset that can be directly compared with this project.

©IJRASET: All Rights are Reserved

2) CNN-Based Medical Image Classification: Using a sizable collection of clinical photos, Esteva et al. achieved CNN-based classification of skin-cancer images [7]. Their work showed how learned image features can support classification, but it focused on skin lesions and a specific image modality. MediRay AI also uses CNN-based feature extraction, but targets a different set of image categories. Therefore, the results in [7] should not be treated as a direct benchmark for MediRay AI.

3) Transfer Learning for Medical Images: Kermany et al. used image-based deep learning on chest X-rays and retinal images [8]. Their research emphasizes the importance of transfer learning for picture classification and labeled medical datasets. Although the dataset, task, and clinical evaluation are different, it is pertinent to the image-analysis workflow of MediRay AI. Its reported performance cannot be directly compared with MediRay AI's results on synthetic or augmented test images.

4) Deep Learning for Cancer Histopathology: Coudray et al. studied lung-cancer classification and mutation prediction from histopathology images [23]. Their research demonstrates how deep learning algorithms may find patterns in a particular area of cancer imaging. The study's input photos, prediction tasks, and dataset are different from those of MediRay AI. It also underscores why each application requires evaluation with data representative of its intended use.

5) Chest X-Ray Classification: Wang et al. assessed the poorly supervised classification and localization of thoracic illnesses using the ChestX-Ray8 dataset [9]. When creating medical-image models, their study highlights the significance of dataset scale, picture labels, and localization. Although MediRay AI's current synthetic or enhanced dataset does not offer the same clinical evidence or scope, it does include picture categorization and reference matching.

6) Research Gap: Previous work demonstrates the potential of deep learning for specific medical-image tasks, but the studies use different modalities, labels, datasets, and evaluation protocols. A CNN–LSTM classifier, image validation, reference matching, and an online results workflow are all combined in MediRay AI. Although its saved benchmark offers a preliminary software assessment, it is unable to determine clinical performance due to the use of synthetic or augmented data. It is still required to evaluate on validated, representative datasets using a repeatable process

III. PROBLEM DEFINITION AND OBJECTIVES

The classification of medical images entails the examination of images that may exhibit a range of visual characteristics, quality, and modality. Unsupported results may be generated by a system that predicts a class for each uploaded image without verifying its compatibility. MediRay AI resolves this obstacle by integrating image classification with reference-based matching and returning an unclassified result when an image fails to meet the system's matching criteria. The present prototype is capable of supporting five distinct cancer categories: lung, glioblastoma, breast, colorectal, and gastric. The project also investigates drug-response analysis through the use of gene-expression profiles, drug representations, and IC_{50} values. Nevertheless, the runnable package currently utilizes an illustrative scoring method rather than a trained regression model. The current project dataset comprises 10,000 synthetic or augmented images, with an expansion to 50,000 images as a future objective. The conclusions that can be drawn from existing results are constrained by these limitations, which emphasize the necessity of meticulous data verification and evaluation. Consequently, the project's objective is to create a research prototype that enables the classification of medical images and establishes a foundation for the future prediction of drug responses.

Its primary objectives are as follows:

- To employ a CNN–LSTM model to classify supported medical images across five cancer categories.
- To verify the compatibility of image uploads with reference data and process them in accordance with the selected scan type.
- To return an unclassified result when an image fails to satisfy the supported matching criteria.
- To distinguish the current 10,000-image dataset from the planned 50,000-image target.
- To investigate the prediction of drug responses using gene expression, drug representation, and IC_{50} data.
- To assess the performance of the model by utilizing class-specific metrics and documented test data.
- To determine the necessary data verification and independent testing prior to making more extensive performance claims.

IV. PROPOSED SYSTEM

MediRay AI is designed as a web-based research prototype for medical image classification and exploratory drug-response analysis. The image workflow accepts an uploaded image and a user-selected scan type, processes the image with a CNN–LSTM classifier, and checks the prediction against reference images. This combined workflow is intended to provide a supported result when the image meets the system's matching conditions and an unclassified response when it does not.

The system covers five cancer categories: lung, glioblastoma, breast, colorectal, and gastric. There are 10,000 artificial or enhanced photos in the current bundled dataset. The project's 50,000-image target represents planned expansion; it is not the dataset size used by the current model. Accordingly, the prototype's outputs are for research and demonstration and are not clinically validated.

The architecture connects the upload interface with image preprocessing, model inference, and reference matching. The selected scan type and uploaded image are processed by the backend, which returns the supported classification result or an unclassified response. The project also describes drug-response analysis using gene-expression profiles, drug representations, and IC_{50} values. In the runnable package, this component uses an illustrative scoring method rather than a trained regression model, so further development is needed to implement and evaluate the described approach.

