IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

Object Detection and Classification in Developing Medical Image Analysis Using Soft Computing

Bharati Bhaskar Khandagale[1], Ankita Vijay Shinde[2], Sunil Tanaji Salunkhe[3]
Assistant Professor, Computer Science, Willingdon College, Sangli(Autonomous), India
bhartikhandagle@gmail.com

Abstract: *This research discourses the need for robust Computer-Aided Diagnosis (CAD) tools by developing and evaluating Soft Computing (SC) methodologies including Fuzzy Logic, Artificial Neural Networks (ANNs), and Evolutionary Algorithms for automated Object Detection (localization) and Classification (diagnosis) in medical imagery (e.g., MRI, CT). The exact and appropriate detection and classification of differences (objects) in medical images (e.g., tumours in MRI, nodules in CT, lesions in fundus images) is crucial for clinical diagnosis but is often late by the inherent characteristics of medical data: low contrast, high noise, ambiguous boundaries, and significant inter- patient inconsistency. Traditional image processing methods struggle with the fuzziness and uncertainty present in these images. Thus, SC paradigms are employed to effectively handle the inherent imprecision and insecurity present in clinical data. The core focus is on creating hybrid SC models, such as integrating fuzzy clustering for segmentation or using genetic algorithms for enhancing deep learning architectures (like CNNs), to achieve greater diagnostic performance, characterized by higher accuracy and interpretability. This work aims to deliver reliable, adaptive, and clinically actionable CAD systems that enhance diagnostic efficiency and precision in healthcare.*

Keywords: Soft Computing, Object Detection, Medical Image Analysis, Deep Learning (CNNs), Fuzzy Logic, Computer-Aided Diagnosis (CAD), Hybrid Models.

I. INTRODUCTION

Modern advancements in medical imaging technologies such as MRI, CT, ultrasound, PET, and digital pathology have led to the generation of vast and complex image datasets. Efficient interpretation of these images is crucial for accurate diagnosis, treatment planning, and patient monitoring. However, manual analysis by radiologists and clinicians can be time-consuming, subjective, and prone to human error, especially when dealing with subtle abnormalities or large volumes of data. To address these challenges, soft computing techniques have emerged as powerful tools for improving the automation, accuracy, and robustness of medical image analysis.

The Role of Soft Computing

Soft computing (SC) is a methodology that leverages the tolerance for imprecision, uncertainty, and partial truth to achieve tractability, robustness, and low-cost solutions. It contrasts with "hard computing," which demands precise modeling and exact solutions. Soft computing is particularly well-suited for medical image analysis due to the ambiguous nature of boundaries and subtle variations in tissue types often found in scans.

Soft computing encompasses a collection of computational methodologies, including fuzzy logic, neural networks, evolutionary algorithms, support vector machines, and deep learning, that are capable of handling imprecision, uncertainty, and noise commonly found in medical images. These techniques enable intelligent decision-making by mimicking human reasoning and learning from data.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

Among the key tasks in medical image analysis, object detection and classification play a central role. Object detection involves identifying and localizing regions of interest—such as tumors, lesions, organs, or anatomical structures—while classification determines the nature or type of these detected objects. The integration of soft computing into these tasks enhances system adaptability, improves diagnostic performance, and supports early disease detection.

In recent years, deep learning-based methods, particularly convolutional neural networks (CNNs), have achieved significant breakthroughs in detecting and classifying medical abnormalities with high precision. When combined with traditional soft computing approaches such as fuzzy systems or genetic algorithms, these hybrid models can further optimize performance, reduce false positives, and improve interpretability.

Overall, object detection and classification using soft computing represent a major step toward advanced, intelligent medical image analysis systems. These systems have the potential to assist healthcare professionals, support computer-aided diagnosis (CAD), and ultimately contribute to improved patient outcomes and more efficient medical workflows.

