INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Region-Focused Patch-Based Mixup for Improving Classification in Low-Resource Medical Imaging

A. Sahaya Mercy[1]*, **Dr. G. Arockia Sahaya Sheela**[2]

1 Research Scholar, Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli-2. Affiliated to Bharathidasan University, Tamil Nadu, India
2 Assistant Professor, Department of Computer Science, St. Joseph's College (Autonomous), Tiruchirappalli-2, Affiliated to Bharathidasan University, Tamil Nadu, India

OPEN ACCESS

Received: 25/07/2025
Accepted: 25/08/2025
Published: 19/09/2025

Citation: Sahaya Mercy A, Arockia Sahaya Sheela DG (2025) Region-Focused Patch-Based Mixup for Improving Classification in Low-Resource Medical Imaging. Indian Journal of Science and Technology 18(34): 2811-2820. https://doi.org/10.17485/IJST/v18i34.1313

* **Corresponding author**.

smercy.amal@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objective: This study aims to enhance performance of classification accuracy in grayscale medical images by addressing the limitations posed by small and imbalanced datasets. The focus is to develop a simple yet effective data augmentation technique that improves model generalization without introducing dynamic or adaptive complexities. Method: A novel augmentation technique, Region-Guided Mixup Augmentation (RGM-Aug), is proposed. The method involves blending a fixed-size patch from one image into another image of the same class at a predetermined location, introducing localized feature diversity while preserving semantic integrity. The augmented dataset trains the lightweight MobileNetV3-Small model. Performance is evaluated based on accuracy, Micro-F1, Macro-F1 scores, and computational efficiency. Findings: The application of RGM-Aug resulted in substantial improvements across all evaluated models. The proposed model achieved 98.63% accuracy. Comparative analysis with existing methods demonstrated superior performance in classification metrics and efficiency. The augmentation technique also led to faster model convergence and reduced validation loss, indicating better generalization. Novelty: Unlike existing augmentation strategies that rely on global transformations or adaptive mechanisms, RGM-Aug introduces a deterministic, region-focused approach that enriches intra-class variability with minimal computational overhead. This method is particularly effective for small-scale medical imaging datasets and is designed to be architecture-agnostic, making it adaptable to various deep learning models without additional complexity.

Keywords: Liver Steatosis; Ultrasound Imaging; Region-Guided Mixup Augmentation; Deep Learning; Data Augmentation

1 Introduction

Non-alcoholic Fatty liver disease (NAFLD) has become a significant health care problem worldwide, potentially developing rapidly and without symptoms to develop even greater liver disease leading to steatohepatitis, fibrosis, and liver cirrhosis [(1)]. There is

still a chaotic system of NAFLD diagnostics: current clinical practice is based on the intensive use of invasive methods of obtaining liver tissue, expensive and inconvenient imaging techniques and interpretation of ultrasound results[2]. Though non-invasive examinations like vibration-controlled transient elastography (FibroScan) and CT/MRI testing of liver have demonstrated superiority in staging and fat quantification, they are not practical enough due the lack of accessibility and scale and ramping up in non-professional or less-resourced healthcare settings[3].

Some of the recent breakthroughs in deep learning have demonstrated a strong potential in automating the interpretation of medical images, as well as extracting clinically relevant data in both visual and structured health data[4]. Nevertheless, most of the studies that have been conducted so far have concentrated on the analysis of imaging modalities independently or on a limited, selected dataset of images that cannot be applied to a wide variety of patients. Moreover, relatively little has been done to investigate comparative performance of the various classes of deep learning architectures when trained on real world, multimodal, clinical datasets that contain not only imaging but also metadata.

A consistent pattern can be observed among the four studies even though they are at different stages of development of progressive automation of liver ultrasound i.e., automation of quantification of the liver, automation of segmentation of the liver and automation as real-time decision support. Bozic et al. severely critiques the use of ultrasound methods of hepatic steatosis by following the variability of B-mode, the applicability of CAP, and MRI-PDFF as a benchmark of fat measurements, thereby instigating the idea of a standardized quantitative instrument[5]. By leveraging the recent advances in automation, Wu et al demonstrate comparable AUCs with current CNN-based architecture and practical lesion outline detection using YOLO, making ultrasound viable to be used as in console-aiding tools during surveillance and intervention[6]. Drazinos et al. focus on NAFLD grading, where fine-tuned, pre-trained CNNs on curated hepatorenal windows compare favorably to or even out-perform the hepatorenal index thresholds, further consolidating curated data plus augmentation as important degrees of freedom towards viable grading[7]. Jiawei Wu et al. focus on an upstream precondition: precise and quick liver segmentation by promoting a boundary-attentive network that considers the balance between Dice/IoU and latency at a CPU level, which provides reliable ROIs in downstream processing pipelines detection and grading[8].

