ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

Melanoma Image Classification Based on Reinforcement Learning and Dynamic Attention Mechanisms

Tongzhu Zhao[1,a], Yu Xiang[1,b,*], Wei Wang[1,c], Yonghao Wu[1,d], Tiancai Zhu[1,e]

[1]School of Information Science and Technology, Yunnan Normal University, Kunming, China
[a]2324100063@ynnu.edu.cn, [b]XIANGYU@ynnu.edu.cn, [c]2324100048@ynnu.edu.cn,
[d]2324100051@ynnu.edu.cn, [e]2324100067@ynnu.edu.cn
**Corresponding author*

Abstract: *Early diagnosis of skin cancer is crucial for improving patient survival rates. Existing deep learning-based classification models predominantly employ fixed, static attention mechanisms, which struggle to adaptively capture subtle features of lesions that vary in scale and morphology. This study proposes a dynamic attention network based on deep reinforcement learning (ResNetRL), aiming to enhance the model's classification performance for melanoma by dynamically regulating multi-scale attention weights in real-time. We embed Dual-Scale Attention Modules (DAS) into the four stages of ResNet50, where the scaling factors for channel attention and spatial attention are dynamically adjusted by the DDPG (Deep Deterministic Policy Gradient) algorithm. A hybrid reward function is innovatively designed, incorporating three metrics: classification accuracy, loss trend, and lesion region stability. Evaluated on the ISIC skin cancer datasets (ISIC 2017 and ISIC 2019), the proposed method achieves classification accuracies of 89.81%, 88.83%, and 91.38%, respectively. This study validates the effectiveness of dynamic attention mechanisms in medical image analysis and provides a novel research approach and methodology for the field.*

Keywords: *Melanoma Classification; Residual Network; Dynamic Attention; Deep Reinforcement Learning; DDPG*

1. Introduction

Melanoma, as the most aggressive type of skin malignancy, requires accurate early classification to significantly improve patient survival rates. However, traditional diagnostic methods heavily rely on physicians' subjective experience, leading to inefficiencies and high rates of misdiagnosis. Data from the International Skin Imaging Collaboration (ISIC) challenge indicates that even the optimal models achieve an average precision (AP) of only 0.691, highlighting the limitations of existing approaches [1]. Within dermatology, the manual diagnosis of melanoma remains a highly challenging task[2].In recent years, deep learning techniques have made significant progress in the field of medical image analysis [3]. Yu et al. [4]achieved a breakthrough in melanoma classification by improving the GoogLeNet architecture and proposing a method based on the feature encoding of local image patches. However, existing methods still face two key challenges: on one hand, merely increasing network depth can lead to gradient vanishing and training difficulties; on the other hand, traditional Convolutional Neural Networks (CNNs) struggle to effectively capture the subtle feature differences of melanoma, such as irregular borders and heterogeneous pigment distribution [5].Zhang et al.[6] designed a deep convolutional neural network model integrating attention mechanisms with residual learning for the automatic classification of skin melanoma. The core architecture of this network includes: multiple stacked attention residual blocks, which enhance the extraction of key features via the attention mechanism; a global average pooling layer to reduce feature dimensionality while preserving spatial information; and a fully connected classification layer for the final lesion classification output. By combining attention mechanisms with residual connections, this model effectively improved the accuracy of melanoma classification.

The main contributions of this paper include:

(1) Proposed an innovative dynamic attention learning paradigm addressing the limitations of existing static attention mechanisms in medical image classification, which struggle to adaptively capture multi-scale and morphologically varying lesion features, this study for the first time deeply integrates deep

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

reinforcement learning with a dynamic attention mechanism, constructing a novel network framework named ResNetRL.

(2) Designed and implemented a dynamically adjustable dual-scale attention module. A self-developed dual-scale attention module is embedded into the four feature stages of the ResNet50 backbone network. This module not only integrates channel and spatial attention but also innovatively introduces scaling fac- tors dynamically adjusted by a reinforcement learning agent, enabling adaptive optimization of attention intensity.

(3) Constructed an end-to-end reinforcement learning control system based on DDPG.A complete stateaction-reward interaction framework is designed.

