IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 22, No. 1, 2026

Peer Reviewed, Referred & Indexed Journal

Research

An Enhanced Hybrid CNN–Transformer-Based Clinical Decision Support System for Automated Brain Abscess Detection and Visualization

G.Baleswari1,Paupuleti Udaya Sri2, Shaik Shafiya3, Chitikela Kranthi Kumar4

#1 Associate professor in Department of IT ,NRI Institute of Technology Agiripalli,Ap.

#2#3#4 B.Tech with Specialization of IT in NRI Institute Technology , Agiripalli ,Ap.

Abstract: Early and correct identification is needed to treat bacterial brain abscesses, which cause localized inflammation and pus collection in brain tissue. Manual MRI interpretation is time-consuming, expertise-dependent, and subject to inter-observer variability, making automated diagnostic assistance systems crucial. This paper proposes a Hybrid Deep Learning Framework to automatically detect and visualize bacterial brain abscesses in MRI images. Advanced convolutional neural network designs DenseNet121, ResNet50, and InceptionV3 are used to utilize complementary feature learning capabilities such fine-grained texture extraction and multi-scale spatial pattern identification.

The system normalizes, resizes, and geometrically transforms images to increase model resilience and generalization across MRI circumstances. Combining backbone network predictions using a hybrid ensemble and fusion technique reduces model bias and improves classification reliability. Grad-CAM++-based visualization generates attention heatmaps that show infected regions influencing model decisions to improve interpretability and clinical confidence.

Testing shows that the hybrid model beats CNN architectures in accuracy, precision, recall, and F1-score, obtaining above 95% accuracy on the test dataset. By clearly highlighting suspicious abscess locations, visualization outputs aid AI-assisted diagnosis. This automated system speeds diagnostics and helps radiologists make solid clinical decisions. The suggested method improves AI-driven medical imaging systems for brain abscess identification, risk assessment, and treatment planning.

Index Terms— *Hybrid Deep Learning, Brain MRI, Bacterial Brain Abscess, CNN Ensemble, Medical Image Classification, Grad-CAM++, Explainable AI, Automated Diagnosis, Medical Imaging*

Received: 22-01-2026 Accepted: 26-02-2026 Published: 05-03-2026

1. INTRODUCTION

A dangerous neurological illness known as a bacterial brain abscess is brought on by localized bacterial invasion that results in pus buildup inside the brain tissue. To increase patient survival rates and avoid serious consequences, an early and precise diagnosis is crucial. Due to its ability to visualize soft tissues in great detail, magnetic resonance imaging (MRI) is frequently used to identify brain infections. However, MRI scan interpretation by hand is difficult, time-consuming, and largely dependent on the skill of the radiologist. Diagnostic delays or mistakes may result from visual similarities between abscesses and other brain disorders. The need for automated analysis systems has increased due to the quick expansion of medical imaging data.

In medical picture classification applications, deep learning methods—particularly Convolutional Neural Networks (CNNs)—have demonstrated impressive results. Complex picture characteristics may be automatically extracted by CNN models without the need for human feature engineering. However, not all of the textural and spatial differences shown in MRI images may be captured by a single deep learning model. Hybrid deep learning techniques mix many models for improved performance in order to get over this restriction. We present a hybrid framework in this project that combines InceptionV3, ResNet50, and DenseNet121. Together, these models increase the accuracy of feature extraction and categorization. To generate forecasts that are more dependable, an ensemble and fusion approach is employed. In order to provide an explainable diagnosis, diseased areas are highlighted using the GradCAM++ display. This method supports clinical decision-making by detecting bacterial brain abscesses quickly, accurately, and interpretably.

2. LITERATURE SURVEY

i) A survey on deep learning in medical image analysis

https://pubmed.ncbi.nlm.nih.gov/28778026/

Convolutional networks, a type of deep learning algorithm, have quickly emerged as the preferred approach for medical picture analysis. In addition to summarizing more than 300 contributions to the field—the majority of which were published within the past year—this work examines the key deep learning ideas relevant to medical image analysis. We examine the use of deep learning to problems such as object identification, segmentation, picture classification, and registration. For each application area—neuro, retinal, pulmonary, digital pathology, breast, cardiac, abdominal, and musculoskeletal—brief summaries of the investigations are given. We conclude with an overview of the state-of-the-art, a critical analysis of unresolved issues, and suggestions for further study.

