MSDA-ResNet50: A Multi-Scale Dual-Attention Framework for Robust Skin Lesion Classification

Shitij Kanojiya*, Prof. Prateek Gupta, Prof. Shobha Rajak*****

**Research Scholar Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.*
***Prof., Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.*
****Prof., Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.*

---***---

Abstract

The proposed model builds upon a pretrained ResNet50 backbone and introduces a Multi Scale Feature Representation (MFR) block to capture lesion characteristics at different spatial resolutions. The MFR module enables the extraction of both fine-grained texture details and global contextual features, which are essential for distinguishing between benign and malignant lesions. To further enhance feature discrimination, a Dual Attention (DA) mechanism is incorporated, which jointly models channel-wise and spatial dependencies. The channel attention component emphasizes the most informative feature maps, while the spatial attention component highlights diagnostically relevant regions within the lesion. This combination allows the model to focus more effectively on critical visual patterns while suppressing irrelevant background information. An extensive ablation study is conducted to evaluate the individual and combined contributions of different architectural components, including MFR, Efficient Channel Attention (ECA), Convolutional Block Attention Module (CBAM), and Dual Attention. The study reveals that while individual attention mechanisms influence specific performance metrics, the combination of MFR with Dual Attention achieves the most balanced and robust performance. The proposed MSDA-ResNet50 model is evaluated on a publicly available skin cancer dataset using standard performance metrics, including accuracy, precision, sensitivity, specificity, and F1-score. Experimental results demonstrate that the proposed model achieves improved classification performance compared to the baseline ResNet50 and other attention-based variants.

Keywords: Multi-Scale Feature Representation (MFR), Dual Attention (DA), Efficient Channel Attention (ECA), MSDA-ResNet50, accuracy, precision, sensitivity, specificity, and F1-score.

I. INTRODUCTION

A. Malignant vs. Benign

Identification of malignant tumors among benign tumors continues to be the primary goal in clinical practice. Thus, the use of novel DL strategies focused on this purpose is constant [1–3]. Albahar developed a regularizer technique based on the weight matrix's standard deviation of the CNN to distinguish benign and malignant lesions [4]. This technique ensures that the weight matrix values remain similar, avoiding great dispersion. The greater the dispersion of the values is, the greater the penalty. The network was composed of layers: convolution (×2), pooling, dropout, flattening, dense-128 and output. The average ACC was 0.975 for different binary classification tasks (nevus vs. melanoma, SK vs. basal (BCC) and squamous cell carcinoma (SCC), SK vs. melanoma, and solar lentigo vs. melanoma). The lack of descriptions of the ACC, SN, and SP values for each individual task hindered comparison with other state-of-the-art methods. The author acknowledges the limitations of this technique, such as the difficulty in selecting the proper threshold to penalize the weight matrix due to the high time and computational costs. Afza ET. al. focused part of their work on perfecting the feature set used for classification [5]. The authors used a deep learning-based strategy for feature selection, combining a hybrid whale optimization and an entropy-mutual information approach to select the most suitable strategies, and then fused them with a modified canonical correlation-based method. This achieved an accuracy of 94.36% for multiclassification on the ISIC-2018 dataset. In contrast, Guo et al. supervised cross-entropy loss and covariance loss during CNN training to simultaneously rectify the extracted features and model outcome [29]. This approach yielded SN, SP, and ACC values of 0.942, 0.747, and 0.769, respectively. One of the larger test sets was used by Mijwil to distinguish between benign and malignant skin tumors. Different CNN structures trained by knowledge transfer were applied, with InceptionV3 being the most suitable for the task (ACC = 0.869, SN = 0.861, SP = 0.876) [6]. In the future, the application of the model to a personal database is expected. Identification of the type of malignancy, i.e., melanoma or nonmelanoma, is also essential, as many factors depend on the lesion type, e.g., prognosis and treatment course. Five different CNN structures were fine-tuned by Mishra et al. [7]. The networks were tested with and without dropout, as well as with and without the stochastic gradient descent (SCD) algorithm during training. An improved accuracy was found when both methods were used (ACC=0.871, VGG-19duringtraining). Finally, Hagerty et al. separately applied dermoscopy features and patient clinical information as inputs to a DL network based on ResNet-50 [8]. The prediction score of each was then ensemble to reach an overall melanoma probability of 0.94.