The unsupervised deep clustering (UDC) of Gd-EOB-DTPA-enhanced MRI was shown to separate steatosis and non-alcoholic steatohepatitis (NASH) in a novel proof-of-concept study. As compared to traditional imaging parameters such as fat fraction (FF) and relative liver enhancement (RLE), UDC demonstrated very good diagnostic performance in all the histological components of NAFLD[9].

The combination of multilayer perceptron and Random Forests in stacked ensemble models outperformed other combinations based on findings of a study performed on the National Health and Nutrition Examination Survey (NHANES) dataset combined with biomarker data and ultrasound images. Segmentation of liver was done by means of a U-Net model refined with a transformer. The last model achieved an accuracy of 99.8 with an AUC value of 0.997 that indicated the power of hybrid model structures in NAFLD classification[10].

A multiparametric MR-based segmentation-based system with deep learning was able to differentiate focal lesions and healthy liver. Having AUCs of 0.925 and 0.852 on training and test data respectively, the system provided standardization of liver diseases by means of automation of lesion classification[11].

The combination of texture features with deep neural networks in the form of radiomic has been implemented in predicting cirrhosis related to NAFLD. A model trained with gray level and gradient level co-occurrence matrices of MRI images resulted in better diagnostic values against conventional ones[12].

Population-scale screening required big data and the evaluation of machine learning classifiers such as SVM, random forest, and XGBoost on over 14,000 adults. SVM model had the highest accuracy of 0.801, whereas RF had the highest AUROC (0.852), a fact that shows their potential in clinical screening[13].

Identical data augmentation methods are the traditional methods of data expansion which include flipping, rotation and scaling. These techniques lack, however, enough variability of semantics, especially of grayscale ultrasound images. Most recent studies concentrated on the traditional augmentation and segmentation techniques. Random or global hybriding methods may produce unrealistic, class-agnostic samples, particularly with fine-grained differences like grayscale texture.

To overcome these restrictions this study presents a new augmentation methodology, termed Region-Guided Mixup Augmentation (RGM-Aug). Added an advantage to the proposed work it combines the RGM-Aug technique with MobileNetV3-Small deep learning model. The central principle is to increase intra-class diversity by spatially fusing a fixed patch area on a single image into another image of the same type. In contrast to the global mixing strategies, RGM-Aug avoids losing locality and context relevancy and hence creates more realistic and class-coherent synthetic samples. Numerous experiments show that RGM-Aug leads to massive performance advancement during classification than the models trained without augmentation. The presented contribution has its value in low-data environment applications, e.g. liver ultrasound diagnostics, where collecting large-scale diverse sets is impractical.

2 Methodology

2.1 Overview of Proposed RGM-Aug Approach

RGM-Aug is proposed ensuring that ultrasound image datasets are augmented by diversity in representation by performing localized patch-level blending within the same class. In contrast to traditional data augmentation methods which apply global modifications to the image (e.g. flips, noise, rotation), RGM-Aug applies deterministic region-specific transformations that focus on augmenting intra-class variability without degrading the class semantics. Figure 1 shows the entire research pipeline, including model-evaluation, from the acquisition of input images in the data and its processing. The ultrasound images are first divided into the classes of healthy and fatty liver.

At the augmentation stage, RGM-Aug combines two corresponding images of same class to create structurally consistent but diverse samples by blending at patch-level locations using a local approach. This augmented data is then added to this original training set and fed into the deep learning model. The best model is later gauged using normal metrics like accuracy and F1-score. In this augmentation treatment, the lightweight convolutional models acquire robust local textures that are related to liver steatosis.

The augmented samples are clinically valid and allow generalization over small data, which is often a problem in doing medical imaging. The last method involves hard constraints imposed to ensure reproducibility and semantic drift by restricting the patch to augmentation by unique patches within a given class and a fixed region and ratio of blend.

2.2 Process of Region-Guided Mixup Augmentation

Figure 2 demonstrates the internal procedure in the RGM-Aug method which is of particular interest to patch extraction, blending and recombining. The steps are starting with loading and classifying the dataset into the binary classes: Class 0 (health liver) and Class 1 (fatty liver). The image pairs are chosen randomly in every class. The process involved in the augmentation is given below:

2.2.1. Input Ultrasound Image

The procedure starts with the initial B-mode ultrasound examination when the area of concern (ROI) is outlined. This ROI denotes the anatomical area of interest upon which additional investigation and augmentation is to be made.

2.2.2. Patch Extraction

The highlighted ROI is subdivided into fixed size image patches through a sliding window method of window overlap of 50 percent. This hybrid approach achieves improved spatial coverage, keeps regional structural information and even the most minute texture variations necessary to accurate characterization of the region.