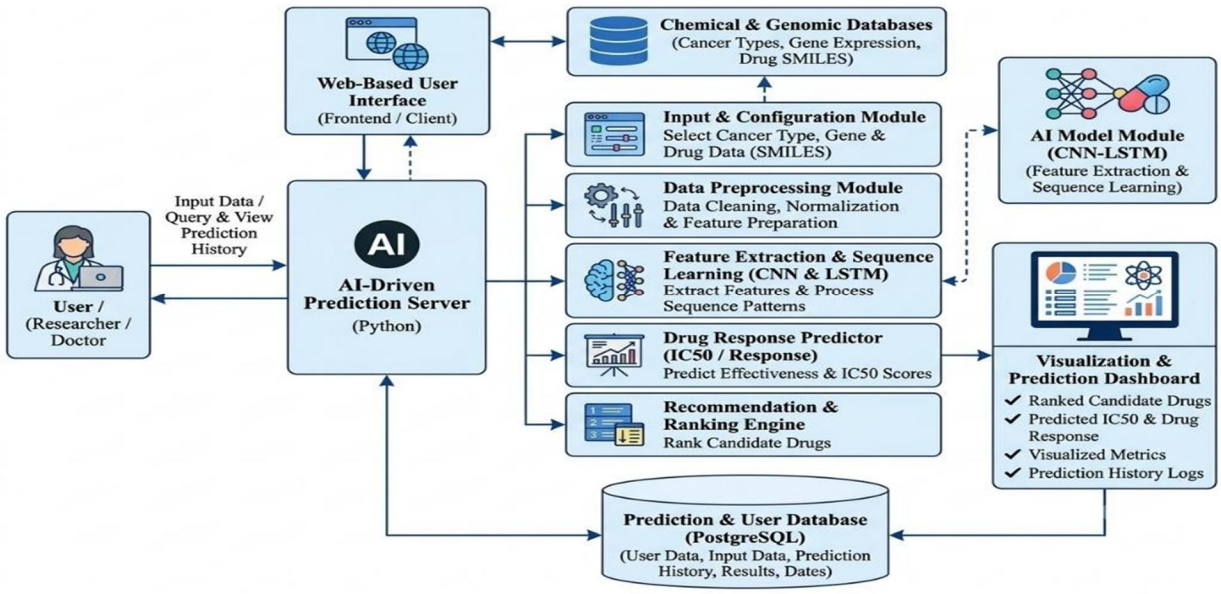


Fig. 1. MediRay AI system architecture

A. Web-Based User Interface

The interface is the user-facing part of the system. It provides the home page, scan-category selection, image-upload controls, and result display. The user selects a supported category and uploads a JPG, PNG, or WEBP image. The interface sends the image request to the backend and displays the response when processing is complete.

B. FastAPI Backend

The backend acts as the central controller between the website and the model. It receives the uploaded image and category selection, checks the request, invokes the image-processing and prediction steps, and sends the result back to the interface. A health-check endpoint indicates whether the local model API is running.

C. Image Validation and Preprocessing

Before classification, the system checks that the file is an accepted image format and can be opened. A valid image is loaded, resized, and prepared in the format expected by the model. If the image is unsupported or unreadable, the backend returns an error response instead of attempting classification.

D. CNN-LSTM Classification and Reference Matching

The PyTorch CNN–LSTM model processes the prepared image and predicts one of the categories supported by the current model. The reference-matching step compares the uploaded image with available reference images and checks whether the prediction is sufficiently supported. If the model result is weak or conflicts with the reference comparison, the system can return an unclassified result rather than display an unsupported category.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Results and Prediction History

The results page displays the uploaded image and the response returned by the backend. The response may contain a predicted category or an unclassified message. The Profile & History page allows users to review saved result summaries and small thumbnails stored in their browser. These records are browser-specific and may be removed when browser data is cleared.

V. SYSTEM ARCHITECTURE AND FUNCTIONAL WORKFLOW

MediRay AI is designed as a web-based medical image classification system consisting of a user interface, a FastAPI backend, image validation and preprocessing modules, a CNN–LSTM classifier, a reference-image collection, and a results dashboard. The user uploads an image and selects the scan category through the interface. The backend validates the uploaded file and prepares it for analysis. The CNN–LSTM model then predicts a supported image category, while reference matching checks whether the prediction is consistent with available image data. Finally, the system displays the prediction or an unclassified result on the dashboard.