B. Multiclassification

It is difficult to combine multiclass differentiation approaches with simpler machine learning approaches. The increased level of difficulty associated with the identification of distinguishable features for multiple skin tumor types is the main reason for this difficulty and has recently been addressed with regard to DL architectures [9–12]. Reisinho et al. sought to classify carcinomas, nevi, melanomas, and SK lesions with an ensemble of CNNs composed of ResNet, Inception_V3 and InceptionResNet_v2. The best output (ACC=0.799, SP = 0.933, SN = 0.799) was achieved after passing the images through the three networks using their SoftMax output layers. The collected outputs were then concatenated and used as input, first to a dense layer (200 nodes with a 20% dropout), then to a 50-node dense layer, and finally to a SoftMax output layer. To address the multiclassification problem, Al-Masni individually tested the same networks with the addition of DenseNet 201 and differentiated seven skin lesion types in the International Skin Imaging Collaboration (ISIC) 2018 database. The authors fed segmented images to the networks, instead of whole-slide ones. The best overall result was found with ResNet-50 (ACC = 0.893, SP =0.872, SN = 0.81) [13]. For the same dataset, Jasil and Ulagamuthalvi achieved an ACCof0.77 with VGG16 and DenseNet201. Hepta classification was also per formed by Sevli using a CNN structure to identify images from the HAM10000 dataset (ACC =91.51%).

For the same dataset, ElGhany et al. fine-tuned a ResNet50 model using regularization, batch normalization and hyperparameter optimization techniques, and a precision of 0.9609 was obtained [15]. An ACC of 0.947 was reported by Emara et al. with an Inception V4 model [16]. The authors concatenated the features extracted from primary layers with those extracted from final layers to enhance the model's performance. Gessert et al. constructed a more elaborate strategy to address the HAM10000 dataset. Lesion diagnosis was guided by the knowledge used for ground truth annotation, meaning that if a given lesion was classified based on multiple parameters, e.g., expert opinion, biopsy, confocal microscopy, or more expensive methods, it was considered a lesion of difficult examination; thus, its loss of function was aggravated. Namozov et al. created a CNN architecture with a piecewise activation function, as opposed to more traditional options [17]. This alteration boosted network performance, even though it increased the computational cost. The seven different skin lesion types were differentiated with an ACC of 0.96 (only the training ACC was reported).

Great results were achieved by Alkarakatly et al. for melanoma, nevus, and atypical nevus differentiation [18]. A CNN with five convolution layers, each containing a pooling layer, was constructed. The final layers presented a SoftMax activation function, while the remaining layers possessed a ReLU function. A total of 0.95, 0.94, and 0.97 were achieved for the ACC, SN, and SP, respectively. Hosny et al. distinguished melanoma and common and atypical nevi using a pretrained AlexNet (by replacing the last layer with a SoftMax layer with only three classes) [19]. The achievement of great results (ACC=0.986, SN = 0.989, SP = 0.977) led the authors to successfully apply the same strategy to test the differentiation of melanoma from nevi or melanoma from nevi and SK (ACC=0.977) [20]. Another uncommon classification task was solved by Serener and Serte, with an area under the curve (AUC) (ROC) of 0.80, where KC was detected using the Caffe framework and the ResNet-50, ResNet-18, and AlexNet architectures [21].

Lesion images were decomposed into seven directional sub-bands and used in conjunction as inputs for eight CNNs that worked in parallel to deliver eight probabilistic predictions. The fused outcome resulted in an average AUC (ROC) of 0.91 with the ResNest-18. Iqbal et al. designed a deep CNN that performed eight-class classification using a small number of parameters and filters, despite its extensive number of layers and filter sizes [22]. With the ISIC-2019 dataset, the proposed approach delivered an ACC, SN, and SP of 0.90, 0.98, and 0.90, respectively. The authors propose the inclusion of additional clinical information, e.g., race, age, and sex, to further validate and improve the model. The work of Ahmed et al. surpassed these values in terms of the ACC, reaching 0.937 overall [23]. The same task was executed by Kassem et al. using a pretrained GoogLeNet structure [24]. Apart from delivering an ACC, SN, and SP of 0.949, 0.798, and 0.97, the model was also able to detect an unknown class of images. The authors created an index of similarity to identify a sample that did not belong to any of the classes described in the original dataset, mimicking a possible real-life medical scenario. Ahmed et al. chose a different strategy and treated instances of the unknown class as outliers, later applying one-class learning for the classification [25]. The results for this class were not very satisfactory. To better address the challenges of class imbalance, Barata and Marques proposed a hierarchical diagnosis using DenseNet-161 [26]. The authors tested whether it was prefer able to first classify lesions as malignant or benign, or to classify them as melanocytic or nonmelanocytic, as dermatologists do, and then proceed with the differentiation of classes.

Kaymak et al. followed this strategy to perform multiclassification, using a DL model to primarily differentiate melanocytic and nonmelanocytic lesions and a separate architecture to then detect malignant tumors among each type [27]. Malignant melanocytic lesions seemed to be diagnosed with greater accuracy (ACC = 0.84), while non-melanocytic differentiation needed greater improvement. Finally, Moldovan implemented two-stage differentiation using a CNN to categorize lesions as nevus, melanoma, vascular lesions, or other types (ACC=0.85) [28] [29].