2.2.3. Region-Guided Mixup (RGM-Aug)

The extracted patches are fed into the proposed RGM-Aug method wherein similar-class patches across different images are merged on the spatial domain. This region-directed fusion can expand intra-class diversity, as well as contextual relevance, addressing shortcomings of global mixing strategies.

2.2.4. Blended Patch Output

The created augmented patches will have a realistic anatomical consistency after using the RGM-Aug. This method does not impose artifacts or unnatural patterns as alternative augmentation methods because synthetic samples that are created using it may be used to train deep learning models more effectively.

2.2.5. Reconstructed Augmented Image

Finally, the augmented patches are merged to create an entire reconstructed image. This synthesized image is almost close to a natural ultrasound scan but with an improved diversity that gives it the ability to model better in situations it has not seen previously in unseen patient cases.

Let us assume given a set of ultrasound images $\mathcal{D} = \{(I_i, y_i)\}_{i=1}^{N}$, where $I_i \in R^{H \times W}$ represents a grayscale ultrasound image and $y_i \in \{0, 1\}$ denotes its corresponding binary class label (healthy or fatty liver), RGM-Aug operates by selecting a pair of images (I_a, I_b) such that $y_a = y_b$. This intra-class selection ensures that the augmented image maintains the original class semantics.

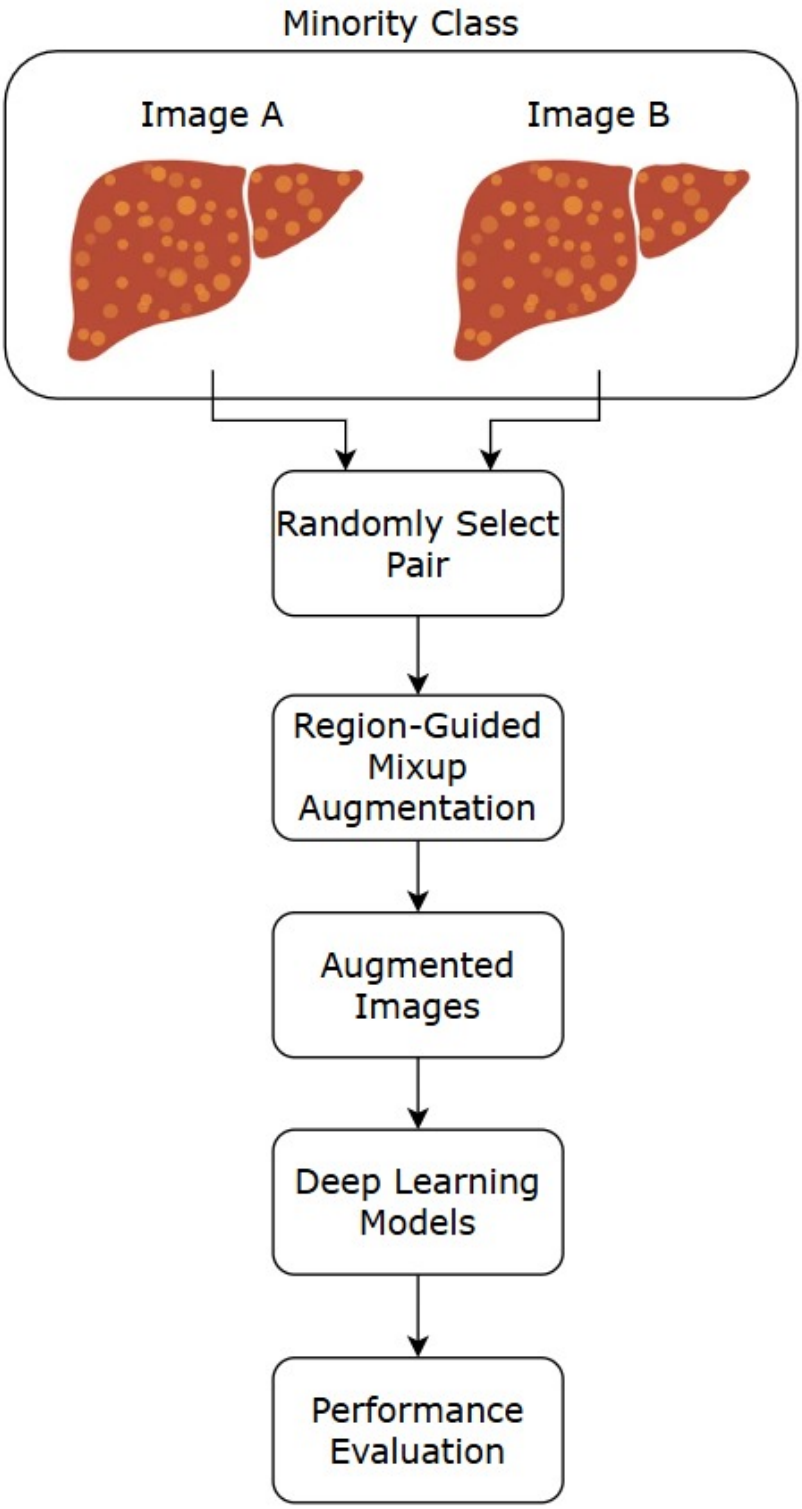