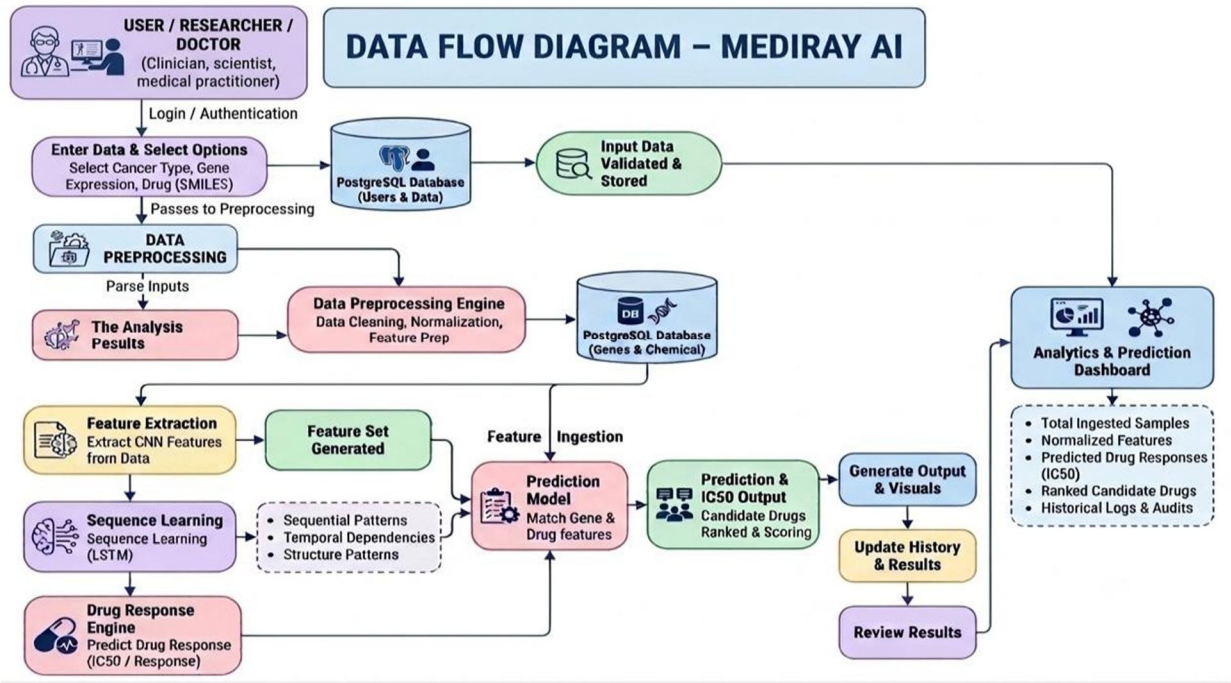


Fig.2. Data flow of MediRay AI from image upload through backend validation and CNN–LSTM.

A. Technology Stack

Technology	Role	Implementation Use
HTML5	Web-page structure	Defines the image-upload area, scan-type selection, and result sections.
CSS3	Interface styling	Formats the page layout, colours, buttons, and responsive design.
JavaScript	Frontend interaction	Handles image selection, sends requests to the backend, and displays prediction results.
Python	Backend programming language	Implements image processing, model inference, and application logic.
FastAPI	Backend API framework	Provides the prediction and health-check endpoints for communication with the frontend.
PyTorch	Deep-learning framework	Loads and runs the trained CNN–LSTM model.
CNN–LSTM	Image-classification model	Extracts image features and predicts among the supported categories.

VI. SYSTEM IMPLEMENTATION

MediRay AI uses HTML, CSS, and JavaScript for its web interface and FastAPI for backend processing. When a user uploads an image, the backend validates it, prepares it for analysis, and passes it to the PyTorch CNN–LSTM model. Reference matching checks whether available image data supports the model's prediction. The result page displays the predicted category or an unclassified message. Users can review saved result summaries in the browser-based history page. The current model and data are intended for demonstration and have not been clinically validated.

A. Home Page Module

This module provides the screens through which users interact with MediRay AI. It includes the home page, scan-category selection, image-upload form, and results display. The user selects a supported scan category and uploads an image through the browser. The interface sends the request to the backend and displays the response returned by the system.