2. Related Work

2.1 ResNet Network Model

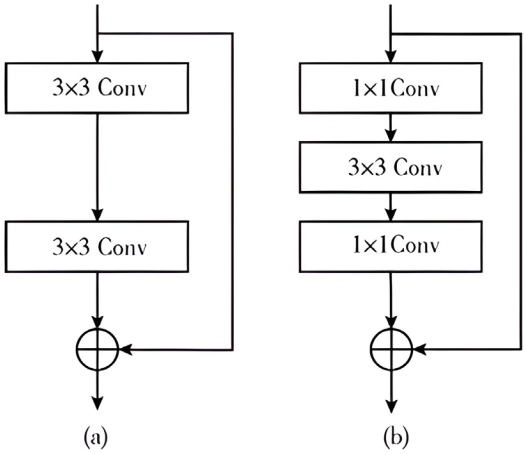

Figure 1: Residual Module

ResNet (Residual Network) addresses the issues of gradient vanishing and gradient explosion in deep network training through the introduction of residual learning. Its core component is the residual block, which employs "skip connections" to add the input to the output, mathematically expressed as $y = f(x) + x$. This design allows gradients to propagate directly through the shortcut connections, effectively alleviating gradient problems and ensuring that performance does not degrade as network depth increases. When the input and output dimensions do not match, a 1×1 convolutional layer is used to adjust the dimensions to ensure correct addition. The basic structure of the residual module is shown in Figure 1.

The residual module illustrated in Figure 1(b) adopts a three-layer convolutional structure, which differs from the traditional two-layer design. Specifically, a 1×1 convolutional layer is introduced, primarily serving to reduce the computational load of the residual module. This design not only optimizes computational efficiency but also maintains the network's depth and performance. Due to the existence of residual structures, ResNet can be designed with considerable depth without encountering gradient vanishing or explosion issues. This deep architecture enables ResNet to excel in handling complex image classification tasks, allowing it to learn richer feature representations.

2.2 Attention Mechanism

Attention Mechanism[7]is a computational technique that mimics the human cognitive focusing process by dynamically assigning weights to highlight critical information in the input data. Its core principle involves calculating the correlation between features, enabling the model to adaptively concentrate on task- relevant portions. Initially achieving breakthroughs in natural language processing (e.g., Transformer), this mechanism is now widely applied in fields such as computer vision and medical image analysis. Its advantages lie in enhancing model interpretability while improving the ability to capture long-range dependencies or subtle features. Channel Attention is an important form of attention mechanism, designed to allow networks to automatically learn the importance of each feature channel, thereby enhancing crucial channels and suppressing the influence of irrelevant or noisy channels. In

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

convolutional neural networks, feature maps consist of multiple channels, each corresponding to different features (e.g., edges, textures, etc.). The goal of channel attention is to dynamically adjust channel weights, enabling the network to focus on task-relevant features and reduce redundant information. SENet [8] is the most classic implementation method for channel attention. Its structure comprises two steps: Squeeze and Excitation. Squeeze: Compresses each channel to produce a channel descriptor vector.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i,j) \tag{1}$$

Among them:x c(i,j) is the activation value of the c-th channel at position (i, j),H×W is the spatial size of the feature map. The operation produces a channel descriptor $z \in \mathbb{R}^C$ (where C is the total number of channels).The excitation step learns inter- channel relationships through a fully connected layer followed by a non-linear activation, generating a channel-wise weight vector s defined as:

$$s = \sigma\big(W_2 \cdot \delta(W_1 \cdot z)\big) \tag{2}$$

Where z denotes the channel descriptor vector obtained from the squeeze operation.W_1 and W_2 represent the weights of two fully connected layers, forming a bottleneck structure to capture non-linear inter-channel dependencies.$\delta$and σ refer to the ReLU and Sigmoid activation functions, respectively. The resulting weight vector s is employed to recalibrate the importance of each feature channel.

The learned weights are multiplied channel-wise with the original feature maps to obtain enhanced features. The spatial attention mechanism is a type of attention mechanism focused on the spatial dimensions of feature maps. By assigning attention weights to each position in the feature map, it enhances critical information and suppresses redundant information, thereby improving the model's ability to perceive important features. It can be implemented in various ways, such as channel-weighted averaging, combining max pooling and average pooling, or dynamic attention based on convolution. The spatial attention mechanism performs well in tasks such as image classification, object detection, and image segmentation, significantly improving model performance and robustness while enhancing interpretability.