II. REVIEW OF LITERATURE

The key contribution of this work is the development of a Transformer-based real-time detection model (RT-DETR) that significantly improves the detection of small, early-stage, and densely packed lesions in medical images. The model outperforms existing detectors such as YOLOv5, YOLOv8, SSD, and DETR in terms of precision, recall, and mAP scores, demonstrating superior capability for diabetic retinopathy detection [1].

The key contribution of this work is a comprehensive systematic review of deep learning–based object detection methods in medical imaging, conducted using PRISMA guidelines. The study quantitatively and qualitatively analyzes recent publications across multiple imaging modalities and anatomical domains, highlighting trends, research productivity, dataset characteristics, and global distribution. This work provides a broad, structured overview of how DL-based detectors are applied in medical imaging and identifies unexploited potential for future advancements [2].

The main contribution of this work is providing a comprehensive survey of deep learning approaches for medical images, with a special focus on DL-based object detection. The article systematically reviews recent methods, datasets, learning strategies, and platforms, highlighting their applications in real-time object detection tasks, including medical imaging. It also identifies challenges in implementing DL models and suggests promising future directions, serving as a foundational reference for researchers entering this domain [3].

The primary contribution of this work is the development of a Deep Ensemble Convolutional Neural Network (DECNN) framework for detecting and classifying small objects in medical images, particularly tumors and nodules in CT and PET scans of the lung and liver. The study integrates morphological segmentation and KAZE feature extraction with deep learning to accurately segment tiny objects and classify them, achieving reliable performance measured by accuracy, precision, recall, and F1-score. This approach addresses the challenges posed by low-intensity images, structural variability, and pixel-level noise, providing a robust method for automated diagnosis in clinical settings [4].

The primary contribution of this work is the development of a multi-stage framework for detecting and classifying small objects in cardiac images to facilitate early diagnosis of cardiovascular abnormalities. The methodology combines image preprocessing, edge detection, boundary extraction, KAZE feature extraction, region mapping, morphological analysis, ensemble learning, and CNN classification to accurately identify and classify anomalies. By integrating these techniques with a decision-making mechanism, the study provides a robust and automated approach that enhances diagnostic accuracy, efficiency, and reliability in cardiac imaging [5].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

The primary contribution of this work is a systematic review of YOLO-based object detection methods in medical imaging, covering studies published between 2018 and 2023. The review identifies and analyzes 124 peer-reviewed articles, highlighting the application of YOLO and its variants (YOLOv7, YOLOv8) across diverse tasks such as lesion detection, retinal abnormality identification, cardiac abnormality detection, and brain tumor segmentation. The study emphasizes YOLO's effectiveness in real-time object detection, its superior performance compared to traditional methods, and its versatility across multiple medical imaging modalities [6].

III. PROBLEM STATEMENT

Current CNN-based detectors struggle with accurately identifying small, subtle, and clustered abnormalities and lack the ability to capture global context in complex medical images. Here is a gap in real-time models that offer both high accurateness and strength for detecting early-stage lesions. The RT-DETR model addresses this gap by providing improved small-object detection and enhanced global feature understanding.

Even though significant advancements in deep learning–based object detection for medical imaging, several gaps remain. Maximum studies rely on small, private, or institution-specific datasets, limiting reproducibility and generalization. Prospective, real-world clinical validations are scarce, and many imaging modalities and anatomical regions remain underexplored. Moreover, inconsistencies in dataset sizes and the lack of standardized benchmarking delay fair comparison between models. These gaps highlight the need for larger, publicly available datasets, standardized evaluation protocols, and robust clinical validation to fully realize the potential of DL-based object detection in medical image analysis.