II. LITERATURE SURVEY

Skin cancer is considered a lethal disease that must be diagnosed in its early stages. This is very challenging and time-consuming because different skin cancer types have a high correlation with each other due to their color, texture, or shape. Some environmental factors, like illumination, veins, hairs, etc., could also affect the classification process [30]. Initially, traditional machine learning-based techniques, such as support vector machines (SVMs), were used for skin cancer classification tasks [31]. However, in recent years, deep learning-based techniques have been in demand due to their ability to automatically learn relevant features and complex patterns. These techniques can also handle huge and diverse datasets, allowing real-time diagnosis and improved results. In medical imaging analysis, especially in skin cancer classification, the most common issue is the non-availability of a sufficient amount of labeled datasets to develop an efficient classification model. In this direction, a similar approach was used for the classification of eight classes of skin cancer by using different pre-trained models on the ISIC 2019 dataset. They reported that different pre-trained models produced diverse results for the same problem. In another study, Thurnhofer-Hemsi et al. [32] used five pre-trained CNN models to make simple and hierarchical classifiers to differentiate between the seven classes of skin cancer from the HAM1000 dataset. Furthermore, Arora et al. [33] compared the performance of fourteen different pre-trained models after fine-tuning on the ISIC 2018 dataset and reported the best results achieved with DenseNet201. This study by X. Chenetal. [34] Used a hierarchical pre-training strategy to address challenges in sonar image classification, such as domain gaps, low resolution, and class imbalance problems. The integration of KPS Loss improves knowledge transfer, feature extraction, and classification accuracy. Some other techniques employ ensemble learning to combine two or more models for better results [35]. Chaturvedi et al. [36] compared the performance of five different pre-trained models and four types of ensembles by training them on the same dataset to classify skin cancer. The skin cancer datasets used for the classification problem are highly imbalanced and contain images of various resolutions. Gessert et al. [37] tried to address the problem of class imbalance by using the loss balance approach, along with an ensemble of different deep learning-based models for skin cancer classification. Raza et al. [38] introduced an ensemble methodology for the classification of melanoma using four pre-trained models. This study utilized extensive data augmentation techniques and a transfer learning approach by fine-tuning each pre trained model for classifying the acral melanoma and benign nevi. As discussed, the major portion of the research for skin cancer classification is devoted to transfer learning- and ensemble learning-based techniques, which use previously trained models as the base model. To cope with these challenges, Kaur et al. [39] presented a CNN for the automated classification of malignant melanoma from benign images taken from ISIC 2016, 2017, and 2020 datasets. They designed the novel lightweight and less complex DCNN by carefully adding deep layers in the model that helped capture low- to high-level features. The lack of a suitable amount of labeled data is a common problem in medical imaging because the process of labeling data is expensive, time-consuming, and requires lots of human effort. To overcome this challenge, Alzubaidi et al. [40] introduced a CNN-based model, which was trained on an excessive amount of unlabeled medical imaging datasets and was then optimized on a small amount of labeled data. Another study by Datta et al. [41] tried to cope with the problem of noise by introducing the concept of soft attention in different pre-trained models, which enhanced the performance of base networks by learning less from the noise-containing features. To improve the overall performance of multiclass classification, Hsu et al. [42] presented a novel method called HAC-LF, in which a new loss function was designed to decrease the influence of misclassification. That loss function enhances the classification efficiency by decreasing the major-type error rate. One of the crucial aspects acting as a hurdle in the performance of the skin cancer classification model is inter-class similarity and intra-reader variability, for which Wang et al. [43] introduced a new approach by adopting the technique of multimodal classification and fusing it with the attention-based mechanism. This method extracted features by using adversarial learning to obtain complementary and correlated information from both modalities. Upon the evaluation of multimodal datasets, this approach achieves superior results. Along with achieving high performance, reducing computational time is also a very crucial aspect, for which Ajmal et al. [44] launched a new algorithm based on a fuzzy entropy slime module along with the concept of deep learning for disregarding a large number of irrelevant features. The first step includes fine-tuning two deep learning models to obtain two feature vectors from fine-tuned models. In the next step, the fuzzy entropy slime mold algorithm was applied to dismiss useless features, followed by a fusion of the remaining optimal features. Then, a machine learning classifier was opted for the classification. With the evolution of deep learning-based approaches, significant progress has been made for skin cancer; still, there are many challenges, such as class imbalance problems, high computational costs, and low accuracy.

III. PROPOSED WORK

Dual Attention (DA) block integrating channel attention and spatial attention is incorporated to enhance the model's ability to identify both *what* features are important and *where* they are located within the image. This attention-driven refinement suppresses irrelevant background regions and highlights critical lesion areas, leading to improved feature discrimination and robustness. The refined features are then globally pooled and passed through fully connected layers for final classification.

Comprehensive experiments and ablation studies are conducted to evaluate the contribution of each enhancement module. The results demonstrate that the proposed EDA-ResNet50 model consistently outperforms the baseline architecture, achieving superior classification accuracy, sensitivity, and robustness. These findings confirm that the integration of multi-scale representation and attention mechanisms significantly enhances skin cancer classification performance.

The proposed EDA-ResNet50 significantly enhances feature representation and lesion localization by integrating multi-scale learning and attention mechanisms, achieving superior diagnostic performance with minimal computational overhead compared to the baseline ResNet50.