Fig 1. Workflow Diagram of the Proposed RGM-Mixup

A fixed rectangular region $R \subset I_a$ is defined, typically centered within the image, with dimensions (h_p, w_p), where h_p and w_p denote the height and width of the patch, respectively. The corresponding patch P_b is extracted from the same region R in image I_b.

The augmentation process blends the extracted patch P_b into image I_a within region R, resulting in an augmented image I_{aug} defined as:

$$I_{aug}(x,y) = \begin{cases} (1-\alpha)\cdot I_a(x,y) + \alpha \cdot I_b(x,y), & if(x,y) \in R \\ I_a(x,y), & otherwise \end{cases} \quad (1)$$

Here, $\alpha \in [0,1]$ represents the blending ratio, controlling the influence of the inserted patch from I_b. In this study, α is fixed at 0.3, ensuring a subtle yet effective introduction of new local features without overpowering the original image content.

The augmented dataset $\mathcal{D}_{aug}$ is constructed by applying RGM-Aug to each image in $\mathcal{D}$, effectively doubling the dataset size and enhancing the diversity of feature representations. This enriched dataset is then utilized to train lightweight CNN model.

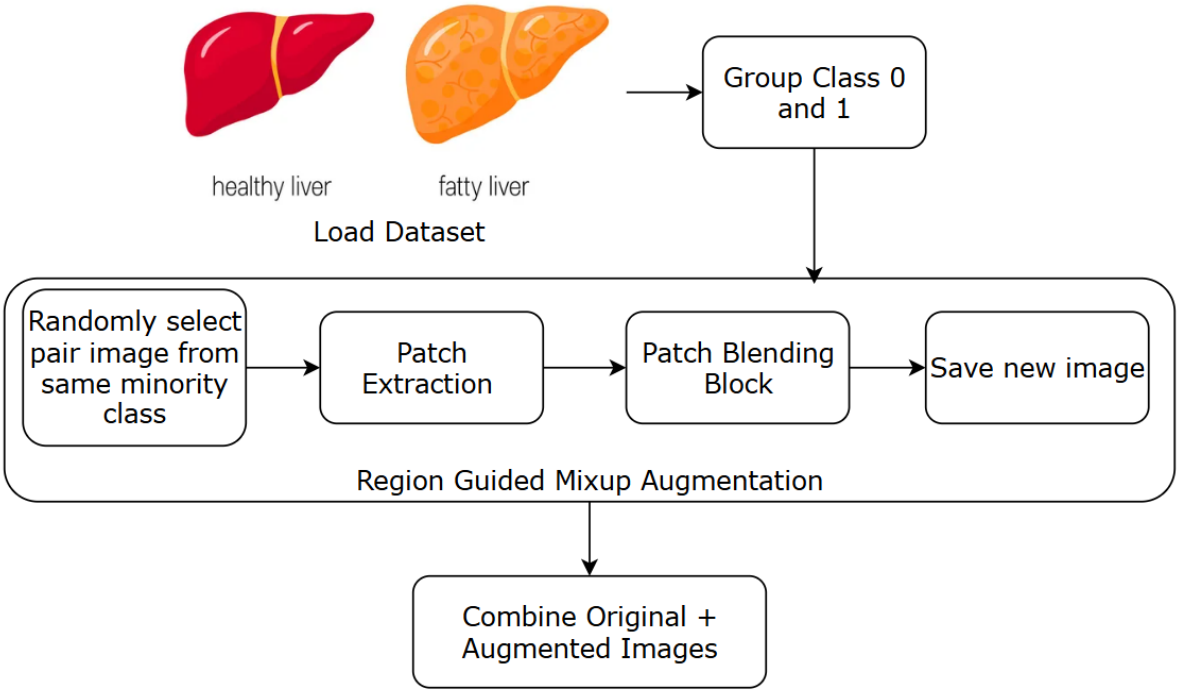

Fig 2. Internal Process Flow of Region-Guided Mixup Augmentation

Algorithm 1: Region-Guided Mixup Augmentation (RGM-Aug)
Input:
- D: Set of ultrasound images grouped by class $\{I_1, I_2, ..., I_n\}$
- (p_{width}, p_{height}): Fixed patch size
- α: Blend ratio (constant value, e.g., 0.3)