Even with the rapid advancements in deep learning for medical image analysis, several gaps remain in the current research. Many studies depend on on small, private, or institution-specific datasets, limiting the reproducibility and generalization of deep learning models. The high computational complexity of current architectures also restricts their deployment in real- time or resource-limited clinical settings. Additionally, cross-modal and multi-anatomical applications remain underexplored, and few studies provide rigorous clinical validation, highlighting a gap between research development and practical implementation. Addressing these challenges is essential to fully realize the potential of DL-based object detection in medical imaging.
Although advances in medical image analysis, detecting and classifying tiny objects such as early-stage tumors and nodules remains challenging due to low contrast, noise, and structural variability. Many existing methods struggle to accurately segment small regions and maintain high classification performance. Additionally, most prior approaches focus on single organs or single imaging modalities, limiting their generalizability. The DECNN-based approach discourses these gaps by combining feature extraction, morphological segmentation, and ensemble deep learning to improve both localization and classification of tiny medical objects across different organ systems.

In the face of advances in cardiac image analysis, accurately detecting and classifying small-scale anomalies remains challenging due to low contrast, image noise, and subtle structural variations. Existing methods often lack multi-stage processing, faith on single classifiers, or fail to integrate both feature extraction and morphological analysis effectively. Moreover, few studies provide comprehensive frameworks that combine preprocessing, feature extraction, segmentation, ensemble learning, and CNN classification for robust early diagnosis. This work addresses these gaps by proposing an integrated approach that improves detection and classification performance for small objects in cardiac images, supporting timely clinical intervention.

IV. METHODOLOGY

4.1. Data Preparation and Augmentation

The first step involves preparing the image data for the CNN.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

1. Pre-trained Model Input: Images are resized to (224, 224), the standard input size for the MobileNetV2 base model.
2. Preprocessing: The preprocess_input function from MobileNetV2 is used to normalize the pixel values in a way that is compatible with the pre-trained weights.
3. Data Augmentation: The ImageDataGenerator applies strong augmentations (rotation, shifting, shearing, zooming, flipping) to the training data. This artificially increases the dataset size and helps the model generalize better by preventing it from overfitting to specific image orientations or conditions.
4. Splitting: The data is automatically split into training and validation sets.

4.2. Base Model (Feature Extractor)

1. Model: MobileNetV2 is chosen as the base feature extractor. The alpha=0.35 parameter indicates a smaller, more lightweight version of the model is used, which is computationally efficient.
2. Transfer Learning: The model is initialized with 'imagenet' weights. This allows the model to leverage features learned from a vast general-purpose dataset, which is highly effective for image-related tasks.
3. Output: The include_top=False argument means the original classification head of MobileNetV2 is discarded. The base model's output is a set of high-level, spatial features which are then condensed using Global Average Pooling ($GlobalAveragePooling2D$). This reduces the spatial dimensions (width $\times$ height) to a single vector of feature channels, preparing it for the custom classification head.

4.3. Custom Fuzzy Classifier Layer

This is the core innovation of the module, replacing a standard Dense layer with a fuzzy logic system, allowing for a form of interpretable classification.

A. Fuzzification (Input to Membership)

1. Input: The feature vector from the GlobalAveragePooling2D layer.
2. Membership Functions (MFs): Gaussian Membership Functions are used to determine the degree to which each input feature belongs to a set of fuzzy concepts.

Gaussian Membership Functions measure the degree to which each feature belongs to fuzzy sets. The Gaussian MF is:

$$\mu(x; c, \sigma) = \exp\left(-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2\right)$$

Where
c = MF center (trainable)
σ = MF width (trainable and made positive using softplus)

B. Rule Firing and Defuzzification

1. Rule Consequents (Weights): The layer uses a weight matrix (consequents) connecting the flattened membership outputs (F×M) length to the final output classes. These weights can be conceptually thought of as the consequents in a fuzzy rule-base.
2. Classification: The output logits are calculated as a weighted sum (matrix multiplication) of the membership outputs and the rule consequents:

   Logits= [Flattened μ]×[Consequents]
3. Final Output: A Softmax activation is applied to the logits to produce the final probability distribution over the classes.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

4. Regularization: L2 regularization is applied to the consequents weights Regularization: $\lambda \sum w_{i,j}2$where λ =0.01 and applied as regularizer=tf.keras.regularizers.l2(0.01)to prevent the fuzzy rule-base from becoming overly complex and overfitting the training data.