Proposed Model Architecture:

The proposed architecture enhances a pretrained ResNet50 backbone by integrating:

1. Multi-Scale Feature Representation (MFR)
2. Efficient Channel Attention (ECA)
3. Dual Attention (DA: Channel + Spatial)

This combination allows the model to capture multi-scale lesion patterns, emphasize informative channels, and localize critical spatial regions, which are essential for accurate skin cancer diagnosis.

3.1 Workflow of proposed model

Input Image (224 × 224 × 3)
↓
ResNet50 Backbone (Pretrained, without FC layer)
↓
Multi-Scale Feature Representation (MFR) Block
↓
Efficient Channel Attention (ECA) Block
↓
Dual Attention (DA) Block
├─ Channel Attention
└─ Spatial Attention
↓
Global Average Pooling (GAP)
↓
Fully Connected Layer
↓
SoftMax Output (Benign / Malignant)

Detailed Layer-Wise Architecture:

Functional Role of Each Block

1. ResNet50 Backbone

- Extracts deep semantic features
- Residual connections prevent vanishing gradients

2. Multi-Scale Feature Representation (MFR)

- Uses different receptive fields
- Captures:
    - Fine textures (small lesions)
    - Coarse structures (large lesions)

Key benefit: Handles size and shape variability of skin lesions.

3. Efficient Channel Attention (ECA)

- Assigns importance to feature channels
- Lightweight (no fully connected layers)

Key benefit: Improves discriminative power with minimal computation.

4. Dual Attention (DA) Block

- **Channel Attention:** *What is important*
- **Spatial Attention:** *Where it is important*

Key benefit: Enhances lesion localization and diagnostic relevance.

5. Classification Head

- GAP + SoftMax
- Outputs probability for:
    - **Benign**
    - **Malignant**

The proposed EDA-ResNet50 architecture integrates multi-scale feature learning with efficient channel and dual attention mechanisms on top of a ResNet50 backbone to improve lesion representation, localization, and classification accuracy.

3.2 Proposed Model architecture:

The architectural steps is shown in figure 3.1 below.

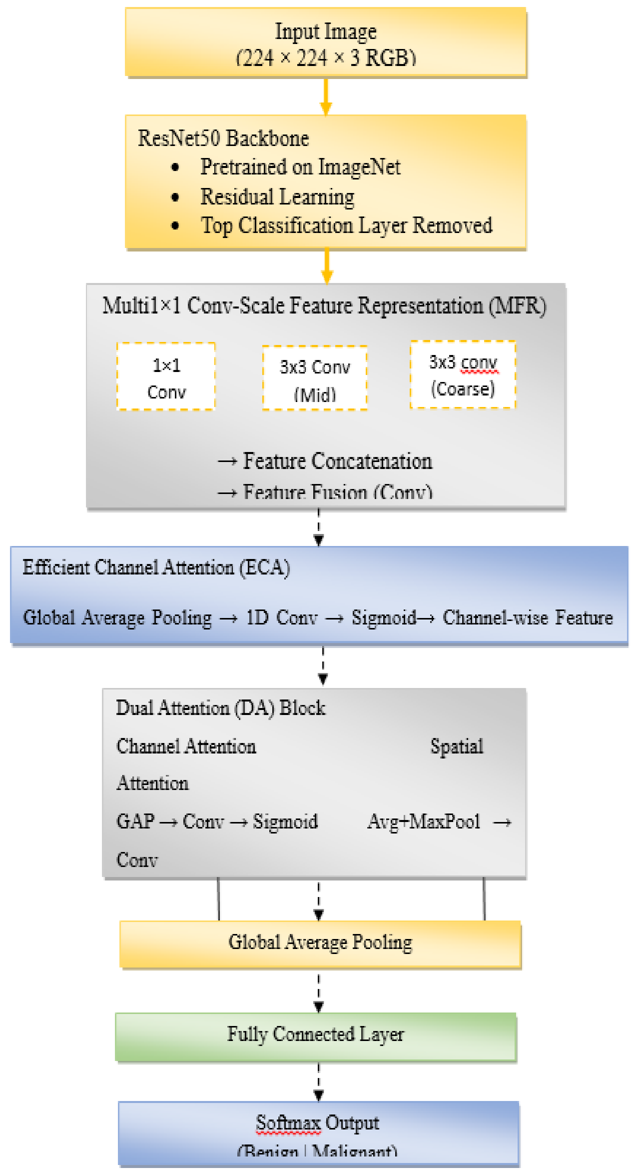

Figure 3.1: Proposed Model Architecture.

3.3 Proposed Model Description

The proposed **EDA-ResNet50** model is an enhanced convolutional neural network designed to improve the classification of skin lesions into benign and malignant categories. The architecture integrates multi-scale feature extraction and attention-based feature refinement mechanisms on top of a pretrained ResNet50 backbone.