Output:
- D_{aug}: Set of augmented images

for each class label C ∈ D do
  Retrieve all images $\{I_1, I_2, ..., I_k\} \in C$
  for each image $I_a \in C$ do
    Randomly select another image $I_b \in C$, where $I_b \neq I < sub > a < /sub >$
    Define fixed rectangular region R:
      $(x_o, y_o) \leftarrow$ Center coordinates of image I_a
      Dimensions of R $\leftarrow (p_{width}, p_{height})$
    Extract patch P_b from I_b within region R
    Blend patch P_b into corresponding region of I_a to generate I_{aug}:
      for each pixel $(x, y) \in R$ do
        $I_{aug}(x, y) \leftarrow (1 - \alpha) \times I_a(x, y) + \alpha \times P_b(x, y)$
      end for
      for each pixel $(x, y) \notin R$ do
        $I_{aug}(x, y) \leftarrow I_a(x, y)$
      end for
    Save I_{aug} in D_{aug} under class label C
  end for
end for
return D_{aug}

2.3 Experimental Setup

The lightweight MobileNetV3-Small convolutional neural network was incorporated with the RGM-Aug methodology. The suggested model is deployed with the help of the TensorFlow framework and conducted on the computer with the NVIDIA GPU accelerator.

The dataset used in this paper was provided by Byra et al.[14], composed of the ultrasound B-mode liver images used in the detection and classification of NAFLD. The dataset contains 145 subjects which fall into two different classes:

- Normal liver: 75 subjects
- Fatty liver: 70 subjects

All the ultrasound images were obtained with the help of standardized clinical protocol utilizing the Siemens Acuson S2000 machine and 6 MHz ultrasound frequency with right intercostal acquisition strategy, providing core similarity in the imaging quality. All the subjects were scanned by an experienced radiologist, and images were annotated by two independent experts manually.

To capture the diversity of patients and generalizability, the data set has a large span of demographics, adults aged 25-68 and distributed equally between men and women. The ground-truth labels were obtained with MRI Proton Density Fat Fraction (PDFF) measurement, so the accuracy and reliability are high and clinically accurate.

This dataset was chosen because of public access, standardized acquisition protocol, and its previous usage in several benchmarking studies. These properties make it a perfect resource towards both formulating and testing the proposed algorithm with a guarantee of reproducibility.

To make the dataset compatible in all the models, the ultrasound images[14] were resized to a fixed value of $224\ X\ 224$. The dataset was divided into two parts comprising of 80:20 training and validation data sets.

All models were trained under the contributed number of epochs and the same preprocessing pipelines to tackle fair assessment. All models were pretrained on ImageNet and fine-tuned on the liver ultrasound dataset. The use of binary cross-entropy as the objective function was done because the nature of the problem is binary cross-entropy. Weight updates were done

using the Adam optimizer and the dropout layer was used to avoid the problem of overfitting. The training was performed in 25 epochs and a mini batch of 32. Table 1 summarizes the hyperparameter settings in detail applied to each model.

Table 1. Hyperparameter Settings Used in the Experiments

Hyperparameter	MobileNetV3-Small
Input Image Size	224 × 224
Pretrained Weights	ImageNet
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Number of Epochs	25
Loss Function	Binary Crossentropy
Dropout Rate	0.5
Blending Ratio (α)	0.3
Patch Size for RGM-Aug	32 × 32

3 Results and Discussion

3.1 Quantitative Results

Models were both trained and tested using the original and the RGM-Augmented datasets. Results in Table 2 summarize the comparative findings. In the original dataset MobileNetV3-Small achieved an accuracy of 97.27%. Training took 23 seconds on mobile NetV3-small. The testing time of all models was low, and the best inference time was 0.96 seconds of MobileNetV3-Small.

With the recommended augmentation being applied, the condition improved significantly. Similar results were displayed by MobileNetV3-Small that has moved to accuracy of 98.63%. The training time was high since it had extended augmented dataset, but the testing time was still computationally efficient with all models having inference time below 4 seconds.

Results of past research were also used to be the benchmark. Using Inception-ResNet-v2 as the model with SVM as the classifier, Byra et al. (2018)[14] achieved an accuracy of 96.30% on the same dataset. Using DenseNet-121, Miriam Naim Ibrahim et al. (2023)[15] were able to achieve the best AUROC of 90.1%, with sensitivity and specificity of 95.00% and 85.20%, respectively. The accuracy according to Petros Drazinos et al. (2025)[7] with DenseNet-201 was 92.53 percent after they used the augmentation technique.

