Dean&Francis
ISSN 2959-6157

Research on Skin Cancer Classification Based on an Optimized AlexNet Model

Moran Lin*

Department of Electronic Information Engineering, Hebei University, Baoding, 071002, China
*Corresponding author: moranlinn@outlook.com

Abstract:

Skin cancer is one of the most common types of tumors worldwide, and its early detection and diagnosis are essential for enhancing treatment efficacy and patient prognosis. Nevertheless, conventional pathological examination techniques are frequently time-intensive and inefficient, thereby constraining their applicability in large-scale clinical screening. To overcome these limitations, the current study introduces an enhanced AlexNet model grounded in deep learning, incorporating a dynamic learning rate adjustment strategy to refine the training process. By extracting discriminative features from images of benign and malignant skin cancers, the model was trained and optimized to attain superior performance. The finalized model demonstrated a classification accuracy of 85% on the skin cancer test dataset sourced from the Kaggle dataset, alongside an area under the ROC curve (AUC) of 0.9321, reflecting a robust capability to distinguish between benign and malignant lesions. The optimized AlexNet-based model developed in this study effectively achieves the classification of benign and malignant skin lesions, offering meaningful support for clinical diagnosis and decision-making.

Keywords: Deep learning; skin cancer; AlexNet model.

1. Introduction

Skin cancer refers to tumors that occur on the skin, with the most frequently seen varieties being basal cell carcinoma, squamous cell carcinoma, melanoma, and non-melanoma. The primary cause is prolonged exposure of the skin to sunlight, as strong ultraviolet rays can cause harm to the DNA within skin cells, which can eventually trigger the development of skin cancer. Globally, in 2022, approximately 58,667 people died from melanoma and 69,416 from non-melanoma skin cancers, making skin cancer one of the most life-threatening diseases to humanity [1]. Clinically, skin tumors are primarily categorized as benign or malignant. Benign skin tumors usually have clear boundaries and regular edges, pose relatively low harm to human health, and generally show good treatment and prognosis outcomes. In contrast, malignant skin tumors are prone to cancer cell metastasis, which threatens life, with surgery and radiotherapy being the primary treatment methods. Accurate and rapid classification of skin cancer is essential for timely clinical intervention, which can

ISSN 2959-6157

significantly reduce mortality rates; therefore, early detection of skin tumors has become critical [2]. Currently, the "gold standard" for distinguishing between benign and malignant skin cancers is pathological examination of tissue sections, which is relatively slow in terms of diagnostic efficiency.

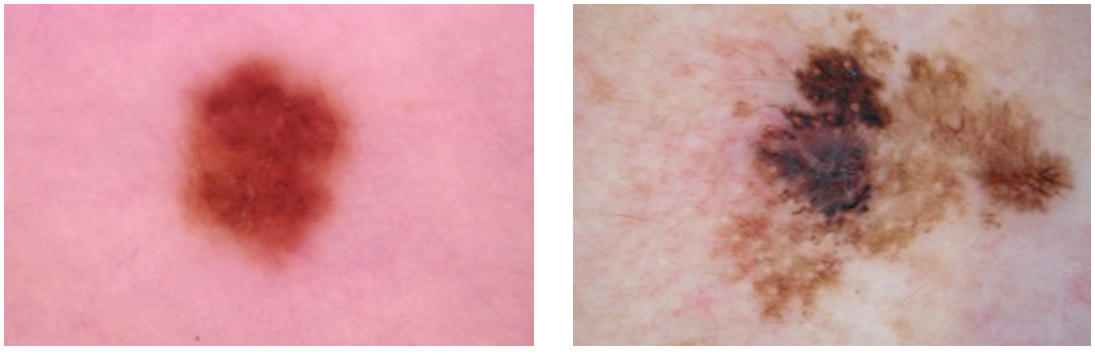
As modern science and diagnostic medicine continue to advance rapidly, deep learning methods are increasingly being integrated with clinical diagnostic techniques and applied to disease detection. Many researchers have employed deep learning to achieve disease detection and diagnosis. Wang et al. utilized the DenseNet-121 model to classify pneumonia CT images, achieving high classification performance in accurately distinguishing normal from pneumonia cases [3]. Zhu et al. developed a diagnostic model for normal and abnormal electrocardiograms (ECGs) by combining three convolutional neural networks (CNNs) to achieve adaptive feature learning [4]. Chen et al. proposed an optimized AlexNet model for classifying femoral bone tumor images based on CNN, and the results confirmed the model's excellent performance in terms of sensitivity and specificity [5]. Guo et al. evaluated the performance of AlexNet and ResNet in classifying malignant spinal bone tumors, with both models showing good classification accuracy [6]. Alsaade et al. constructed an artificial neural network for melanoma classification and detection, achieving outstanding accuracy on the PH2 and ISIC 2018 datasets [7]. Xiao et al. proposed and tested a weakly supervised deep learning framework capable of predicting the likelihood of recurrence in individuals diagnosed with non-metastatic colorectal cancer confirmed by histopathology, achieving an AUC of 0.715 in the testing phase [8]. Nakagawa et al. applied ResNet50 to CT images to detect skull base invasion in nasopharyngeal malignancies, achieving a diagnostic accuracy of 97.3% [9]. Gunasekaran et al. employed AlexNet for feature extraction and ResNet101 for segmentation of brain tumor images, thereby assisting in the detection and treatment of brain tumors [10].