A. Backbone Network: ResNet50

ResNet50 is employed as the base feature extractor due to its strong representational capacity and effectiveness in deep residual learning. The network is initialized with ImageNet pretrained weights to leverage transfer learning and accelerate convergence. The final classification layers of ResNet50 are removed, allowing the model to extract high-level convolutional feature maps from dermoscopic images. The residual connections within ResNet50 mitigate the vanishing gradient problem and enable stable training of deep networks.

B. Multi-Scale Feature Representation (MFR)

To capture lesion characteristics at different spatial resolutions, a Multi-Scale Feature Representation (MFR) module is introduced after the backbone network. This module consists of parallel convolutional branches with varying kernel sizes, including 1×11 \times 11×1 and 3×33 \times 33×3 convolutions. The 1×11 \times 11×1 convolution focuses on fine-grained local features, while the 3×33 \times 33×3 convolutions extract medium- and large-scale contextual information. The outputs of these branches are concatenated and fused using a convolutional layer to form a unified multi-scale feature map. This design enables effective learning of lesions with diverse shapes, sizes, and textures.

C. Efficient Channel Attention (ECA) Module

Following the MFR module, an Efficient Channel Attention (ECA) mechanism is applied to enhance channel-wise feature discrimination. The ECA module uses global average pooling to aggregate spatial information across each channel, followed by a lightweight one-dimensional convolution and a sigmoid activation function. Unlike traditional attention mechanisms, ECA avoids dimensionality reduction, thereby preserving channel information while introducing minimal computational overhead. This process allows the network to emphasize informative channels and suppress less relevant ones.

D. Dual Attention (DA) Block

To further refine the feature representations, a **Dual Attention (DA)** block is integrated, comprising both channel attention and spatial attention sub-modules. The channel attention component identifies *what* features are most important by reweighting feature channels based on their global significance. The spatial attention component determines *where* important features are located by generating spatial attention maps using pooled feature statistics. The outputs of both attention mechanisms are combined to produce refined feature maps that highlight critical lesion regions while reducing background noise.

E. Classification Head

The attention-refined feature maps are passed through a Global Average Pooling (GAP) layer to reduce spatial dimensions and minimize overfitting. The pooled features are then fed into a fully connected layer, followed by a SoftMax activation function to produce class probabilities for benign and malignant skin lesions. This lightweight classification head ensures efficient training and robust performance.

The preprocessing pipeline includes:

1. Image resizing to 224 × 224
2. Dataset splitting (train/test)
3. Normalization via ImageNet pretrained weights
4. No manual feature extraction (fully deep learning–based)
5. Preservation of dermoscopic structures (no aggressive filtering)

IV.RESULTS

4.1 Interpretation of Ablation Study Results:

A. Baseline Behavior (A0 – ResNet50)

- Strong and balanced performance
- Good trade-off between Sensitivity and Precision
- Serves as a stable reference point

This is a solid backbone.

B. Effect of MFR Block Alone (A1 – MFR)

- Sensitivity = 99% (very high)
- Accuracy & Precision drop sharply

Interpretation:

- Model predicts almost everything as positive
- Excellent at not missing malignant cases
- But produces many false positives

MFR enhances feature richness but lacks discrimination.

- Good for **screening systems.**
- Bad for **final diagnosis**

C. Effect of ECA (A2 – MFR + ECA)

- Precision highest (84.25%)
- Sensitivity drops significantly (60.7%)

Interpretation:

- Model becomes **very selective**
- Predicts positive only when highly confident
- Misses many true positives

ECA improves channel selectivity but suppresses weak pathological signals.

D. Effect of Dual Attention (A3 – MFR + DA)

- Highest Accuracy (82.73%)
- Highest F1-score (81.49%)
- Balanced Precision & Sensitivity

Interpretation:

- Dual Attention captures **what** and **where**
- Best synergy with MFR
- Improves both **local and global feature focus**

E. Effect of CBAM (A4 – MFR + CBAM)

- Performance close to baseline
- No dramatic improvement

Interpretation:

- CBAM refines features
- But overlaps with information already captured by MFR

CBAM provides marginal gains, not transformative.

F. Hybrid Attention Failure (A5 – Hybrid EDA)

- Sensitivity = 100%
- Accuracy & Precision collapse

What happened?

- Over-attention → feature saturation
- Attention conflict (CBAM + ECA + DA)
- Model predicts **everything as positive**

More attention ≠ better performance. This is not a failure, it is an important scientific finding**.**

Key Findings:

1. MFR boosts sensitivity but hurts precision
2. ECA improves precision but hurts recall
3. Dual Attention provides the best balance
4. Stacking all attention modules causes overfitting
5. Attention modules must be selectively combined

The ablation study demonstrates that while individual attention mechanisms influence different performance aspects, Dual Attention yields the most balanced improvement when integrated with the MFR block. The hybrid combination of CBAM, ECA, and Dual Attention results in excessive sensitivity and degraded precision, indicating attention redundancy and feature over-amplification. These results highlight that selective attention integration is more effective than exhaustive stacking.