Table 2. Comparative Results

Model	Dataset	Accuracy (%)	Micro-F1 (%)	Macro-F1 (%)	Training Time (s)	Testing Time (s)
MobileNetV3-Small	Original	97.27	97.27	96.53	23.00	0.96
RGM-Aug- MobileNetV3-Small (Proposed)	Augmented	98.63	98.63	98.23	42.24	1.19
Byra et al.[14]	Original	96.30	-	-	-	-
Miriam Naim Ibrahim et al.[15]	Original	90.1 (AUROC)	95.00	85.20	-	22.64s per patient
Petros Drazinos et al.[7]	Augmented	92.53	-	-	-	-

Figure 3 represents the training behavior of MobileNetV3-Small and its classification ability using the RGM-Augmented dataset. In the figure (the left panel), a trend graph of accuracy is represented where training and validation accuracy are increasing continuously as the epoch progresses. The validation accuracy is nearly 97% at epoch 20 and it does not vary much as the model is able to generalize effectively. The right plots of loss curves demonstrate that both the training and validation loss are declining significantly and smoothly, and there are no symptoms of overfitting. The validation loss steadily follows the training loss, which also confirms the steady learning process assisted by the offered augmentation approach.

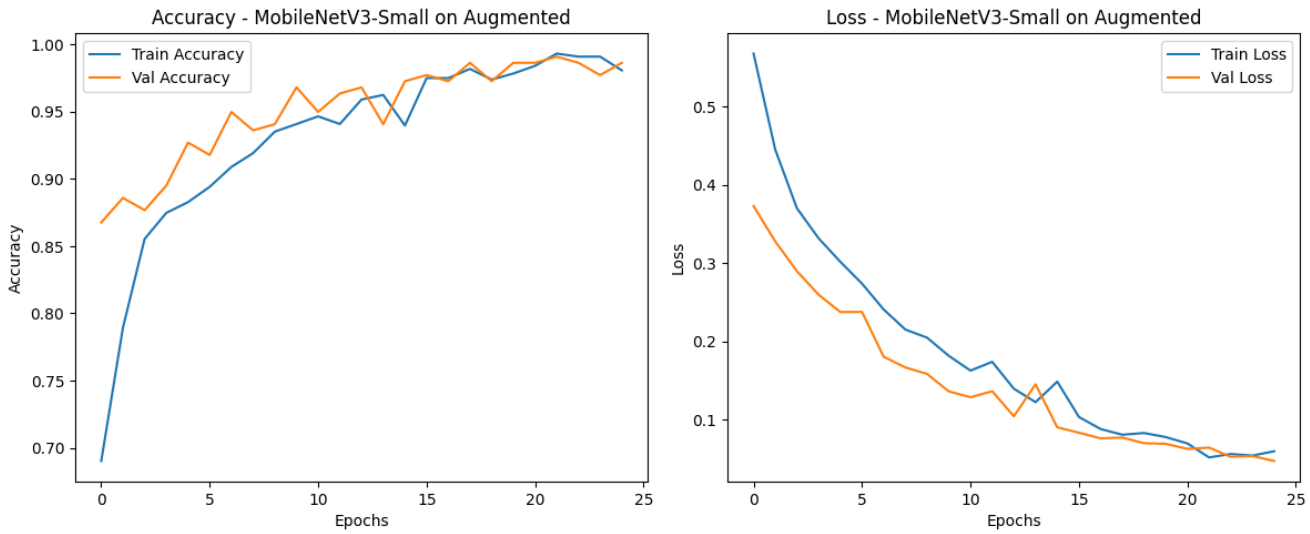

Fig 3. Accuracy and loss curves of MobileNetV3-Small trained on the RGM-Augmented dataset. *(Left: Training vs. validation accuracy; Right: Training vs. validation loss across 25 epochs)*

3.2 Discussion

Experimental findings prove that the suggested RGM-Aug powerfully increases the classification ability of ultrasound liver steatosis detection by lightweight deep learning models. RGM-Aug is also contrary to the traditional augmentation methods using global augmentations (e.g., rotate, flip, noise) that inject random intra-class variability by blending the patches at the global level. Such a mechanism provides local texture that attempts to achieve natural differences within a class via simulation of natural intra-class variation facilitating enhanced generalization ability without voided semantic integrity.

The augmentation method produced outstanding results in all the tested models. These results confirm the effectiveness of focalized feature boosting as an aid to learning of models especially in the data limited medical imaging contexts. Such improvements are not only present in both indicators of performance but also training behavior, validating the effectiveness of RGM-Aug as an augmentation strategy.