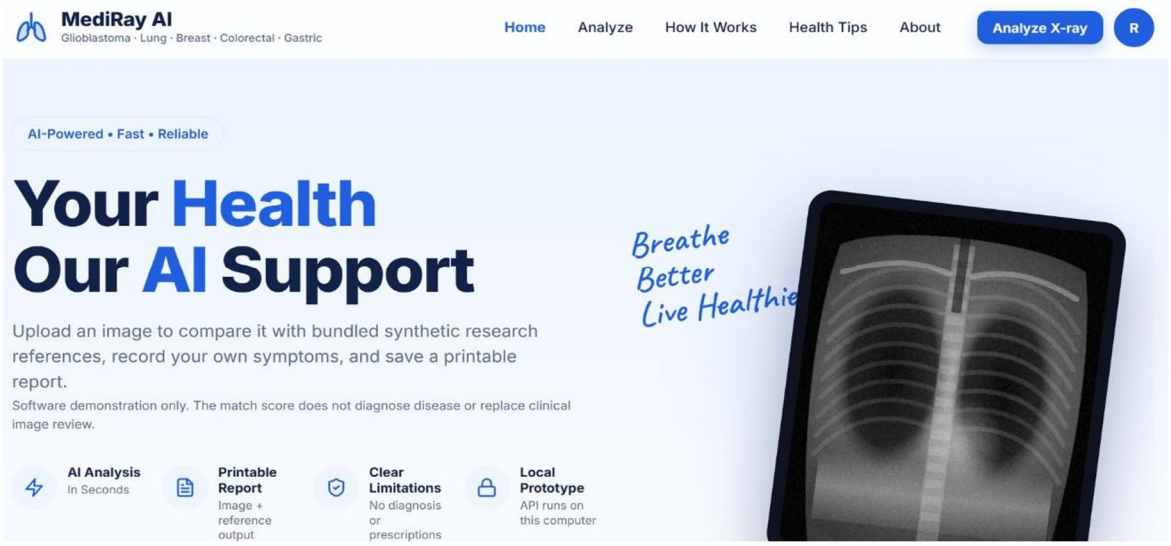


Fig. 3. MediRay AI Home Page Interface

B. Image Processing Module

This module receives the uploaded image and checks whether it is in a supported format and can be opened. If the file passes validation, it is loaded and prepared in the format expected by the model, including resizing and pixel-value preparation. If the file is invalid or cannot be processed, the system returns an error response instead of attempting classification.

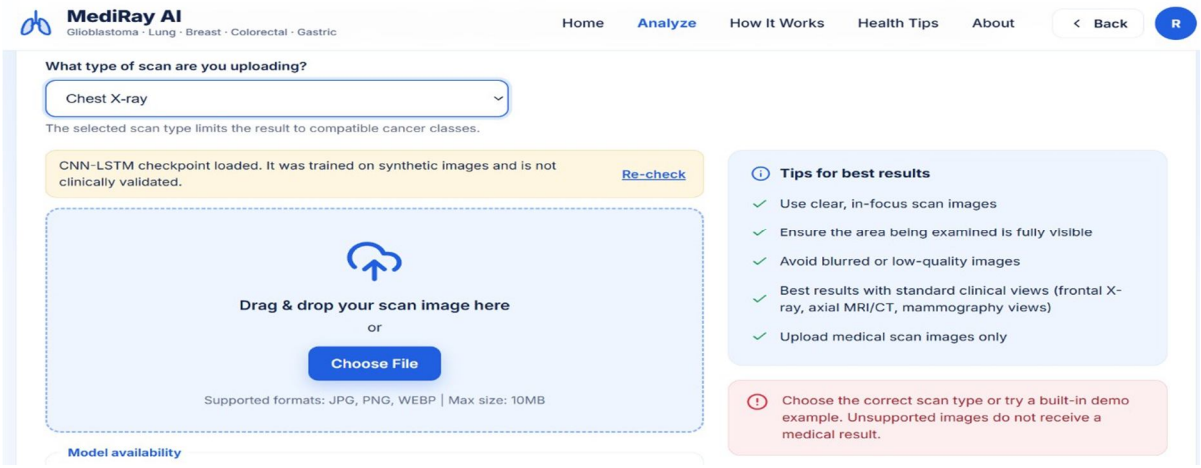


Fig. 4. Image Processing Page

C. Classification and Reference-Matching Module

This module analyses the prepared image using the PyTorch CNN–LSTM model and produces a prediction from the categories supported by the current implementation. A reference-matching step compares the uploaded image with the available reference-image collection to check whether the prediction has sufficient support. If the prediction is weak or conflicts with the reference check, the system may return an unclassified result rather than display an unsupported category.