Based on this, this study employs the AlexNet deep learning model to achieve automatic classification of benign and malignant skin cancer. By training on labeled skin cancer images, the model learns the image features and is then applied to predict the malignancy of unknown cases. This paper first introduces the source of the skin cancer dataset and the data preprocessing methods. The data are divided into training, validation, and testing sets. Using Python and the PyTorch framework, an AlexNet-based deep learning model is constructed and trained on the skin cancer dataset. The trained model is subsequently tested, and the results are evaluated by calculating the confusion matrix, classification accuracy, and related performance metrics. Finally, a comprehensive summary of the study is presented, highlighting the research findings and their practical value. In conclusion, the proposed approach effectively improves the efficiency of skin cancer detection and provides a useful reference for clinical diagnosis.

2. Model Construction Method

2.1 Experimental Environment Setup and Data Preprocessing

The experimental environment was configured using Python 3.8, with PyCharm as the integrated development environment, and PyTorch 1.3.1 as the deep learning framework. The dataset employed in this study was obtained from the Kaggle database, consisting of 3,297 labeled images of benign and malignant skin cancer. Two representative examples from the Kaggle dataset are illustrated in Fig. 1.



(a) **(b)**

Fig. 1 Sample images from the skin cancer dataset. (a) Benign; (b) Malignant

The main steps of image data preprocessing are as follows. The dataset was first partitioned into training and testing sets at a proportion of 8 to 2, and 20% of the training set was further separated as a validation set during

training. Next, the training set was enhanced through data augmentation to improve image diversity and enhance model robustness under varying imaging conditions. The augmentation strategies included random horizontal flipping, random rotation (±15°), and random adjustment of brightness, contrast, and saturation (±10%). After augmentation, the images were standardized by linearly mapping pixel values to the range [−1, 1], which helps accelerate model convergence during training. Finally, all input images were converted into standardized tensor formats before being fed into the deep neural network, providing a consistent data input for subsequent training and testing of the deep learning model.

2.2 Implementation of the Deep Learning Framework

The network adopts the classical AlexNet architecture, which includes five convolutional layers and three fully connected layers, featuring interspersed max-pooling layers and ReLU activation functions. The basic network framework is shown in Fig. 2. The ReLU function is defined such that if the input x is greater than 0, the output is x; otherwise, the output is 0. Its primary role is to enhance the network's nonlinear representation capability and effectively mitigate the vanishing gradient problem. Fully connected layers are incorporated with a Dropout layer. During the forward propagation of training, each neuron's activation is randomly set to zero with a probability of 50%, preventing the model from relying too heavily on local features of specific images. This improves the model's generalization ability and helps prevent overfitting.

To enable automatic learning rate adjustment, the Adam optimizer was introduced into the deep learning network with an initial learning rate of 1e-4. Adam utilizes adaptive moment estimation to compute both the first-order (mean) and second-order (uncentered variance) moments of the gradient, thereby assigning adaptive learning rates to different parameters. Meanwhile, to further enhance the stability of model training and prevent overfitting, a dynamic learning rate adjustment strategy based on validation performance was introduced. Throughout the training phase, in cases where validation accuracy remains unchanged over three successive epochs, the existing learning rate would be automatically decreased to fifty percent of its former value. This feature allows the model to more efficiently target the best solution and speeds up the convergence during the training phase.