Another important finding that was noted in previous works on this problem[7],[14],[15] is that traditional augmentation methods that allowed to lift classification rates to over 90%, that Micro-F1 and the Macro-F1 metrics tended to demonstrate a significant difference. This disparity suggests that the classifiers in the said studies, although having reported high overall accuracy, had suffered class imbalance in terms of predictive ability, namely most would perform well when predicting the majority class and worse when predicting the minority class. Micro-F1, which is a combination of the performance of all classes, can be kept high even in case the model does not work well on minority class samples, but Macro-F1, which counts each class equally, will fall in those circumstances. Such unbalancing shows that the introduced augmentation strategies are inefficient since they do not achieve representative and sufficiently diverse data across all classes.

The novelty of RGM-Aug is that it incorporates the concept region-guided intra-class patch blending, an algorithm that synthesizes images with more local texture diversity that is at the same time semantically unperturbed. In contrast to MixUp and CutMix that frequently result in slightly improbable boundaries or lesion confusion, which is especially undesirable in grayscale ultrasound images, RGM-Aug blends are limited to a realistic area, and augmented images remain anatomically valid. This specifically overcomes a significant limitation that occurs in medical image augmentation: the coherence of classes in maintaining accuracy in diagnosis.

Although the extra data generation step was added, the net result was a relatively small increase in training time and almost no change in testing time. In addition, RGM-Aug is a preprocessing strategy that can be used offline; as such, it can be deployed in environments with limited resources.

There are several limitations, even though there are benefits. On the other hand, the predetermined mixing ratio and patch size and size are to ensure the reproducibility, but they might not be the best options available to support all the aspects of the image features or the pathological grades. The approach lacks more dynamic representation of the complexity and spatial location of the diagnostic features that may constrain its use to more diverse data sets. Moreover, although the augmentation

provides intra-class diversity, it does not directly concern inter-class enhancement of bars, which can be profoundly vital when applied to more grading-types of diseases. Finally, the augmentation itself remains intra-class based, which may be expanded to a system of adversarial augmentation or contrastive learning to enhance the discrimination of features further.

3.2.1. Clinical Relevance and Societal Benefits:

The presented RGM-Aug + MobileNetV3-Small model could be of vital use to clinical practice with its enhanced efficiency in detecting NAFLD in non-invasive ways through the pulse wave ultrasound images. Timely and accurate NAFLD diagnosis is essential since this disease has no apparent symptoms at the initial stages but may degenerate into cirrhosis or liver cancer in case it is not diagnosed in time. The reason behind the high accuracy of the proposed model is its ability to show balanced macro and micro F1-scores, which serves as the guaranty that the model will perform well regardless of whether it is a mild or severe case.

Clinically, the framework can be plugged into existing B-mode systems without the need to install additional equipment or provide manual support. Radiologists can utilize the automatically augmented patches and enhanced classification outputs to make timely and more certain diagnostic decisions and may also reduce inter-observer variability in a timely manner.

About the social good, the proposed approach aids in the implementation of early screening projects in clinics and low-resource areas where sophisticated diagnostic devices are not available. The approach facilitates preventive healthcare by allowing efficient and automatized analysis, minimizing the long-term load on tertiary health institutions and improving patient outcomes. Also, the mobile design of MobileNetV3-Small makes it possible to implement handheld ultrasounds, which benefits the expansion of mobile medical initiatives in a rural setting and telemedicine.

4 Conclusion

The paper proposes RGM-Aug as a potential method to augment the limited and homogeneous data of liver steatosis using ultrasound imaging to facilitate liver steatosis classification. The proposed approach further diversifies intra-class samples through integrating localized areas of similar class instances and leads deep learning models to generalize better. The MobileNetV3-Small model was augmented, and the results included an average of classification accuracy, and micro- and macro-F1 scores.

Model evaluation (of ones trained using RGM-Aug) on the experimental dataset consistently showed an improvement in results compared to the models used on the original dataset. It is worth noting that the MobileNetv3-Small obtained the accuracy of 98.63% and Macro F1-score of 98.23%, which proves the robustness and reliability of the introduced technique. The findings also indicate that localized augmentation as compared to conventionally used global perturbations helps lead to better feature discrimination and enhanced learning efficiency.

The research loading determines that RGM-Aug is a computationally efficient and practical strategy to the extent of data augmentation leading to the increased diagnostic performance in medical image classification in data-scarce responsibilities. Meanwhile, its knack of simplicity in concept grants an easy seamless adoption to a variety of different CNN-based architecture with no structural changes required to fit properly. In the future, the model can be oriented towards combining adaptive patch sizing with context-adaptive blending ratios or include the RGM-Aug model into self-supervised learning or transformer-based architectures to capture semantic features at a higher level. These improvements would additionally increase the usability of the method to more domains of medical images analysis, with the same computational efficiency.

5 Acknowledgement

The authors thank, DST-FIST, Government of India for funding towards infrastructure facilities at St. Joseph's College (Autonomous), Tiruchirappalli – 620002.