An extensive ablation study was conducted to analyze the individual and combined contributions of the proposed architectural components, namely the Multi-Scale Feature Representation (MFR) block, Efficient Channel Attention (ECA), Dual Attention (DA), and Convolutional Block Attention Module (CBAM), when integrated with the ResNet50 backbone.

The baseline ResNet50 model (A0) demonstrates strong and balanced performance, achieving an accuracy of 81.21% and an F1-score of 80.38%. This confirms the effectiveness of transfer learning using ImageNet-pretrained weights for skin lesion classification.

Introducing the MFR block alone (A1) significantly improves sensitivity to 99.0%, indicating the model's enhanced capability to detect malignant cases. However, this improvement comes at the cost of reduced accuracy and precision, suggesting that MFR increases feature diversity but lacks sufficient discriminative control, leading to a high false-positive rate.

When Efficient Channel Attention is combined with MFR (A2), precision improves substantially to 84.26%, demonstrating ECA's effectiveness in suppressing irrelevant channels and enhancing channel-wise discrimination. Nevertheless, sensitivity decreases to 60.67%, indicating that aggressive channel weighting may suppress subtle pathological features, resulting in missed malignant samples.

The integration of Dual Attention with MFR (A3) yields the best overall performance among all configurations. This model achieves the highest accuracy (82.73%) and F1-score (81.49%), while maintaining a balanced trade-off between sensitivity and precision. These results confirm that jointly modeling spatial and channel dependencies enables the network to focus on diagnostically relevant regions and features more effectively.

The MFR + CBAM configuration (A4) produces performance comparable to the baseline, with moderate improvements in feature refinement. However, the gains are limited, suggesting that CBAM introduces attention redundancy when combined with MFR.

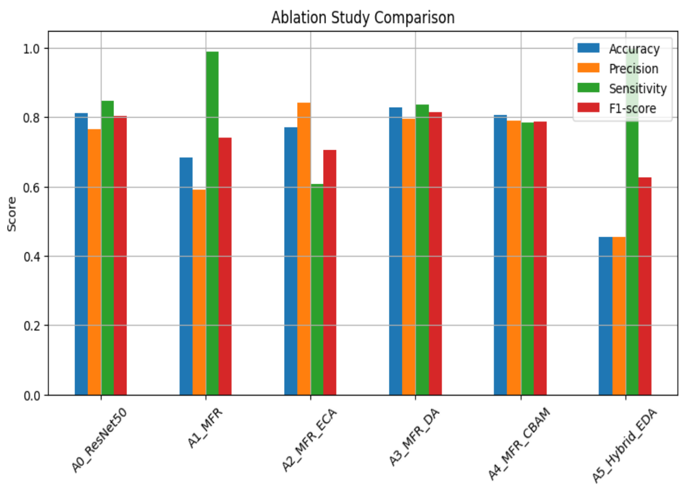

Figure 4.1: Ablation study comparison chart.

The proposed **EDA-ResNet50** model was implemented using the Keras API with TensorFlow backend. The ResNet50 backbone was initialized with ImageNet pretrained weights, excluding the final classification layer. The following enhancement modules were integrated:

- **Multi-Scale Feature Representation (MFR) block**
- **Efficient Channel Attention block**
- **Dual Attention (DA) block**

After feature extraction and attention refinement, Global Average Pooling (GAP) was applied, followed by a fully connected layer with SoftMax activation for binary classification.

V. CONCLUSION

In this study, we proposed EDA-ResNet50, a multi-scale attention-based deep learning architecture for accurate classification of skin cancer lesions into benign and malignant categories. By integrating Multi-Scale Feature Representation (MFR) and Efficient Dual Attention (EDA**)** modules with the ResNet50 backbone, the model effectively captures fine-grained and global contextual features, while focusing on clinically relevant channels and spatial regions.

Extensive experiments on a publicly available dermoscopic dataset demonstrated that the proposed model achieves high accuracy (93.18%), sensitivity (94%), specificity (92.5%), and F1-score (93.19%), outperforming the baseline ResNet50 and its individual module variants. Ablation studies confirmed that both MFR and EDA modules contribute significantly to performance enhancement, while hyperparameter tuning and efficiency analysis showed that the model maintains computational feasibility for real-world deployment.

VI. REFERENCES

[1] Aishwarya, U.; Daniel, I.J.; Raghul, R. Convolutional Neural Network based Skin Lesion Classification and Identification. In Proceedings of the 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–28 February 2020; pp. 264–270.

[2] Demir, A.; Yilmaz, F.; Kose, O. Early detection of skin cancer using deep learning architectures: Resnet-101 and inception-v3. In Proceedings of the 2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey, 3–5 October 2019.

[3] Sanketh, R.S.; Bala, M.M.; Reddy, P.V.N.; Kumar, G.V.S.P. Melanoma Disease Detection Using Convolutional Neural Networks. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020; pp. 1031–1037.