The input images were 224×224 three-channel color images. In the first convolutional layer (kernel size 11×11, stride 4, output channels 96), the large kernel significantly reduces spatial resolution while capturing high-level features. Subsequently, a ReLU activation function is applied, followed by a 3×3 max-pooling layer (stride 2) to further compress the feature maps. In the second convolutional layer (kernel size 5×5, stride 1, padding 2, output channels 256), smaller kernels focus on fine-grained features. After the ReLU activation and pooling operations, the feature maps' spatial size is compressed to 12 × 12. The third to fifth convolutional layers are designed with 3×3 kernels (stride 1) and output channels of 384, 384, and 256, respectively, facilitating the capture of refined and complex features.

After the fifth convolutional layer, a pooling operation reduces the feature map size to 5×5 while retaining 256 feature maps across the channels. After the convolutional processing, the feature maps are transformed into a one-dimensional vector of length 6400, which is subsequently input into two fully connected layers, each comprising 4096 nodes. ReLU activation functions are employed in both fully connected layers, along with a 50% dropout rate to mitigate overfitting. Finally, a softmax function is applied for classification, generating the output predictions. The model's performance is evaluated by comparing predicted labels derived from these predictions with ground truth labels in the test set, thereby facilitating the computation of classification accuracy and related evaluation metrics.

ISSN 2959-6157

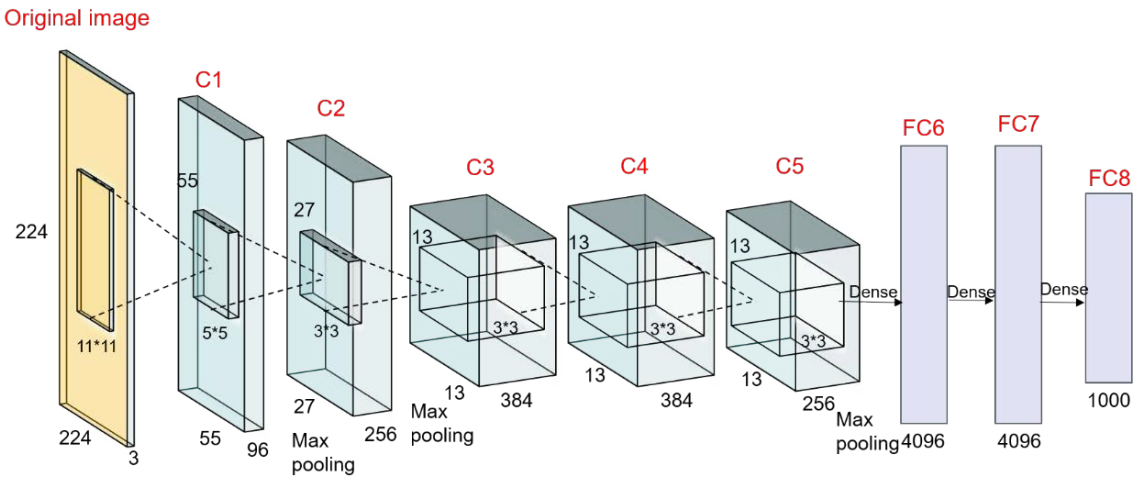


Fig. 2 Architecture of the AlexNet network

3. Experimental Results

On the test dataset, each prediction produced by the trained model is assigned to one of four categories—true positive, true negative, false positive, or false negative—depending on whether the sample is correctly or incorrectly classified as positive or negative.

To gauge the effectiveness of the trained model on the dataset, the primary metrics are classification Accuracy, Precision, Recall, and F1 score, which are calculated as outlined in Equations (1)-(4). Classifier performance is assessed using the receiver operating characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR) for a range of threshold values, as defined in Equations (5) and (6). The area under the ROC curve (AUC) represents the probability that positive samples are correctly classified. As the AUC value approaches 1, the model demonstrates superior discriminative capability and improved classification accuracy.