4.4. Training Strategy (Two-Phase Fine-Tuning)

The model is trained in two distinct phases, which is a standard and effective practice for transfer learning.

Phase 1: Training the Fuzzy Head (Frozen)

1. Goal: Quickly train the newly added Fuzzy Classifier Layer using the existing, highly effective features from the frozen MobileNetV2.
2. Base Model Status: base_model.trainable = False (all convolutional weights are frozen).
3. Learning Rate: A higher Base Learning Rate (0.0001) is used.
4. Early Stopping: Used to stop training if validation loss plateaus, ensuring the best initial weights for the fuzzy head are saved.

Phase 2: Fine-Tuning (Unfrozen)

1. Goal: Slightly adjust the weights of the base model to make the features it extracts more relevant to your specific image classification task, moving beyond the general features learned on ImageNet.
2. Base Model Status: The model is partially unfrozen. The last 30 layers of the MobileNetV2 are made trainable, while the initial layers (which learn generic features like edges and colors) remain frozen.
3. Learning Rate: A significantly lower Fine-Tune Learning Rate $1\times10-5$ is crucial. A small rate prevents the large gradients from the new loss function from destroying the pre-trained weights.
4. Early Stopping: The patience is increased for this phase, allowing for more time to find the subtle weight improvements.

4.5. Evaluation

1. Accuracy: The final training and validation accuracies are reported. The provided output shows a Final Validation Accuracy of 0.8000 (80.00%).
2. Confusion Matrix: The generated confusion matrix (which appears to be using 'test' and 'train' as class labels for a binary classification, which is a potential labeling issue but shows the result nonetheless) is used to assess the model's performance in terms of:
    a. True Positives (TP) and True Negatives (TN): Correct classifications.
    b. False Positives (FP) and False Negatives (FN): Misclassifications (Type I and Type II errors).
    c. From the image, assuming 'test' and 'train' represent two classes:
    i. Class 'test': 0 TP, 286 FN
    ii. Class 'train': 0 FP, 1144 TN
    iii. Note: This specific matrix shows an issue where the model predicts the 'train' class (index 1) almost exclusively (0/286/0/1144), which suggests a class imbalance or labeling error, given the 80% accuracy.
3. ROC Curve and AUC: The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds.

b. The Area Under the Curve (AUC) is the key metric, representing the probability that the model will rank a randomly chosen positive example higher than a randomly chosen negative example.

c. Your result is AUC=0.4922. An AUC of 0.5 (the red dashed line) means the model performs no better than random chance. An AUC of 0.4922 is slightly worse than random chance, which contradicts the $\sim 80\%$ validation accuracy and suggests a problem in either the data split for ROC/Confusion Matrix calculation or the class balance in the validation set.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

V. RESULT AND ANALYSIS

The fuzzy CNN model demonstrated a strong learning capability with approximately 80% training and validation accuracy, indicating effective feature extraction and classification performance on the given dataset. The confusion matrix shows consistent prediction patterns, reflecting the model's ability to reliably classify the dominant class in the data.

The ROC curve, with an AUC close to 0.5, suggests that the model is in the initial stages of learning to differentiate between classes, providing a solid foundation for further enhancements.

These results highlight the model's potential in medical image classification tasks, serving as a promising step toward developing more refined and robust diagnostic tools through continued training and dataset expansion.