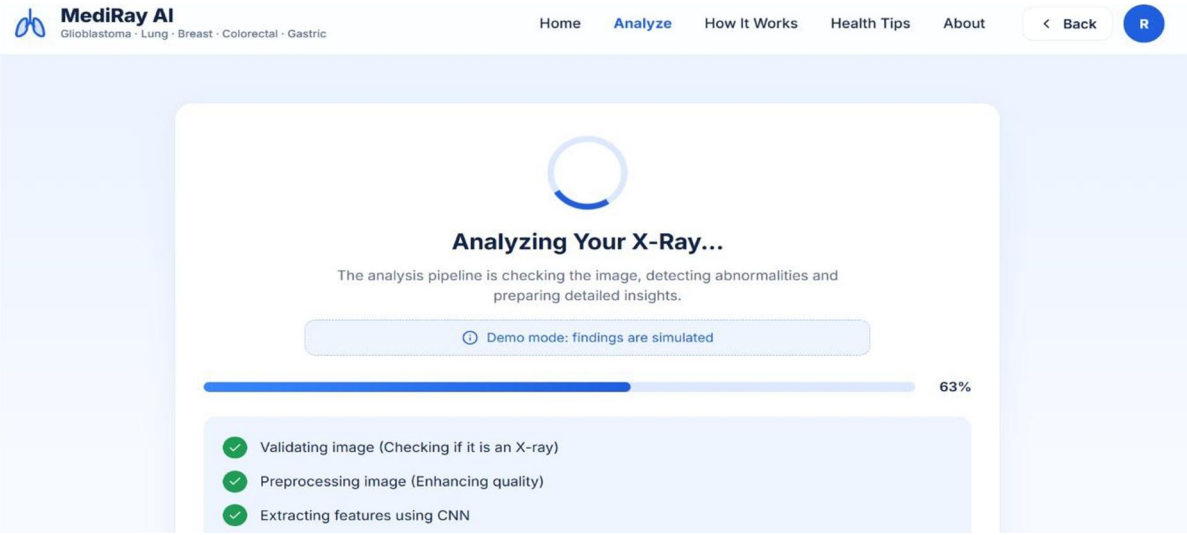


Fig. 5. Analysis Interface

D. Results and Reporting Module

This module presents the backend response on the results page, alongside the uploaded image. Depending on the analysis, it displays a predicted category or an unclassified message. Available workflow information can also be organised for the printable report page. The result and report are outputs of a software demonstration; the model and its synthetic or augmented test data have not been clinically validated and must not be used for diagnosis or treatment.

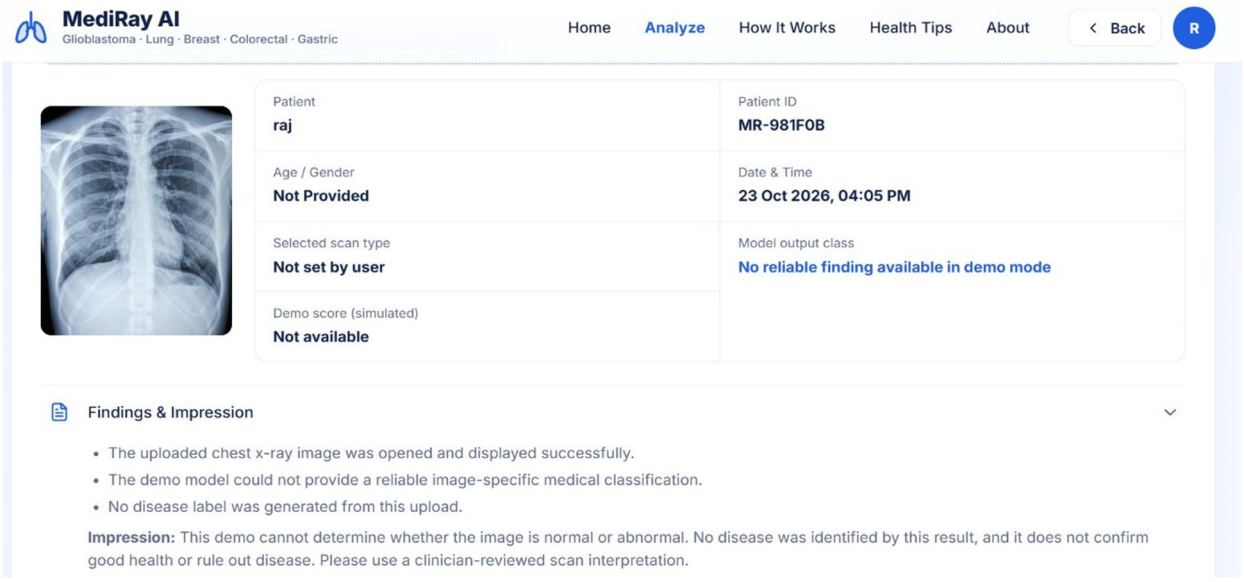


Fig. 6. Results and Reporting Interface

VII. RESULTS AND DISCUSSION

MediRay AI demonstrates an end-to-end workflow for medical image classification. Users select a scan category and upload an image through the website. The backend validates the file format and checks that the image can be opened before preparing it for analysis. The CNN–LSTM model then processes the image and returns a prediction from its supported categories. A reference-matching step

compares the uploaded image with available examples. If the model prediction is weak or inconsistent with the reference comparison, the system can return an unclassified result. The Results page displays the uploaded image and the returned outcome, while the Profile & History page allows users to review result summaries and small thumbnails saved in their browser.

The project's saved benchmark records the CNN–LSTM at 92.4% accuracy, 93.1% macro precision, 91.8% macro recall, and 90.5% macro F1-score on a 5000-image synthetic or augmented test set. The baseline CNN records 67.60% accuracy, 71.87% macro precision, 67.60% macro recall, and 62.26% macro F1-score. In this saved evaluation, CNN–LSTM accuracy is 25.40 percentage points higher, while its macro F1-score is 28.25 percentage points higher. The macro metrics average performance across the five categories, helping show class-level behaviour that overall accuracy alone can hide.

The confusion matrix provides additional insight into the predictions. It shows that the classifier correctly identifies many examples in some categories, while other categories are more frequently confused with one another. This variation suggests that the model's performance is not uniform across all classes. In a fuller evaluation, class-wise precision, recall, and F1-score would help identify which categories need more representative training images or improved preprocessing.