2.3 Deep Reinforcement Learning

Reinforcement Learning (RL) is a crucial paradigm in machine learning [9], focusing on how an agent[10] learns an optimal policy through continuous interaction with its environment to maximize cumulative rewards. Its core principle is trial-and-error learning: based on the environmental state, the agent selects an action, receives immediate reward feedback, and through continuous exploration and exploitation, gradually optimizes its policy to achieve long-term objectives [11]. Due to its capability to address sequential decision-making problems, reinforcement learning is regarded as one of the key pathways toward achieving general artificial intelligence. The interactive model of this process is illustrated in the figure, reflecting the dynamic cyclical relationship among state, action, reward, and policy. Deep Reinforcement Learning (DRL) effectively addresses the challenges of discretization and combinatorial explosion faced by traditional methods in high-dimensional continuous state and action spaces by introducing deep neural networks as function approximators within the traditional reinforcement learning framework. Its core innovation lies in: the policy network directly maps high-dimensional states into a probability distribution of actions or continuous parametric outputs; the value network enables accurate evaluation of arbitrary state-action pairs. By integrating the representational power of deep learning [12] with the sequential decision-making mechanism of reinforcement learning, deep reinforcement learning can jointly update policies and value functions using gradient optimization methods (such as Deep Q- Networks and their extensions). This approach demonstrates robust learning capabilities across a variety of complex tasks, ranging from video games to robotic control.

3. Methods

3.1 Overall Model Architecture

In the task of melanoma image classification, this study adopts ResNet50 as the backbone network and introduces improvements upon it. ResNet50 consists of four convolutional stages, comprising 3, 4, 6, and 3 residual blocks, respectively. When the input image size is 224×224 pixels, the output feature maps from each stage have dimensions of 56×56,28×28,14×14,and 7×7 sequentially. This hierarchical downsampling structure facilitates the extraction of multi-scale lesion features. The specific architectural parameters of the ResNet50 network are detailed in table 1.

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

Table 1: ResNet Module Architecture

Block	Output Size	Layer Details
Conv Block 1	56×56	[1 × 1, 64 3 × 3, 64 1 × 1, 256] × 3
Conv Block 2	28×28	[1 × 1, 128 3 × 3, 128 1 × 1, 512] × 4
Conv Block 3	14×14	[1 × 1, 256 3 × 3, 256 1 × 1, 1024] × 6
Conv Block 4	7×7	[1 × 1, 512 3 × 3, 512 1 × 1, 2048] × 3

The overall architecture of the model proposed in this paper is based on a pre-trained ResNet50 as the backbone network. The convolutional layers, batch normalization layers, activation layers, and max pooling layers of the original ResNet50 constitute the front-end feature extractor. At the output ends of its four core residual blocks (layer1 to layer4), we have inserted DAS modules respectively. These modules integrate channel attention and spatial attention mechanisms, with their attention weights dynamically adjusted through reinforcement learning, thereby adaptively enhancing the multi-scale features extracted at each stage. Finally, the enhanced features are aggregated via a global average pooling layer and mapped to the target class space by a fully connected classification layer, completing the binary classification task for melanoma. The model structure is shown in Figure 2:

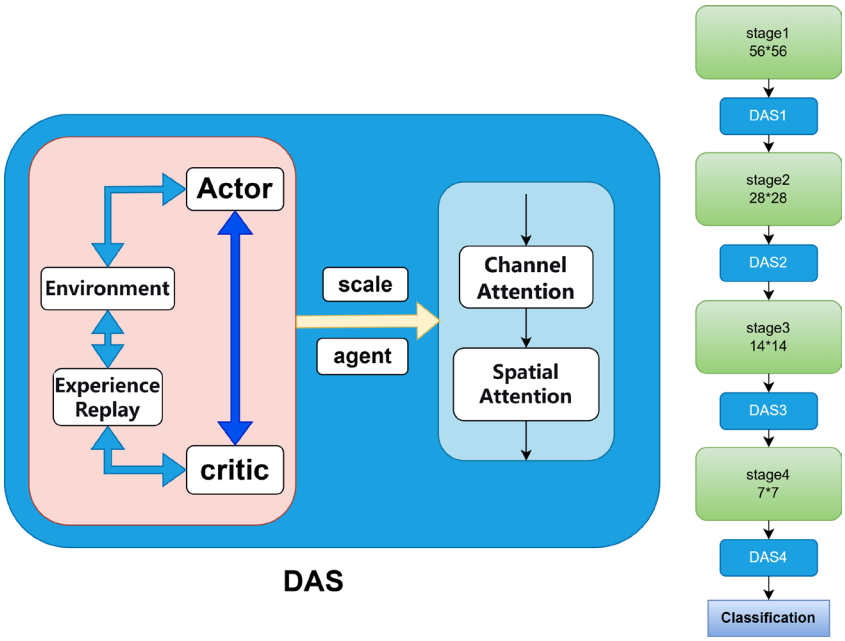

Figure 2: Model Architecture

3.2 DAS Module

The DAS module proposed in this study integrates channel attention and spatial attention, and innovatively introduces dual-scale scaling parameters dynamically regulated by a reinforcement learning agent. The module consists of two core components: a channel attention sub-module based on the Squeeze-and-Excitation structure, and a spatial attention sub-module based on spatial pooling fusion. Both sub-modules employ residual connection designs, and their respective enhancement processes can be expressed as fol- lows:

$$x_{\text{channel}} = x \times (1 + \text{channel\_att} \times \text{channel\_scale}) \tag{3}$$

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

$$x_{\text{spatial}} = x \times (1 + \text{spatial\_att} \times \text{spatial\_scale}) \tag{4}$$

Where channel_att and spatial_att represent the channel and spatial attention weights, respectively. channel_scale and spatial_scale are learnable scaling factors constrained by the tanh activation function[13]. These two scaling factors are initialized with a distribution of (-0.5, 0.5) and are not updated through conventional backpropagation during training. Instead, they are dynamically adjusted by the action vectors generated by the DDPG algorithm, enabling adaptive optimization of the attention mechanism's intensity.

3.3 DDPG Dynamic Regulation Mechanism

To dynamically optimize the scaling parameters within the DAS module, this study designs a reinforcement learning framework based on the Deep Deterministic Policy Gradient (DDPG) algorithm [14]. Through continuous interaction between the agent and the training environment, this framework achieves fine-grained regulation of eight scaling factors (corresponding to four DAS modules).

State Representation: The state observed by the agent is a 20-dimensional vector, comprising four components:(1) Model performance metrics: sliding window accuracy and loss calculated based on a 100- sample window; (2) Training progress information: the proportion of completed epochs and batches; (3) Attention parameters: the current values of the eight scaling factors from the four DAS modules; (4) Trend indicators: the most recent four change trend values for accuracy and loss. Action Space and Network Output: The action space is designed as an 8-dimensional continuous vector, where each dimension corresponds to an adjustment command for a scaling factor. This action vector is generated by the Actor policy network. The Actor network takes the 20-dimensional state as input, processes it through a shared feature extractor consisting of two fully connected layers (each with 128 neurons and using the ReLU activation function), and finally generates raw action values via a linear output layer. To ensure the numerical range of the output actions is controllable, a tanh activation function is applied after this output layer, constraining each dimension's action value to the interval -1,1. Positive values indicate a recommendation to enhance the corresponding scaling factor, while negative values suggest suppression. These values directly correspond to the adjustment direction and magnitude for each scaling factor. Parameter Update Strategy: When executing actions, a fine-grained incremental update strategy is employed:

$$\text{scale}^{(i)}\text{new} = \text{clip}\left(\text{scale}^{(i)}\text{old} + \alpha \cdot \text{a}_{\text{t}}^{(i)},\ -0.5,\ 0.5\right) \tag{5}$$

Where,$\text{a}_{\text{t}}^{(i)}$is the i-th dimensional component of the action $a\,t$, $\text{alpha} = 0.0001$ is the fine-tuning coefficient. This design ensures the smoothness of parameter variations and prevents training instability.Function $\text{clip}(\cdot, -0.5, 0.5)$constrains the final value of the scaling factor to the hard limit of [-0.5, 0.5], preventing it from excessively deviating from the initial distribution. Reward Function: A multi-objective hybrid reward function is designed, comprehensively considering immediate performance, long-term optimization, and training stability:

$$\begin{aligned} R_t = \quad & w_1 \cdot R_{\text{immediate}} + w_2 \cdot R_{\text{discounted}} + w_3 \cdot R_{\text{long-term}} \\ & -\lambda \sum |\ \text{scale}\ | + R_{\text{trend}} \end{aligned} \tag{6}$$

Specifically, the immediate reward encourages an increase in accuracy (weight 1.2) and a decrease in loss (weight 0.1) for the current batch. The discounted reward applies a discount factor (gamma = 0.99) to future rewards, encouraging the agent to plan for long-term benefits. The long-term breakthrough reward is dynamically weighted and granted when model performance surpasses the historical best, based on the extent of improvement; Parameter Stability Penalty($-\lambda\sum|\{scale\}|$)A stability penalty term with lambda = 0.1 is applied to suppress severe parameter fluctuations; the trend reward additionally encourages sustained improvement in the trends of both accuracy and loss.