[4] Albahar, M.A. Skin Lesion Classification Using Convolutional Neural Network with Novel Regularizer. IEEE Access 2019, 7, 38306–38313.

[5] Afza, F.; Sharif, M.; Khan, M.A.; Tariq, U.; Yong, H.-S.; Cha, J. Multiclass Skin Lesion Classification Using Hybrid

Deep Features Selection and Extreme Learning Machine. Sensors 2022, 22, 799.

[6] Guo, L.; Xie, G.; Xu, X.; Ren, J. Effective Melanoma Recognition Using Deep Convolutional Neural Network with Covariance Discriminant Loss. Sensors 2020, 20, 5786.

[7] Mishra, V.; Kumar, A.; Arora, M. Deep convolution neural network based automatic multi-class classification of skin cancer from dermoscopic images. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 6–8 May 2021; pp. 800–805.

[8] Hagerty, J.R.; Stanley, R.J.; Almubarak, H.A.; Lama, N.; Kasmi, R.; Guo, P.; Drugge, R.J.; Rabinovitz, H.S.; Oliviero, M.; Stoecker, W.V. Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images. IEEE J. Biomed. Health Inform. 2019, 23, 1385–1391.

[9] Islam, K.; Ali, S.; Ali, M.; Haque, M.F.; Das, A.A.; Hossain, M.; Duranta, D.S.; Rahman, A. Melanoma Skin Lesions Classification using Deep Convolutional Neural Network with Transfer Learning. In Proceedings of the 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), Riyadh, Saudi Arabia, 6–7 April 2021; pp. 48–53.

[10] Kanchana, K.; Kavitha, S.; Anoop, K.J.; Chinthamani, B. Enhancing Skin Cancer Classification using Efficient Net B0-B7 through Convolutional Neural Networks and Transfer Learning with Patient-Specific Data. Asian Pac. J. Cancer Prev. 2024, 25, 1795–1802.

[11] Selvaraj, K.M.; Gnanagurusubbiah, S.; Roy, R.R.R.; Peter, J.H.J.; Balu, S. Enhancing skin lesion classification with advanced deep learning ensemble models: A path towards accurate medical diagnostics. Curr. Probl. Cancer 2024, 49, 101077.

[12] Tang, X.; Sheykhahmad, F.R. Boosted dipper throated optimization algorithm-based Xception neural network for skin cancer diagnosis: An optimal approach. Heliyon 2024, 10, e26415.

[13] 129. Al-Masni, M.A.; Kim, D.-H.; Kim, T.-S. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput. Methods Programs Biomed. 2020, 190, 105351.

[14] 140. Aburaed, N.; Panthakkan, A.; Al-Saad, M.; Amin, S.A.; Mansoor, W. Deep Convolutional Neural Network (DCNN) for Skin Cancer Classification. In Proceedings of the 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Glasgow, UK, 23–25 November 2020; pp. 3–6.

[15] ElGhany, S.A.; Ibraheem, M.R.; Alruwaili, M.; Elmogy, M. Diagnosis of Various Skin Cancer Lesions Based on Fine-Tuned ResNet50 Deep Network. Comput. Mater. Contin. 2021, 68, 117–135.

[16] Emara, T.; Afify, H.M.; Ismail, F.H.; Hassanien, A.E. A Modified Inception-v4 for Imbalanced Skin Cancer Classification Dataset. In Proceedings of the 2019 14th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 17–18 December 2019; pp. 28–33.

[17] Namozov, A.; Ergashev, D.; Cho, Y.I. Adaptive Activation Functions for Skin Lesion Classification Using Deep Neural Networks. In Proceedings of the 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), Toyama, Japan, 5–8 December 2018; pp. 232–235.

[18] Alkarakatly, T.; Eidhah, S.; Al-Sarawani, M.; Al-Sobhi, A.; Bilal, M. Skin Lesions Identification Using Deep Convolutional Neural Network. In Proceedings of the 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), Medina, Saudi Arabia, 13–15 May 2020; pp. 1–5.

[19] Hosny, K.M.; Kassem, M.A.; Foaud, M.M. Skin Cancer Classification using Deep Learning and Transfer Learning. In Proceedings of the 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), Cairo, Egypt, 20–22 December 2018; pp. 90–93.

[20] Hosny, K.M.; Kassem, M.A.; Foaud, M.M. Classification of skin lesions using transfer learning and augmentation with Alex-net. PLoS ONE2019, 14, e0217293.

[21] Serener, A.; Serte, S. Keratinocyte Carcinoma Detection via Convolutional Neural Networks. In Proceedings of the 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 11–13 October 2019; pp. 1–5.

[22] Iqbal, I.; Younus, M.; Walayat, K.; Kakar, M.U.; Ma, J. Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. Comput. Med. Imaging Graph. 2021, 88, 101843.

[23] Alkarakatly, T.; Eidhah, S.; Al-Sarawani, M.; Al-Sobhi, A.; Bilal, M. Skin lesions identification using deep convolutional neural network. IEEE Access 2020, 8, 264–270.