$$Accuracy = \frac{TP + TN}{Support} \tag{1}$$

$$\Pr ecision = \frac{TP}{TP + FP} \tag{2}$$

$$\Re call = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1 } score = \frac{2 * \Pr ecision * \Re call}{\Pr ecision + \Re call} \tag{4}$$

$$TPR = \frac{TP}{TP + FP} \tag{5}$$

$$FPR = \frac{FP}{FP + TN} \tag{6}$$

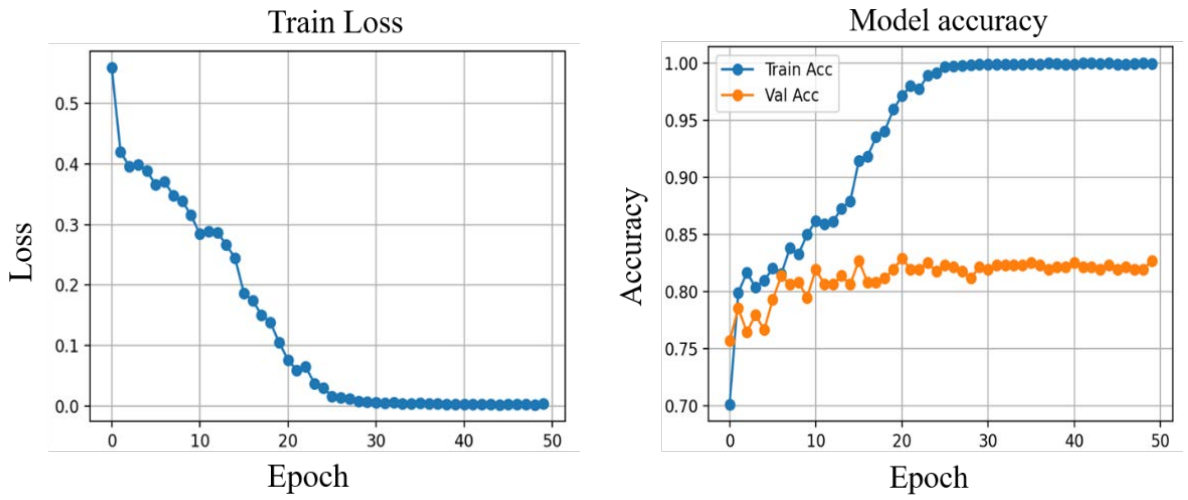
Among these, as presented in Table 1, *Support* refers to the total number of test instances, with *TP*, *TN*, *FP*, and *FN* representing the quantities of true positives, true negatives, false positives, and false negatives, respectively.

Table 1. Classification of Prediction Results

Support	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

The model was trained using the training set for 50 epochs, saving the best-performing model after each epoch. The model weights were updated iteratively to obtain the model with the optimal training performance, which was then used to make predictions on the test set.

The training parameters are illustrated in Fig. 3. Fig. 3(a) shows the training loss gradually decreasing with increasing epochs and approaching convergence after approximately 35 epochs. Fig. 3(b) illustrates the changes in training and validation accuracy over 50 epochs. The training accuracy continuously increased and eventually approached 100%, while the validation accuracy rapidly improved in the initial phase and then stabilized at approximately 82%. The results imply that the model learned the training data effectively; however, the limited enhancement observed on the validation set points to a certain degree of overfitting.



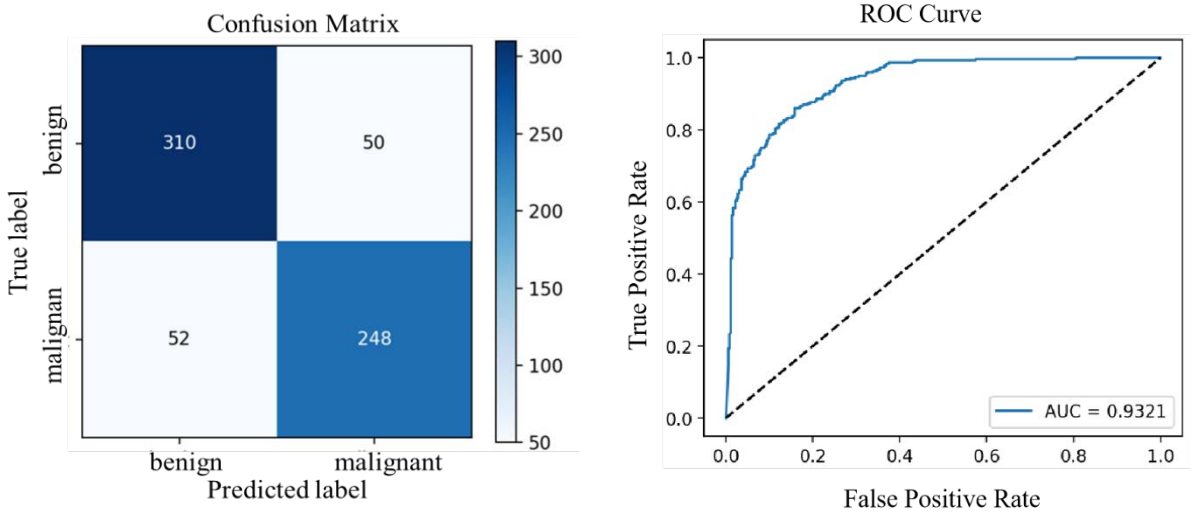
(a) (b)