Network Architecture and Training Strategy: A standard Actor-Critic architecture is adopted, where both the Actor and Critic networks consist of two fully connected layers with 128 dimensions each [15]. Experience replay employs a prioritized mechanism with a buffer capacity of 20,000, a priority exponent alpha = 0.6, and an importance sampling exponent beta that gradually increases from 0.4 to 1.0. The exploration strategy uses an epsilon-greedy approach, with epsilon decaying from 1.0 to 0.01 at a rate of 0.995. During training, attention parameters are updated every 5 batches, while the DDPG network is updated per batch using 32 samples, achieving synergistic optimization between classification training and reinforcement learning.

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

4. Experiment

4.1 Experimental Setup and Data

The experiments in this study were conducted on a Windows 10 operating system, utilizing an NVIDIA GeForce RTX 3060 GPU with 12GB of memory and 32GB of RAM. The datasets used in the experiments are from the skin disease database provided by the International Skin Imaging Collaboration (ISIC). The specific datasets used—the ISIC dataset, ISIC 2017, and ISIC 2019—are detailed in Table 2.

Table 2: Datasets

Dataset	Isic	Isic2017	Isic2019
Training Set Melanoma (Images)	438	348	3639
Training Set Non-Melanoma (Images)	1739	1652	16625
Test Set Melanoma (Images)	80	117	883
Test Set Non-Melanoma (Images)	136	438	4184

The data within each set can be classified into two major categories: melanoma and non-melanoma. In the experiments, melanoma is designated as the positive class.

4.2 Data Preprocessing

To ensure the consistency and quality of the input data, this study systematically preprocesses the image data. During preprocessing, all dermoscopic images are uniformly resized to 224×224 pixels to meet the standard input requirements of the backbone network ResNet50. To enhance data diversity and improve model generalization, random horizontal flips, brightness and contrast adjustments (±20%), and translations of up to 10% are applied to augment the data. Subsequently, the images are converted into PyTorch tensors and normalized to the [0, 1] range. Finally, standardization is performed using the mean and standard deviation from the ImageNet dataset to make the input data distribution more suitable for neural network training.

4.3 Algorithm Evaluation Metrics

This paper evaluates the classification results using Accuracy (ACC) and the Area Under the ROC Curve (AUC), calculated as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{7}$$

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}\left(\mathrm{FPR}^{-1}(u)\right) du \tag{8}$$

Where: TP and FP represent True Positives and False Positives, respectively. TN and FN represent True Negatives and False Negatives, respectively. A larger value for AUC and ACC indicates better classification performance.

4.4 Analysis of Experimental Results

This model employs a transfer learning strategy for training, using a ResNet50 backbone pre-trained on the ImageNet dataset. The top fully connected classification layer is replaced to adapt to the binary classification task in this study. The main training hyperparameters are set as follows: the optimizer is Adam with an initial learning rate of 0.001, the total number of training epochs is 100, and the batch size is 32. To ensure reproducibility, the random seed is set to 42, and the cross-entropy loss function is used as the loss function, which is suitable for multi-class classification tasks. To systematically validate the effectiveness of the proposed dynamic attention mechanism, we designed multiple ablation experiments. First, to investigate the impact of the attention module insertion position and multi-layer attention combinations on performance, we inserted CBAM attention modules after different layers of ResNet50 and compared them with a baseline variant where CBAM was inserted after all four layers. Specifically, CBAMx indicates that the CBAM module is inserted only after the x-th convolutional stage (layerx) of ResNet50, while CBAM indicates that CBAM modules are inserted after all four stages. The experimental results are shown in Table 3:

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

Table 3: Experimental Results

Model	Isic		2017		2019	
	acc	auc	acc	auc	acc	auc
Resnet50	83.80	86.71	84.00	82.69	84.09	83.34
CBAM1	84.26	90.55	84.5	83.45	86.86	88.43
CBAM2	86.57	90.80	86.83	85.10	87.53	88.11
CBAM3	87.04	91.03	87.00	84.37	87.11	88.46
CBAM4	86.11	91.96	87.67	84.24	87.72	88.43
CBAM	87.04	91.44	87.33	83.73	88.36	88.46
Ours	89.81	92.38	88.83	88.73	91.38	91.81