References

1) Yu NR, Huang S, Deng ZT, Wang J, Shen Y, Leng Y, et al. Design, synthesis and biological evaluation of Alisol B derivatives for potential treatment of non-alcoholic steatohepatitis. *Bioorganic Chemistry*. 2025;155:108132. https://doi.org/10.1016/j.bioorg.2025.108132.
2) Hudson D, Afzaal T, Bualbanat H, AlRamdan R, Howarth N, Parthasarathy P, et al. Modernizing metabolic dysfunction-associated steatotic liver disease diagnostics: the progressive shift from liver biopsy to noninvasive techniques. *Therapeutic Advances in Gastroenterology*. 2024;17:17562848241276334. https://doi.org/10.1177/17562848241276334.
3) Lazaros K, Adam S, Krokidis MG, Exarchos T, Vlamos P, Vrahatis AG. Non-Invasive Biomarkers in the Era of Big Data and Machine Learning. *Sensors*. 2025;25(5):1396. https://doi.org/10.3390/s25051396.
4) Thakur GK, Thakur A, Kulkarni S, Khan N, Khan S. Deep learning approaches for medical image analysis and diagnosis. *Cureus*. 2024;16(5). https://doi.org/10.7759/cureus.59507.

5) Bozic D, Podrug K, Mikolasevic I, Grgurevic I. Ultrasound methods for the assessment of liver steatosis: a critical appraisal. *Diagnostics*. 2022;12(10):2287. https://doi.org/10.3390/diagnostics12102287.
6) Wu CH, Sheu JC, Chou PL, Lee J, Nien HC. Automatic Real-Time Detection and Diagnosis of Liver Tumor with Ultrasound. *Journal of Hepatocellular Carcinoma*. 2025;p. 1599–611. https://doi.org/10.2147/JHC.S524311.
7) Drazinos P, Gatos I, Katsakiori PF, Tsantis S, Syrmas E, Spiliopoulos S, et al. Comparison of deep learning schemes in grading non-alcoholic fatty liver disease using B-mode ultrasound hepatorenal window images with liver biopsy as the gold standard. *Physica Medica*. 2025;129:104862. https://doi.org/10.1016/j.ejmp.2024.104862.
8) Wu J, Liu F, Sun W, Liu Z, Hou H, Jiang R, et al. Boundary-aware convolutional attention network for liver segmentation in ultrasound images. *Scientific Reports*. 2024;14(1):21529. https://doi.org/10.1038/s41598-024-70527-y.
9) Bastati N, Perkonigg M, Sobotka D, Poetter-Lang S, Fragner R, Beer A, et al. Correlation of histologic, imaging, and artificial intelligence features in NAFLD patients, derived from Gd-EOB-DTPA-enhanced MRI: a proof-of-concept study. *European Radiology*. 2023;33(11):7729–7743. https://doi.org/10.1007/s00330-023-09735-5.
10) Geetha R, Rani D, Anbarasu K. Applying Deep Learning Methods to Non-Alcoholic Fatty Liver Disease Management. *2024 5th International Conference on Circuits, Control, Communication and Computing (I4C)*. 2024;p. 282–286. https://doi.org/10.1109/i4c62240.2024.10748435.
11) Skwirczyński M, Tabor Z, Lasek J, Schneider Z, Gibała S, Kucybała I, et al. Deep learning algorithm for differentiating patients with a healthy liver from patients with liver lesions based on MR images. *Cancers*. 2023;15(12):3142. https://doi.org/10.3390/cancers15123142.
12) Prakash KN, Saradha S. A Deep Learning Approach for Classification and Prediction of Cirrhosis Liver: Non Alcoholic Fatty Liver Disease (NAFLD). *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*. 2022;p. 1277–1284. https://doi.org/10.1109/ICOEI53556.2022.9777239.
13) Qin S, Hou X, Wen Y, Wang C, Tan X, Tian H, et al. Machine learning classifiers for screening nonalcoholic fatty liver disease in general adults. *Scientific Reports*. 2023;13(1):3638. https://doi.org/10.1038/s41598-023-30750-5.
14) Byra M, Styczynski G, Szmigielski C, Kalinowski P, Michałowski Ł, Paluszkiewicz R, et al. Transfer learning with deep convolutional neural network for liver steatosis assessment in ultrasound images. *International Journal of Computer Assisted Radiology and Surgery*. 2018;13:1895–1903. https://doi.org/10.1007/s11548-018-1843-2.
15) Ibrahim MN, Blázquez-García R, Lightstone A, Meng F, Bhat M, Kaffas AE, et al. Automated fatty liver disease detection in point-of-care ultrasound B-mode images. *Journal of Medical Imaging*. 2023;10(3):034505. https://doi.org/10.1117/1.JMI.10.3.034505.