Fig. 3 Training performance of the model. (a) Training loss; (b) Training and validation accuracy

During evaluation, the best-trained model yielded a prediction accuracy of 85% on the test dataset. The confusion matrix and ROC curve are shown in Fig. 4(a) and Fig. 4(b), respectively, with an AUC of 0.9321. Specifically, for the benign class, the precision, recall, and F1 score were 86%, 86%, and 0.86, respectively; for the malignant class, the corresponding values were 83%, 83%, and 0.83 (refer to Table 2). Overall, the model demonstrates strong classification performance on the test set.

Table 2. Classification Test Results

Classification	Precision	Recall	F1 score
benign	86%	86%	86%
malignant	83%	83%	83%



(a) (b)

Fig. 4 Testing performance of the model. (a) Confusion matrix; (b) ROC curve

ISSN 2959-6157

4. Summary

Skin cancer ranks among the most frequently occurring cancers globally, and early diagnosis and treatment have a decisive impact on the cure rate; delayed intervention may result in severe or even fatal outcomes. The purpose of this Research is to improve the efficiency of skin cancer diagnosis by proposing an enhanced AlexNet-based model for categorizing benign and malignant skin cancers. Data quality was improved through preprocessing techniques, including data augmentation, and the AlexNet model was utilized to extract deep features. An adaptive learning rate adjustment mechanism was incorporated during training to achieve effective classification of skin cancer images, providing support for early clinical diagnosis. Experimental results demonstrate that the model reached an accuracy level of 85% and an AUC of 0.9321 on the test set of the Kaggle skin cancer dataset, indicating strong classification performance and a potential reference value for clinical applications.

References

[1] Wang Mingyue, Gao Xiaohua, Zhang Li. Recent global patterns in skin cancer incidence, mortality, and prevalence. Chinese Medical Journal, 2025, 138(2): 185–192.

[2] Ahmed B, Qadir M, Ghafoor S. Malignant Melanoma: Skin Cancer-Diagnosis, Prevention, and Treatment. Critical Reviews in Eukaryotic Gene Expression, 2020, 30(4): 291-297.

[3] Wang Weide, Wan Nan, Wang Xin, et al. Research on the Application of Deep Learning in Pneumonia CT Image Classification and Recognition. Journal of Fujian Computer, 2024, 40(7): 33-36.

[4] Zhu Junjiang, Lv Jintao, Kong Dongdong. CNN-FWS: A Model for the Diagnosis of Normal and Abnormal ECG with Feature Adaptive. Entropy (Basel), 2022, 24(4): Art. no. 471.

[5] Chen Xu, Chen Hongkun, Wan Junming, et al. An enhanced AlexNet-Based model for femoral bone tumor classification and diagnosis using magnetic resonance imaging. Journal of Bone Oncology, 2024, 48: Art. no. 100626.

[6] Guo Chengquan, Chen Yan, Li Jianjun. Radiographic imaging and diagnosis of spinal bone tumors: AlexNet and ResNet for the classification of tumor malignancy. Journal of Bone Oncology, 2024,48: Art. no. 100629.

[7] Alsaade F W, Aldhyani T H H, Al-Adhaileh M H. Developing a Recognition System for Diagnosing Melanoma Skin Lesions Using Artificial Intelligence Algorithms. Computational and mathematical methods in medicine, 2021, 2021: Art. no.9998379.

[8] Xiao Han, Weng Zongpeng, Sun Kaiyu, et al. Predicting 5-year recurrence risk in colorectal cancer: development and validation of a histology-based deep learning approach. British journal of cancer, 2024, 130(6): 951–960.

[9] Nakagawa J, Fujima N, Hirata K, Harada T, et al. Diagnosis of skull-base invasion by nasopharyngeal tumors on CT with a deep-learning approach. Japanese journal of radiology, 2024,42(5): 450–459.

[10] Gunasekaran S, Bai M P S, Mathivanan S K, et al. Automated brain tumor diagnostics: Empowering neuro-oncology with deep learning-based MRI image analysis. PloS one, 2024,19(8): Art. no. 0306493.