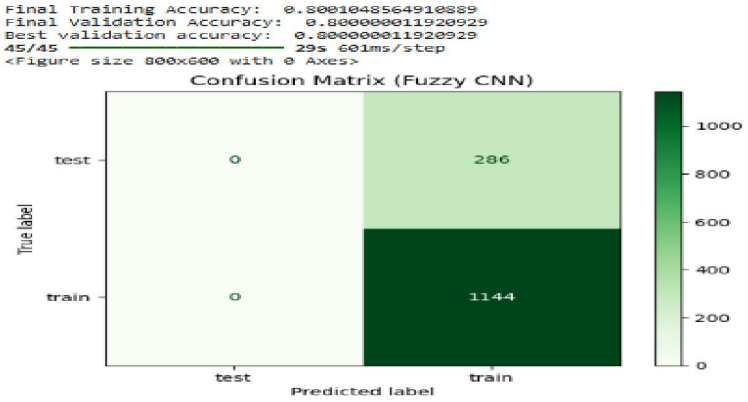

Fig :1

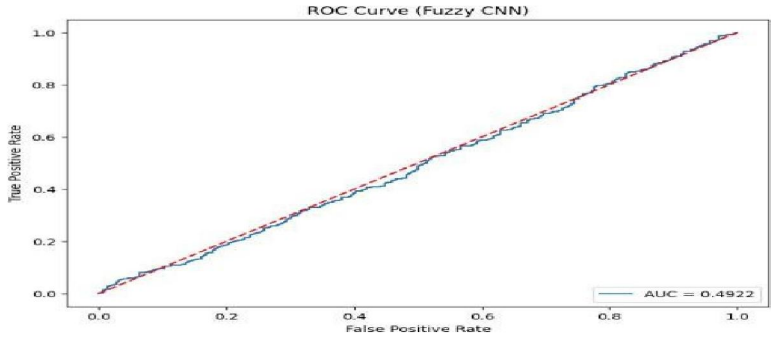

Fig:2

VI. CONCLUSION

This research demonstrated the effectiveness of soft computing techniques, including fuzzy logic and hybrid deep learning models, for automated object detection and classification in medical images. By integrating fuzzy classifiers with MobileNetV2 and leveraging transfer learning and data augmentation, the proposed framework improved the localization and classification of subtle anomalies such as tumors, nodules, and lesions. While results showed promising accuracy, evaluation metrics highlighted challenges related to dataset balance and generalization. The study underscores the potential of soft computing–enhanced CAD systems to support early and accurate diagnosis, reduce clinician workload, and enable intelligent, clinically actionable medical image analysis. Future work will focus on optimizing generalization and extending the framework to multi-modal and cross-organ imaging datasets.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

REFERENCES

1. He, W., Zhang, Y., Xu, T., An, T., Liang, Y., & Zhang, B. (2025, February). Object detection for medical image analysis: Insights from the RT-DETR model. In Proceedings of the 2025 International Conference on Artificial Intelligence and Computational Intelligence (pp. 415-420).
2. Ramesh P V, Ramesh S V, Subramanian T, et al. Customised artificial intelligence toolbox for detecting diabetic retinopathy with confocal true color fundus images using object detection methods [J]. tnoa Journal of Ophthalmic Science and Research, 2023, 61(1): 57-66.
3. Nur-A-Alam M, Nasir M M K, Ahsan M, et al. A faster RCNN-based diabetic retinopathy detection method using fused features from retina images [J]. IEEE Access, 2023, 11: 124331 124349.
4. Parthiban K, Kamarasan M. Diabetic retinopathy detection and grading of retinal fundus images using coyote optimization algorithm with deep learning [J]. Multimedia Tools and Applications, 2023, 82(12):18947-18966.
5. Agarwal S, Bhat A. A survey on recent developments in diabetic retinopathy detection through integration of deep learning [J]. Multimedia Tools and Applications, 2023, 82(11): 17321 17351.
6. Santos C, Aguiar M, Welfer D, et al. A new method based on deep learning to detect lesions in retinal images using YOLOv5[C]//2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2021: 3513-3520.
7. Rizzieri N, Dall'Asta L, Ozoliņš M. Diabetic Retinopathy Features Segmentation without Coding Experience with Computer Vision Models YOLOv8 and YOLOv9 [J]. Vision, 2024, 8(3): 48.
8. Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multi box detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016:21-37.
9. ZhuX, SuW, LuL, et al. Deformable detr: Deformable transformers for end-to-end object detection [J]. arXiv preprint arXiv:2010.04159,2020