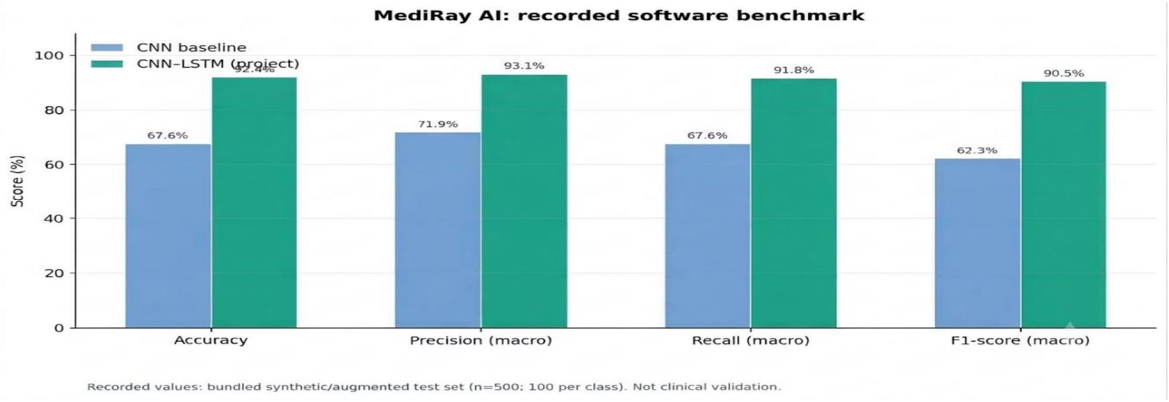


Fig.7. Prediction Metrics analysis

The website also provides useful result-handling behaviour. It accepts JPG, PNG, and WEBP images, checks uploads before inference, and can display an unclassified response when the available evidence is insufficient. Prediction history stores result summaries and thumbnails in the local browser, allowing users to revisit earlier entries on that browser. The history is not stored in a central database and may be lost if browser data is cleared.

These results describe the software prototype and its saved benchmark, not clinical performance. The images are synthetic or augmented rather than a verified patient dataset, and the project's benchmark file and script contain inconsistent training-epoch information. The reported metrics should be confirmed by rerunning the evaluation with a fixed protocol, documented data split, and class-wise analysis before being presented as final research findings. The current model also lacks a normal or benign category, so it cannot determine that an image is healthy or rule out disease. Overall, the project demonstrates the integration of a web interface, FastAPI backend, image-processing pipeline, CNN–LSTM classifier, reference matching, results page, and browser-based prediction history; it has not been clinically validated and must not be used for diagnosis or treatment.

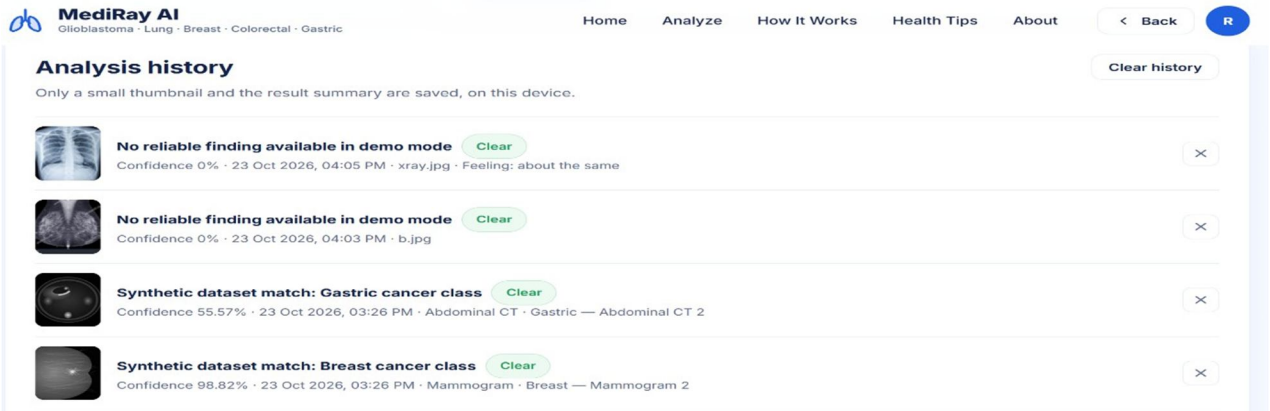


Fig.8. Analysis History

VIII. LIMITATIONS

MediRay AI is a research prototype, and its performance is limited by the available data and evaluation. The saved benchmark uses synthetic or augmented images rather than a verified, diverse collection of patient scans, so its metrics do not show how the model would perform across hospitals, scanners, or patient populations. The benchmark files also contain inconsistent training-epoch information; the scores should be confirmed through a reproducible evaluation before being treated as final. The classifier supports a fixed set of categories and lacks a normal or benign class, so it cannot confirm that an image is healthy or rule out disease. Reference matching can flag weak or conflicting predictions, but it cannot guarantee correctness.