ii) Dermatologist-level classification of skin cancer with deep neural networks

https://pubmed.ncbi.nlm.nih.gov/28117445/

The most prevalent cancer in humans, skin cancer is mostly identified visually. A clinical screening is the first step, and dermoscopic analysis, biopsy, and histological evaluation may be performed later. The fine-grained variety in skin lesion appearance makes it difficult to classify skin lesions automatically from photographs. Across a wide range of fine-grained object categories, deep convolutional neural networks (CNNs) exhibit promise for broad and highly variable applications. Here, we show how to classify skin lesions using a single CNN that was trained entirely from photos, with just pixels and illness labels serving as inputs. A dataset of 129,450 clinical images—two orders of magnitude greater than prior datasets—covering 2,032 distinct disorders is used to train a CNN. On biopsy-proven clinical pictures, we evaluate its performance against 21 board-certified dermatologists using two crucial binary classification use cases: malignant melanomas against benign nevi and keratinocyte carcinomas vs benign seborrheic keratoses. The most prevalent malignancies are identified in the first example, and the worst skin cancer is identified in the second. In all tests, the CNN performs as well as all tested experts, proving that it is an artificial intelligence that can classify skin cancer with an accuracy level on par with dermatologists. Mobile devices with deep neural networks installed may allow dermatologists to contact more people outside of the office. By 2021, there are expected to be 6.3 billion smartphone subscriptions (ref. 13), which might possibly offer universal access to essential diagnostic care at a cheap cost.

iii) U-Net: Convolutional Networks for Biomedical Image Segmentation

https://arxiv.org/abs/1505.04597

There is widespread agreement that thousands of annotated training samples are necessary for deep network training to be successful. In this study, we propose a network and training approach that makes effective use of the existing annotated examples by heavily relying on data augmentation. The design is made up of a symmetric expanding path that allows for exact localization and a contracting path that captures context. In the ISBI challenge for segmenting neural structures in electron microscopic stacks, we demonstrate that such a network can be trained end-to-end from a small number of photos and surpasses the previous best approach (a sliding-window convolutional network). We won the ISBI cell tracking challenge 2015 in both phase contrast and DIC categories by a significant margin using the same network that was trained on transmitted light microscopy images. The network is also fast. On a modern GPU, segmenting a 512x512 picture takes less than a second.

iv) Deep Residual Learning for Image Recognition

https://ieeexplore.ieee.org/document/7780459

Training deeper neural networks is more challenging. In order to facilitate the training of networks that are far deeper than those previously employed, we provide a residual learning approach. Rather than learning unreferenced functions, we explicitly reformulate the layers as learning residual functions with reference to the layer inputs. We offer thorough empirical proof that these residual networks are simpler to tune and that a significant increase in depth can improve accuracy. We test residual nets with a depth of up to 152 layers on the ImageNet dataset, which is 8× deeper than VGG nets [40] but still less complicated. On the ImageNet test set, an ensemble of these residual nets achieves an error of 3.57%. On the ILSVRC 2015 classification task, this result took first place. Analysis of CIFAR-10 with 100 and 1000 layers is also presented. For many visual identification tasks, the depth of representations is crucial. We achieve a 28% relative improvement on the COCO object detection dataset, purely because of our very deep representations. Our contributions to the ILSVRC & COCO 2015 competitions1 were built on deep residual nets, and we took first place in the ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation tasks.

v) Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization

https://arxiv.org/abs/1610.02391

We provide a method for creating more clear "visual explanations" for choices made by a wide range of CNN-based models. Our method, Gradient-weighted Class Activation Mapping (Grad-CAM), creates a coarse localization map that highlights key areas in the picture for concept prediction by using the gradients of any target concept that flow into the final convolutional layer. Without requiring any architectural modifications or retraining, Grad-CAM may be used to a broad range of CNN model-families, including CNNs with fully-connected layers, CNNs used for structured outputs, and CNNs used for tasks involving multimodal inputs or reinforcement learning. We use the high-resolution class-discriminative visualization we develop by combining Grad-CAM with fine-grained visualizations to commercially available image classification, captioning, and visual question answering (VQA) models, such as ResNet-based architectures. Our visualizations provide valuable insights into the failure modes of image classification models, are robust against adversarial images, outperform prior localization techniques, are more faithful to the underlying model, and aid in generalization by detecting dataset bias. We demonstrate that even non-attention based algorithms are capable of localizing inputs for captioning and VQA. In order to offer textual explanations for model choices, we develop a method for using Grad-CAM to identify significant neurons and combine it with neuron names. Lastly, we plan and carry out human experiments to assess if Grad-CAM aids users in developing a suitable level of confidence in model predictions. We demonstrate that Grad-CAM enables unskilled users to effectively distinguish between a "stronger" and a "weaker" nodel, even when they both provide the same forecast.