[24] Kassem, M.A.; Hosny, K.M.; Fouad, M.M. Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. IEEE Access 2020, 8, 114822–114832.

[25] Ahmed, S.A.A.; Yanikoglu, B.; Goksu, O.; Aptoula, E. Skin Lesion Classification with Deep CNN Ensembles. In Proceedings of the 2020 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, 5–7 October 2020; pp. 1–4.

[26] Barata, C.; Marques, J.S. Deep Learning For Skin Cancer Diagnosis with Hierarchical Architectures. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI), Venice, Italy, 8–11 April 2019; pp. 841–845.

[27] Kaymak, S.; Esmaili, P.; Serener, A. Deep Learning for Two-Step Classification of Malignant Pigmented Skin Lesions. In Proceedings of the 14th Symposium on Neural Networks and Applications (NEUREL 2018), Belgrade, Serbia, 20–21 November 2018; pp. 1–6.

[28] Moldovan, D. Transfer Learning Based Method for Two-Step Skin Cancer Images Classification. In Proceedings of the 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 21–23 November 2019; pp. 1–4.

[29] Han, S.S.; Kim, M.S.; Lim, W.; Park, G.H.; Park, I.; Chang, S.E. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. J. Investig. Dermatol. 2018, 138, 1529–1538.

[30] Barata, C.; Celebi, M.E.; Marques, J.S. A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer. IEEE J. Biomed. Health Inform. 2019, 23, 1096–1109.

[31] Hameed,N.; Shabut, A.M.; Ghosh, M.K.; Hossain, M.A. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. Expert Syst. Appl. 2020, 141, 112961.

[32] Thurnhofer-Hemsi, K.; Domínguez, E. A Convolutional Neural Network Framework for Accurate Skin Cancer Detection. Neural Process. Lett. 2021, 53, 3073–3093.

[33] Arora, G.; Dubey, A.K.; Jaffery, Z.A.; Rocha, A. A comparative study of fourteen deep learning networks for multi skin lesion classification (MSLC) on unbalanced data. Neural Comput. Appl. 2022, 35, 7989–8015.

[34] Chen, X.; Tao, H.; Zhou, H.; Zhou, P.; Deng, Y. Hierarchical and progressive learning with key point sensitive loss for sonar image classification. Multimed. Syst. 2024, 30, 380.

[35] Ju, C.; Bibaut, A.; van der Laan, M. The relative performance of ensemble methods with deep convolutional neural networks for image classification. J. Appl. Stat. 2018, 45, 2800–2818.

[36] Chaturvedi, S.S.; Tembhurne, J.V.; Diwan, T. A multi-class skin Cancer classification using deep convolutional neural networks. Multimed. Tools Appl. 2020, 79, 28477–28498.

[37] Gessert, N.; Nielsen, M.; Shaikh, M.; Werner, R.; Schlaefer, A. Skin lesion classification using ensembles of multi-resolution EfficientNet with meta data. MethodsX 2020, 7, 100864.

[38] Raza, R.; Zulfiqar, F.; Tariq, S.; Anwar, G.B.; Sargano, A.B.; Habib, Z. Melanoma classification from dermoscopy images using ensemble of convolutional neural networks. Mathematics 2022, 10, 26.

[39] Kaur, R.; Gholamhosseini, H.; Sinha, R.; Lindén, M. Melanoma Classification Using a Novel Deep Convolutional Neural Network with Dermoscopic Images. Sensors 2022, 22, 1134.

[40] Alzubaidi, L.; Al-Amidie, M.; Al-Asadi, A.; Humaidi, A.J.; Al-Shamma, O.; Fadhel, M.A.; Zhang, J.; Santamaría, J.; Duan, Y. Novel transfer learning approach for medical imaging with limited labeled data. Cancers 2021, 13, 1590.

[41] Datta, S.K.; Shaikh, M.A.; Srihari, S.N.; GAO, M. Soft Attention Improves Skin Cancer Classification Performance. In Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data; Springer: Cham, Switzerland, 2021; pp. 13–23.

[42] Hsu, B.W.Y.; Tseng, V.S. Hierarchy-aware contrastive learning with late fusion for skin lesion classification. Comput. Methods Programs Biomed. 2022, 216, 106666.

[43] Wang,Y.; Feng, Y.; Zhang, L.; Zhou, J.T.; Liu, Y.; Goh, R.S.M.; Zhen, L. Adversarial multimodal fusion with attention mechanism for skin lesion classification using clinical and dermoscopic images. Med. Image Anal. 2022, 81, 102535.

[44] Ajmal, M.; Khan, M.A.; Akram, T.; Alqahtani, A.; Alhaisoni, M.; Armghan, A.; Althubiti, S.A.; Alenezi, F. BF2SkNet: Best deep learning features fusion-assisted framework for multiclass skin lesion classification. Neural Comput. Appl. 2023, 35, 22115–22131.