The experimental results indicate that introducing CBAM attention modules at various layers of ResNet50 consistently improves performance. All CBAM variants significantly outperform the original ResNet50 baseline across major metrics, with particularly notable gains in AUC. However, analysis of the impact of the attention module insertion position reveals that there is no universally optimal single insertion layer.The optimal insertion strategy varies depending on the dataset and evaluation metric: on the ISIC dataset, deep-layer insertion (CBAM4) performs best in capturing discriminative features, achieving the highest AUC (91.96%); whereas on the ISIC 2019 dataset, the full-layer insertion strategy (CBAM) yields the best overall performance (accuracy 88.36%, AUC 88.46%). This demonstrates that while static attention mechanisms can provide performance gains, their effectiveness is highly dependent on the tuning of insertion positions, and performance improvements have inherent limitations. The ResNetRL model proposed in this study circumvents the cumbersome search for insertion strategies through its innovative dynamic attention regulation mechanism. As shown in Table 3, this method achieves leading performance across all six evaluation metrics on three datasets. To further validate the general applicability of the proposed framework, we applied its core principles to various backbone networks, including the lightweight ResNet18[16], the deeper ResNet101 [17], the cardinality-based ResNeXt50[18], and the multi- scale feature-based Res2Net50. The experimental results show that our method consistently outperforms its corresponding baseline models and the optimal CBAM variants on each backbone network. This strongly demonstrates that our proposed approach is a universal, robust dynamic attention enhancement framework whose effectiveness does not rely on specific network architectures. It provides a widely applicable performance improvement solution for medical image classification tasks. The experimental results are shown in Table 4:

Table 4: Results of Other Models

Model		Isic		2017		2019	
		acc	auc	acc	auc	acc	auc
Resnet18	Baseline	85.65	89.17	84.17	85.15	87.23	88.66
	BestCBAMVariant	88.43	92.03	88.00	87.52	88.83	90.15
	Ours	91.2	93.43	89	87.54	91.32	91.02
ResNeXt50	Baseline	85.19	86.3	82.83	76.16	87.41	88.25
	BestCBAMVariant	89.81	88.85	85.83	85.95	90.55	92.23
	Ours	91.67	91.49	88.33	85.88	91.2	93.02
Res2Net50	Baseline	84.72	88.91	84.5	84.6	86.86	87.27
	BestCBAMVariant	88.43	92.9	86.5	87.71	91.04	92.64
	Ours	90.28	93.17	90.17	88.05	91.34	92.80
ResNet101	Baseline	86.57	88.65	86.00	84.20	86.4	86.49
	BestCBAMVariant	89.81	90.8.	87.5	87.04	90.17	90.69
	Ours	90.28	89.41	89.67	87.12	91.02	91.83

5. Conclusion

To address the challenges of subtle lesion features and low classification accuracy in the binary

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

classification task of melanoma images, this paper proposes a dynamic attention network based on deep reinforcement learning (ResNetRL). By embedding dual-scale attention modules into the four stages of ResNet50 and innovatively employing the DDPG algorithm to dynamically regulate the scaling factors of channel and spatial attention, the model achieves adaptive enhancement of multi-scale features. The reinforcement learning agent interacts with the environment to autonomously make decisions and optimize the distribution of attention weights, while the dynamic attention mechanism effectively focuses on critical lesion regions and suppresses irrelevant background interference. Experimental validation on the ISIC series of datasets demonstrates that the proposed model outperforms traditional methods in melanoma image classification, significantly improving classification accuracy. It is hoped that the method proposed in this paper will provide new technical insights for the intelligent classification of melanoma images and serve as a reference for the further application of dynamic structural learning in the field of medical image analysis.

References

[1] Kong, X.: Research on Melanoma Classification Based on Deep Learning [D]. Zhejiang University of Technology, 2019. DOI: 10.27463/d.cnki.gzgyu.2019.000407.

[2] Ruffano, L.F., Takwoingi, Y., Dines, J., et al.: Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults [J]. Cochrane Database of Systematic Reviews, 12: CD012526 (2018).