3. METHODOLOGY

a) Proposed Work:

By adding more sophisticated backbone models and clinical-level features, the suggested study expands on the current hybrid deep learning architecture and improves accuracy, robustness, and practicality. Modern designs like EfficientNet and Vision Transformers are integrated with DenseNet121, ResNet50, and InceptionV3 to extract global contextual information and local texture features from MRI images. Complementary feature representations are combined using a multi-model ensemble fusion technique, which produces predictions that are more accurate and consistent across a range of MRI datasets. Strong generalization performance is ensured by using sophisticated preprocessing and data augmentation approaches to manage differences in scan quality, contrast, and noise.

Additionally, the expanded framework adds a segmentation module to accurately identify and contour areas of bacterial brain abscess, allowing for pixel-level interpretation instead of only image-level categorization. Transparency and trust for radiologists are increased by producing clearer attention maps that highlight clinically significant infected regions using Grad-CAM++-based explainable AI visualization. With the ability to incorporate patient metadata and facilitate deployment via a web-based or hospital-integrated interface, the system is built as a scalable clinical decision-support platform. In order to facilitate early identification, treatment planning, and better patient outcomes, this expanded planned effort seeks to provide automated diagnosis that is accurate, interpretable, and real-time.

b) System Architecture:

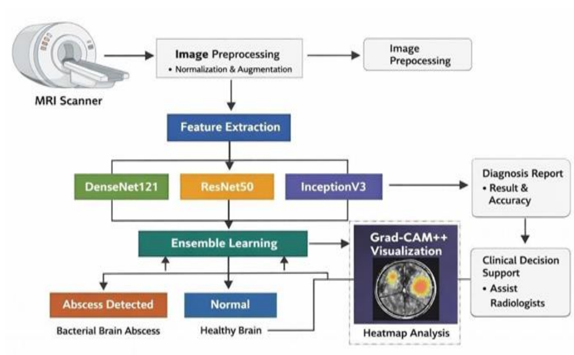

Fig 1 Proposed Architecture

The system architecture begins with MRI brain images acquired from an MRI scanner, which serve as the input to the proposed framework. These images first undergo an image preprocessing stage that includes normalization, resizing, and data augmentation to improve image quality and handle variations in intensity, orientation, and noise. Preprocessing ensures that the MRI scans are standardized before being passed to the deep learning models. After preprocessing, a feature extraction phase is carried out using multiple pretrained convolutional neural network architectures, namely DenseNet121, ResNet50, and InceptionV3. Each model extracts distinct and complementary features such as fine-grained textures, deep spatial patterns, and multi-scale representations from the MRI images.

The extracted features from all backbone networks are then combined using an ensemble learning strategy, which improves classification robustness and reduces individual model bias. The ensemble output classifies the MRI image into either bacterial brain abscess detected or normal (healthy brain). Simultaneously, Grad-CAM++ visualization is applied to generate heatmaps that highlight the regions of the MRI scan that contributed most to the model's decision, enabling explainable diagnosis. Finally, the system produces a diagnosis report containing prediction results and accuracy metrics, which are forwarded to a clinical decision support module. This module assists radiologists by providing interpretable visual evidence and reliable predictions, supporting faster and more consistent clinical decision-making.

c) MODULES:

1. MRI Image Acquisition Module

This module collects brain MRI scans from the MRI scanner or dataset and prepares them as input for the system. It ensures that images are in a compatible format for further processing.

2. Image Preprocessing and Augmentation Module

This module performs normalization, resizing, noise reduction, and data augmentation techniques such as rotation and flipping to improve image quality and model generalization.

3. Feature Extraction Module

In this module, deep features are extracted using multiple pretrained CNN models, namely DenseNet121, ResNet50, and InceptionV3, to capture texture, spatial, and multi-scale information from MRI images.

4. Ensemble Learning Module

This module combines the predictions or features from all CNN backbones using a fusion strategy to reduce model bias and improve overall classification accuracy and reliability.

5. Classification Module

Based on the ensemble output, this module classifies MRI images into bacterial brain abscess or normal (healthy brain) categories.