The model has not been clinically validated and should not be used for diagnosis or treatment. Prediction history is saved in the user's browser rather than a central database, so it may not transfer to another device and may be lost if browser data is cleared. The system's predictions and other displayed content are intended for educational and research demonstration only.

IX. FUTURE ENHANCEMENT

Future work can improve MediRay AI by training and evaluating it with larger, verified, patient-derived datasets that cover diverse scanners, image qualities, and populations. The dataset should include normal and benign cases alongside disease categories. A fixed training and evaluation protocol, independent test sets, class-wise performance metrics, and external validation would help produce more reliable and reproducible results. Model calibration and uncertainty estimates could also help the system communicate when a prediction is uncertain.

The application could be extended with secure server-side accounts and encrypted prediction history so users can access records across devices. Further enhancements may include clinician-reviewed explanations, improved image-quality checks, accessibility and multilingual support, and stronger privacy controls. Any clinical use would require appropriate medical oversight, rigorous validation, and applicable regulatory review before deployment.

X. CONCLUSION

MediRay AI demonstrates a web-based workflow for medical image classification, connecting image upload and validation with a CNN–LSTM model, reference matching, results display, and browser-based prediction history. The system illustrates how these components can be integrated into an accessible prototype. Its reference-matching step can flag weak or conflicting predictions and return an unclassified result when the available evidence is insufficient. The project's saved benchmark reports higher scores for CNN–LSTM than for the baseline CNN on a 5000-image synthetic or augmented test set. These results provide an initial software-level comparison, but they do not establish clinical accuracy. The data is not a verified patient dataset, benchmark metadata requires confirmation, and the model lacks a normal or benign category. Consequently, the system cannot confirm that an image is healthy, rule out disease, or provide a diagnosis.

Overall, the project provides a foundation for further research and development in automated image-analysis workflows. Future work should prioritise verified and diverse medical datasets, reproducible evaluation, external validation, calibrated uncertainty, improved privacy protections, and clinician oversight. Until these steps are completed and the system is appropriately validated, MediRay AI should remain an educational and research prototype rather than a tool for clinical decision-making.