[3] Shen, D., Wu, G., Suk, H.I.: Deep learning in medical image analysis [J]. Annual Review of Biomedical Engineering, 19: 221–248 (2017).

[4] Yu, Z., Ni, D., Chen, S., et al.: Hybrid dermoscopy image classification framework based on deep convolutional neural network and Fisher vector [C]. In Proceedings of the IEEE 14th International Symposium on Biomedical Imaging, Melbourne: IEEE, pp. 301–304 (2017).

[5] Tian, D., Li, F., Zhang, X., Zhang, J.: Construction and Application of a Classification Model Combining Feature Selection and Deep Reinforcement Learning [J]. Electronic Design Engineering, 30(12): 93–97 (2022).

[6] Zhang, J., Xie, Y., Xia, Y., et al.: Attention residual learning for skin lesion classification [J]. IEEE Transactions on Medical Imaging, 38(9): 2092–2103 (2019).

[7] Lin, G., Gong, T., Lin, T.: A Survey on Attention Mechanisms in Image Classification Tasks [J]. Telecommunication Engineering, 65(08): 1349–1362 (2025). DOI: 10.20079/j.issn.1001-893x.241105001.

[8] Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks [C]. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, pp. 7132–7141 (2018). DOI: 10.1109/CVPR.2018.00745.

[9] Thrun, S., Littman, M.L.: Reinforcement Learning: An Introduction [J]. IEEE Transactions on Neural Networks, 16(1): 285–286 (2005).

[10] Busoniu, L., Babuska, R., De Schutter, B., Bart, P.: A comprehensive survey of multiagent reinforcement learning [J]. IEEE Transactions on Systems, Man, and Cybernetics Part C (Applications and Reviews), 38: 156–172 (2008).

[11] Chen, Z., Liu, Z., Wan, L., et al.: A Survey on Multi-Agent Reinforcement Learning Theory and Appli-cations [J]. Pattern Recognition and Artificial Intelligence, 37(10): 851–872 (2024). DOI: 10.16451/j.cnki.issn1003- 6059.202410001.

[12] Meng, X., Li, J., Yu, C., et al.: A Review of Deep Learning in Skin Lesion Image Segmentation [J/OL]. Journal of Frontiers of Computer Science and Technology, 1–27 (2025-12-05).

[13] Rao, A., M.V.M.K., Moonesar, I.A., Atalla, S., Prashanth, B.S., Joshi, G., Soni, T.K., Le, T., Verma, A., Marashdeh, H.: Cross Country Determinants of Investors' Sentiments Prediction in Emerging Markets Using ANN [J]. Frontiers in Artificial Intelligence, 5: 912403 (2022). DOI: 10.3389/frai.2022.912403.

[14] Zhou, S.K., Le, H.N., Luu, K., Nguyen, H.V., Ayache, N.: Deep reinforcement learning in medical imaging: A literature review [J]. Medical Image Analysis, 73: 102193 (2021). DOI: 10.1016/j.media.2021.102193.

[15] Wu, M., Gao, Y., Zheng, J., Zhang, Q., Du, S.: Actor-Dueling-Critic Methods in Reinforcement Learn- ing [J]. Sensors (Basel), 19(7): 1547 (2019). DOI: 10.3390/s19071547.

[16] Liu, Y., Wu, P., Liu, W.: Research on Intelligent Diagnosis of Knee Joint Fracture Medical Images Based on ResNet18 and MSC Attention Mechanism [J]. Software Engineering, 28(11): 16–21 (2025). DOI: 10.19644/j.cnki.issn2096-1472.2025.011.004.

[17] Li, Y., Zhao, P., Zhang, X., et al.: MRI Radiomics Combined with ResNet101 Deep Learning for Dis- criminating Lumbar Brucellar Spondylitis and Spinal Metastatic Cancer [J]. Chinese Journal of

ISSN 2616-5775 Vol. 9, Issue 1: 24-32, DOI: 10.25236/AJCIS.2026.090103

Medical Imaging Technology, 41(06): 958–962 (2025). DOI: 10.13929/j.issn.1003-3289.2025.06.023.
[18] Cheng, C., Wang, Y., Song, Q., et al.: A Lightweight ECA-ResNeXt for Smart Agriculture and Its Application in Rice Disease Recognition [J]. Agricultural Engineering, 15(07): 28–35 (2025). DOI: 10.19998/j.cnki.2095-1795.202507304.