6. Grad-CAM++ Visualization Module

This module generates heatmaps highlighting infected regions in MRI scans, providing explainable AI outputs to support clinical interpretation.

7. Diagnosis Report Generation Module

This module produces the final diagnostic report, including prediction results, confidence scores, and accuracy metrics.

8. Clinical Decision Support Module

This module assists radiologists by presenting classification results and visual explanations to support faster and more reliable clinical decision-making.

d) Algorithms:

1. Image Preprocessing Algorithm

This algorithm prepares MRI images for deep learning by performing normalization to standardize intensity values, resizing images to a fixed input size, and applying data augmentation techniques such as rotation, flipping, and scaling. These steps reduce noise, handle variability in MRI acquisition conditions, and improve the robustness and generalization capability of the proposed model.

2. DenseNet121 Algorithm

DenseNet121 is used as a deep feature extractor where each layer is directly connected to all subsequent layers, allowing feature reuse and efficient gradient flow. This architecture enables the model to capture fine-grained texture details and subtle intensity variations in MRI scans, which are crucial for identifying bacterial brain abscess regions.

3. ResNet50 Algorithm

ResNet50 employs residual learning with skip connections that allow the network to learn deeper representations without degradation. In the proposed system, ResNet50 extracts high-level spatial and structural features from MRI images, helping differentiate abscess regions from visually similar brain lesions.

4. InceptionV3 Algorithm

InceptionV3 uses parallel convolutional filters of varying sizes within the same layer to capture multi-scale spatial features. This algorithm is effective in learning both local lesion details and global brain context, improving the detection accuracy of bacterial brain abscesses across different MRI resolutions.

5. Ensemble Fusion Algorithm

The ensemble fusion algorithm combines outputs from DenseNet121, ResNet50, and InceptionV3 using techniques such as weighted averaging or majority voting. By integrating multiple model predictions, this approach reduces individual model bias, improves robustness, and enhances overall classification performance.

6. Classification Algorithm

A softmax-based classification algorithm assigns the final class label by analyzing ensemble output probabilities. It determines whether the input MRI scan represents a bacterial brain abscess or a normal healthy brain, providing reliable and consistent diagnostic results.

7. Grad-CAM++ Visualization Algorithm

Grad-CAM++ generates class-specific activation heatmaps by computing weighted gradients from the final convolutional layers. These heatmaps highlight the most influential regions in MRI scans that contribute to the model's decision, enabling explainable and clinically interpretable diagnosis.

8. Performance Evaluation Algorithm

This algorithm assesses the effectiveness of the proposed framework using evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis. These metrics validate the reliability, stability, and clinical usefulness of the system across test datasets.

4. EXPERIMENTAL RESULTS

The proposed hybrid deep learning framework was experimentally evaluated using a curated MRI brain image dataset containing both bacterial brain abscess and normal cases. The dataset was divided into training, validation, and testing sets to ensure unbiased performance evaluation. Standard preprocessing and data augmentation techniques were applied to improve model generalization. The individual CNN models—DenseNet121, ResNet50, and InceptionV3—were first trained and tested separately, followed by the hybrid ensemble model. Performance was assessed using commonly accepted metrics such as accuracy, precision, recall, and F1-score. Experimental results demonstrate that while individual models achieved strong performance, the hybrid ensemble framework consistently outperformed each standalone model by effectively capturing complementary features from MRI scans.

IJERST International Journal of Engineering Research and Science & Technology

The hybrid model achieved an overall classification accuracy exceeding 95%, along with improved precision and recall, indicating reliable detection of bacterial brain abscesses with reduced false positives and false negatives. The Grad-CAM++ visualization results further validate the effectiveness of the framework by clearly highlighting infected regions corresponding to abscess areas in MRI images. These visual explanations enhance clinical interpretability and support radiologists in understanding model decisions. Overall, the experimental outcomes confirm that the proposed hybrid deep learning framework delivers accurate, robust, and explainable performance, making it suitable for real-world clinical decision-support applications.