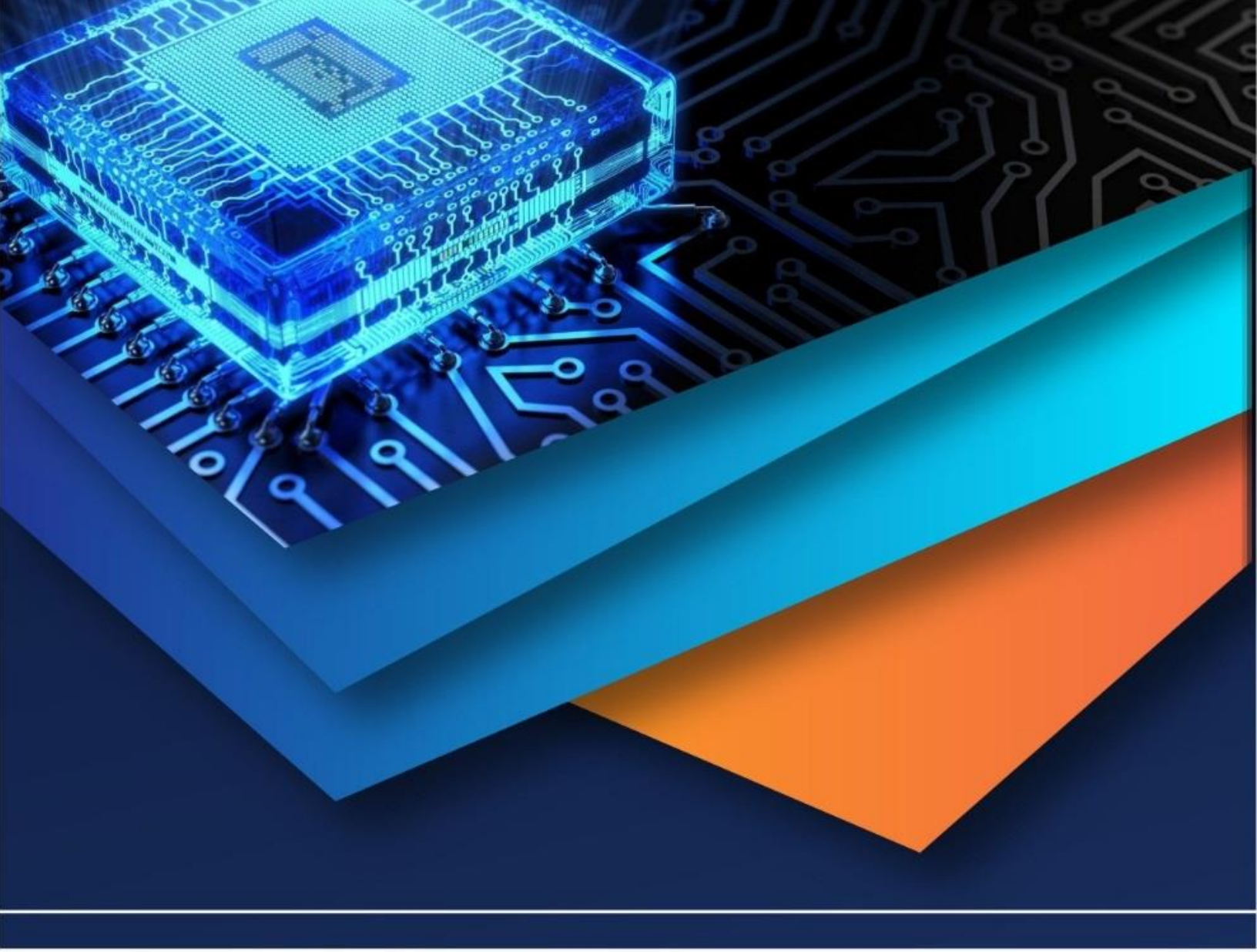
REFERENCES

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105, 2012.
[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," International Conference on Learning Representations, 2015.
[4] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, 2016.
[6] G. Litjens et al., "A Survey on Deep Learning in Medical Image Analysis," Medical Image Analysis, vol. 42, pp.60–88, 2017.
[7] A. Esteva et al., "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks," Nature, vol.542, pp. 115–118, 2017.
[8] D. S. Kermany et al., "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," Cell, vol. 172, no. 5, pp. 1122–1131.e9, 2018.
[9] X. Wang et al., "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localisation of Common Thorax Diseases," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3462–3471, 2017.
[10] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," arXiv preprint, arXiv:1711.05225, 2017.
[11] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer-

©IJRASET: All Rights are Reserved

Assisted Intervention, vol. 9351, pp. 234–241, 2015.
[12] A. Paszke et al., “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” Advances in Neural Information Processing Systems, vol. 32, pp. 8024–8035, 2019.
[13] M. Sokolova and G. Lapalme, “A Systematic Analysis of Performance Measures for Classification Tasks,” Information Processing & Management, vol. 45, no. 4, pp. 427–437, 2009.
[14] C. Shorten and T. M. Khoshgoftaar, “A Survey on Image Data Augmentation for Deep Learning,” Journal of Big Data, vol. 6, article 60, 2019.
[15] C. J. Kelly et al., “Key Challenges for Delivering Clinical Impact with Artificial Intelligence,” BMC Medicine, vol. 17, article 195, 2019.
[16] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” International Journal of Computer Vision, vol. 115, pp. 211–252, 2015.
[17] J. Deng et al., “ImageNet: A Large-Scale Hierarchical Image Database,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255, 2009.
[18] R. R. Selvaraju et al., “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localisation,” Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626, 2017.
[19] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, “On Calibration of Modern Neural Networks,” Proceedings of the 34th International Conference on Machine Learning, vol. 70, pp. 1321–1330, 2017.
[20] J. R. Zech et al., “Variable Generalisation Performance of a Deep Learning Model to Detect Pneumonia in Chest Radiographs: A Cross-Sectional Study,” PLOS Medicine, vol. 15, no. 11, article e1002683, 2018.
[21] S. M. McKinney et al., “International Evaluation of an AI System for Breast Cancer Screening,” Nature, vol.
[22] 577, pp. 89–94, 2020.
[23] B. Ehteshami Bejnordi et al., “Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer,” JAMA, vol. 318, no. 22, pp. 2199–2210, 2017.
[24] N. Coudray et al., “Classification and Mutation Prediction from Non-Small Cell Lung Cancer Histopathology Images Using Deep Learning,” Nature Medicine, vol. 24, pp. 1559–1567, 2018.
[25] G. Campanella et al., “Clinical-Grade Computational Pathology Using Weakly Supervised Deep Learning on Whole Slide Images,” Nature Medicine, vol. 25, pp. 1301–1309, 2019.
[26] D. D. Mongan, L. Moy, and C. E. Kahn Jr., “Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers,” Radiology: Artificial Intelligence, vol. 2, no. 2, article e200029, 2020.
[27] H. J. W. L. Aerts et al., “Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach,” Nature Communications, vol. 5, article 4006, 2014.
[28] R. T. H. Leijenaar et al., “Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis,” European Journal of Cancer, vol. 48, no. 4, pp. 441–446, 2012.
[29] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How Transferable Are Features in Deep Neural Networks?” Advances in Neural Information Processing Systems, vol. 27, pp. 3320–3328, 2014.
[30] K. Clark et al., “The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository,” Journal of Digital Imaging, vol. 26, pp. 1045–1057, 2013.
[31] S. G. Armato III et al., “The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans,” Medical Physics, vol. 38, no. 2, pp. 915–931, 2011.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)