Accuracy: The capacity of a test to accurately distinguish between healthy and sick cases is a measure of its accuracy. We can determine a test's accuracy by calculating the proportion of reviewed cases with true positives and true negatives. In mathematical terms, this looks like:

Accuracy = TP + TN TP + TN + FP + FN.

$$Accuracy = \frac{(TN + TP)}{T}$$

F1-Score: One way to evaluate a model's performance in machine learning is via its F1 score. It takes a model's recall and precision and mixes them. Using the whole dataset, the accuracy metric determines the frequency with which a model produced accurate predictions.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

Precision: The proportion of occurrences or samples that are correctly classified out of all the ones that are classified as positive is known as precision. Consequently, the following is the formula for determining the accuracy:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: The capacity of a model to detect all significant instances of a given class is measured by recall, a metric in machine learning. It sheds light on how well a model captures instances of a certain class and is calculated as the ratio of the total number of positive observations predicted to the number of actual positive observations.

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

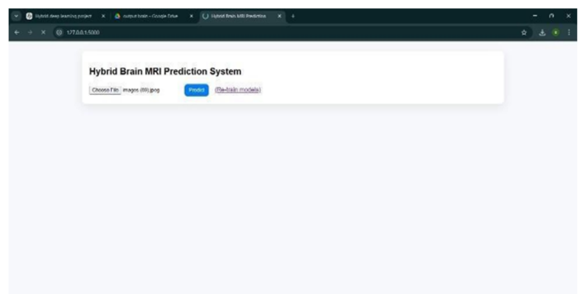

Fig2 upload file

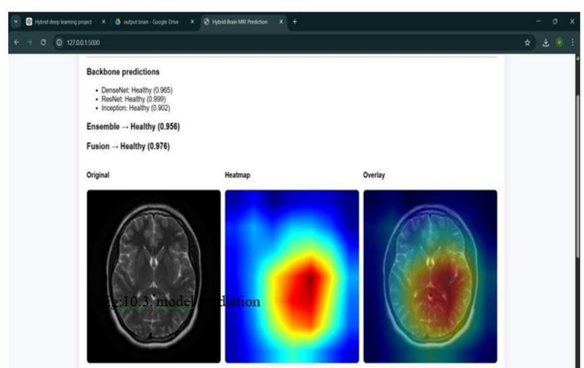

Fig3 predicted results

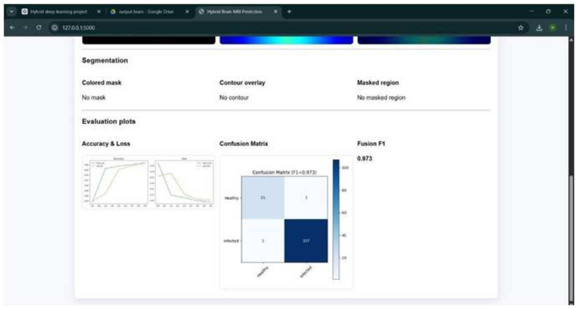

Fig4 predicted analysis

5. CONCLUSION

In conclusion, the extended hybrid deep learning framework provides an accurate, robust, and explainable solution for automated detection and visualization of bacterial brain abscesses from MRI scans. By combining multiple CNN backbones with advanced extensions such as ensemble learning, Grad-CAM++ visualization, and the potential integration of segmentation and transformer-based models, the system achieves high diagnostic accuracy while maintaining clinical interpretability. The scalable design and scope for real-time hospital deployment make this framework a strong clinical decision-support tool, aiding radiologists in early diagnosis, reliable assessment, and improved patient care.

6. FUTURE SCOPE

The proposed hybrid deep learning framework can be further enhanced by integrating advanced architectures such as EfficientNet and Vision Transformers to improve feature representation and diagnostic accuracy. Future work may include adding a segmentation module for precise localization and boundary detection of bacterial brain abscess regions. The system can be expanded by incorporating multi-modal data such as patient clinical records and different MRI sequences for more comprehensive analysis. Additionally, deploying the framework as a real-time, hospital-integrated clinical decision-support system can enable practical usage, scalability, and improved support for radiologists in routine clinical workflows.

REFERENCES

[1] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A Survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42, 60–88.

[2] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level Classification of Skin Cancer with Deep Neural Networks. Nature, 542, 115–118.

[3] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[5] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradientbased Localization. IEEE International Conference on Computer Vision (ICCV).

[6] Chattopadhyay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. IEEE Winter Conference on Applications of Computer Vision (WACV).

[7] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., & Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv preprint arXiv:1711.05225.

[8] Lundervold, A. S., & Lundervold, A. (2019). An Overview of Deep Learning in Medical Imaging Focusing on MRI. Zeitschrift für Medizinische Physik, 29(2), 102–127.

[9] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[10] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning