doi:10.6041/j.issn.1000-1298.2025.09.053

EL－DenseNet: Mushroom Recognition Based on Erasing Module Using DenseNet

WANG Yaojun　ZHAO Weiting　BIE Yuhui　JIA Lu
(*College of Engineering*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Target occlusion poses a significant challenge in computer vision, particularly in agricultural applications, where occlusion of crops can obscure key features and impair the model's recognition performance. To address this challenge, a mushroom recognition method was proposed based on an erase module integrated into the EL－DenseNet model. EL－DenseNet, an extension of DenseNet, incorporated an erase attention module designed to enhance sensitivity to visible features. The erase module helped eliminate complex backgrounds and irrelevant information, allowing the mushroom body to be preserved and increasing recognition accuracy in cluttered environments. Considering the difficulty in distinguishing similar mushroom species, label smoothing regularization was employed to mitigate mislabeling errors that commonly arose from human observers. This strategy converted hard labels into soft labels during training, reducing the model's overreliance on noisy labels and improving its generalization ability. Experimental results showed that the proposed EL－DenseNet, when combined with transfer learning, achieved a recognition accuracy of 96.7% for mushrooms in occluded and complex backgrounds. Compared with the original DenseNet and other classic models, this approach demonstrated superior accuracy and robustness, providing a promising solution for intelligent mushroom recognition.

Key words: mushroom recognition; erase module; label smoothing; DenseNet

CLC number: TP3　**Document Code**: A　**Article ID**: 1000-1298(2025)09-0628-10　**OSID**:

0 Introduction

Mushrooms belong to the phylum Basidiomycetes, a group of succulent and edible Fungi that typically grow on the ground or in plant substrates such as straw and wood[1-2]. Due to their significant nutritional, medicinal, and economic value, they are highly regarded. The nutritional benefits of edible mushrooms include their low calorie content, high protein levels, and low lipid and cholesterol content, along with a rich supply of essential vitamins. In addition, several bioactive metabolites in mushrooms exhibit various beneficial properties, such as anti-cancer, anti-diabetic, anti-inflammatory, and antiviral effects[3-4].

Early mushroom identification methods primarily included morphological feature analysis, chemical detection, and fungal classification techniques[5]. These methods could benefit from improved professionalism, efficiency, and daily applicability. With the rapid advancement of artificial intelligence and hardware capabilities, deep learning techniques have increasingly contributed to mushroom-related research[6]. Among these, convolutional neural networks (CNNs), a cornerstone of deep learning, can offer high accuracy in visual tasks, making them particularly effective for image classification[7-9]. For instance, FENG et al.[10] proposed a mushroom recognition model based on residual modules, achieving over 90% by effectively capturing detailed texture and structural features. In 2021, ZAHAN et al.[11] collected 8 190 mushroom images spanning 45 categories. After applying image enhancement techniques, they used VGG16, ResNet50, and Inception V3 models for classification. The results showed that the accuracy of these models reached 84.44%, 58.65%, and 88.40%, respectively.

WU et al.[12] performed a classification study on 8 516 mushroom images from the Kaggle dataset, focusing on nine categories. The MobileNet V2 model achieved an accuracy of 81.25% on the validation set,

Received date: 2025－01－23　Accepted date: 2025－02－18
Foundation item: National Key R&D Program (2023YFD1701000)
First author: WANG Yaojun, associate professor. E-mail: wangyaojun@cau.edu.cn

an improvement of nearly 9% over the original data. KHASANAH et al.[13] used InceptionV3, VGG16, and ResNet50 networks for mushroom classification and combined InceptionV3 with contrast-limited adaptive histogram equalization to enhance recognition performance. Their results demonstrated that InceptionV3 achieved the highest accuracy at 88.40%.

The erase module enhanced model robustness and generalization by selecting a region in an image and setting its pixel values to zero, simulating occlusion or partial target absence. CHEN et al.[14] employed residual learning to capture the residual features of side outputs, erasing salient regions within these outputs to help the network focus on missing target parts and finer details, thereby improving resolution and accuracy. HOU et al.[15] introduced SeeNet, a self-erasing network designed to prevent attention from diffusing to irrelevant background regions. This model incorporated two self-erasure strategies that guided the network to prioritize more reliable information, effectively emphasizing the object while minimizing background interference. Experimental results demonstrated that compared with the DCSP segmentation network generated using the same method, SeeNet achieved over a 2% performance improvement on the validation set.

Despite these advancements in mushroom recognition, challenges still remained, particularly in standardizing mushroom datasets and improving recognition performance for occluded mushrooms in complex natural backgrounds. These issues can be summarized as follows:

(1) Traditional methods primarily relied on clear and complete image features. As a result, when mushroom images were partially occluded, most existing algorithms struggled to recognize them due to the loss of crucial visual information.

(2) Moreover, mushrooms exhibited significant variation in their growth habits, and their characteristics changed at different growth stages. This variability complicated model recognition, particularly for species with similar appearances, especially when set against complex backgrounds.

To address these challenges, an improved EL－DenseNet model was proposed to achieve recognition accuracy for mushrooms with occlusion or complex backgrounds. The main contributions of this study were summarized as follows:

(1) The proposed method for mushroom recognition, termed EL－DenseNet, leveraged an erase module integrated with the DenseNet architecture. By incorporating the erase attention module, the model became more sensitive to unobstructed features. This mechanism enhanced the model's ability to focus on visible mushroom features, thereby improving classification accuracy. Additionally, the erase module helped eliminate complex backgrounds and irrelevant information, allowing the mushroom body to be preserved and increasing recognition accuracy in cluttered environments. Given the challenges posed by visually similar mushroom species, which can lead to mislabeling and inaccuracies, a label smoothing regularization strategy was employed. This approach mitigated the model's over-reliance on the original hard labels by transforming them into soft labels during training, thereby enhancing the model's generalization capabilities.

(2) A canonical dataset, Fungi, was also constructed, containing 14 836 mushroom images. This dataset included 59 mushroom species, such as *Amanita flavoconia*, *Laricifomes officinalis*, *Omphalotus olivascens*, and *Trametes versicolor*, all captured in natural environments.

1 Dataset Collection and Construction

1.1 MO106 Dataset

The dataset used was based on the MO106 dataset created by KISS et al.[16]. This dataset integrated two primary data sources: the Fungi classification challenge (FGVCx) dataset (https://github.com/visipedia/fgvcx_fungi_comp) and image data from the mushroom observer website (https://mushroomobserver.). The FGVCx dataset, published in 2018, included 1 394 fungal species with 85 578 training images, 4 182 validation images, and 9 758 test images, all contributed by the Danish Svampe Atlas. The mushroom observer website aims to document various mushroom species, assisted in the identification of unfamiliar mushrooms, and fostered a community dedicated to mycological research. Users can contribute by uploading mushroom observations,

sharing images, and commenting on others' submissions, all in accordance with the site's protocol. The website hosted a large and dynamically updated collection of mushroom data. The MO106 dataset contained a total of 106 types of mushrooms, with 29 090 images. The category with the least number of images had 105 images, while the category with the most images had 581 images. The resolution of the image was 640 pixel × 640 pixel. The parts of mushroom images in the MO106 dataset were shown in Fig. 1.

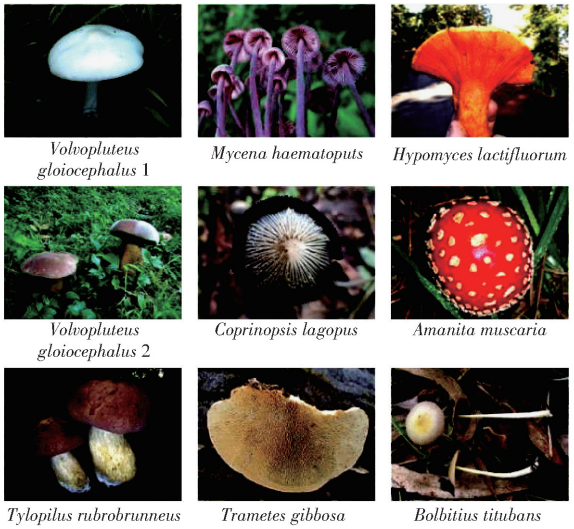


Fig. 1 Examples in MO106 dataset

1.2 Fungi Dataset

The MO106 dataset, which included a wide variety of mushrooms with relatively balanced category distributions, served as the foundation for the creation of the Fungi dataset in this experiment. The processing steps were as follows: firstly, images in the MO106 dataset were duplicated, originally compiled by KISS et al. [16], were removed to reduce data redundancy. This step involved manual verification to address cases of multiple uploads of identical images. Next, images with overly complex backgrounds, which made it difficult for mushroom experts to accurately identify the mushrooms (as shown in Fig. 2), and images where the mushroom features were not clearly visible (such as incomplete mushroom slices caused by irregular photography, as shown in Fig. 3), were discarded. Finally, a curated 59-class mushroom dataset, referred to as Fungi, was compiled, containing 14 836 images. The distribution was relatively balanced, with approximately 300 images per category. Fig. 4 provided distribution of mushroom images.

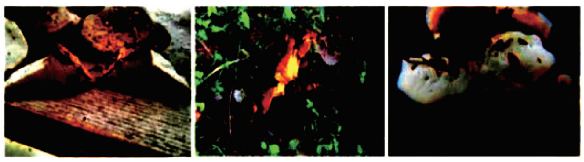


Fig. 2 Examples with complex background



Fig. 3 Examples with incomplete cut

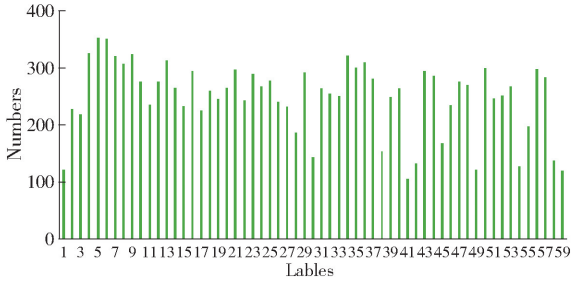


Fig. 4 Distribution of mushroom images in Fungi

1.3 Data Augmentation and Pre-processing

To enhance the diversity of dataset, data augmentation techniques was applied. Given the large size of the Fungi dataset, offline data augmentation would be computationally intensive, so online data augmentation was employed. Before each training epoch, the dataset underwent transformations such as rotation, translation, and others. In each subsequent epoch, these transformations were applied with random parameters, ensuring that the training data varied across epochs. This approach effectively doubled the dataset size for each epoch. The specific transformations, implemented using OpenCV, included operations such as "Adjust brightness", "Rotation", "Cut", "Gaussian blur", and "Random occlusion". This was shown in Fig. 5.

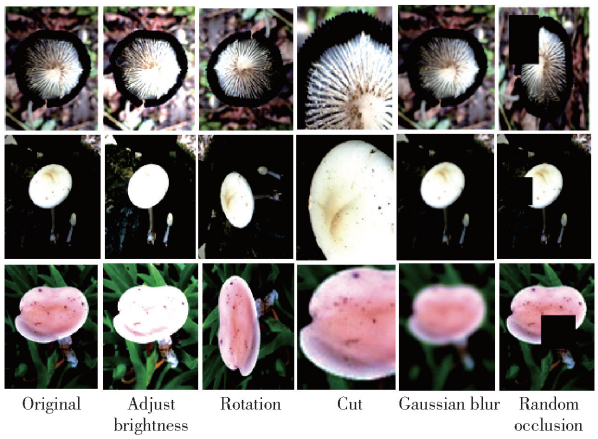


Fig. 5 Enhanced sample mushroom images

2　Methods

2.1　DenseNet Architecture

The DenseNet architecture[17] was used as the base model. DenseNet is a densely connected network composed of multiple dense blocks. Its dense connectivity facilitates cross-connections between features at different scales, enabling a more effective representation of complex semantic relationships in multi-scale objects[18]. The direct connections ensure efficient information flow. Each layer receives input from all preceding layers, promoting feature reuse, as show in Fig. 6. For each layer, the output characteristics of Dense Block were as follows:

$$x_L = H_L([x_0, x_1, \cdots, x_{L-1}]) \quad (1)$$

Where x_L was the output of layer L; x_{L-1} represented the output of layer $L-1$; $H(\cdot)$ represented a sequence of nonlinear calculations.

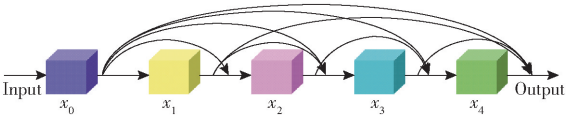


Fig. 6　Network structure of DenseNet

2.2　Erase Module

Occlusion[19] is another critical factor affecting the robustness of convolutional neural networks (CNNs). In practical applications, improving the ability of CNN models to recognize images with varying degrees of occlusion is essential. When the target region in a training image is occluded, a robust classification model should be able to accurately classify the target based on its remaining features and background information. In agricultural applications, occlusion of the target crop can obscure its features, hindering the ability to effectively recognize and analyze the object.

Inspired by this, an erasing mechanism was used to improve the accuracy and robustness of CNN. The erasing mechanism can erase complex backgrounds and interference information, enhancing the model focus on key features. As shown in Fig. 7, mushrooms themselves were relatively small and difficult to identify. Mushroom color and background increased the difficulty of model recognition. There was interference information around the mushroom (such as flowers, backpacks, etc., especially when the interfering objectswere similar in color to the mushroom), which greatly increased the difficulty of recognition.

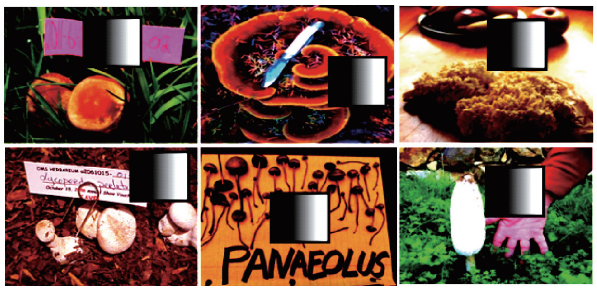


Fig. 7　Images of a mushroom with occlusion

By integrating attention modules, the network can dynamically focus on the most relevant regions in the image while ignoring or suppressing non discriminatory background elements. In this case, the model would be able to learn more about the unique features of mushrooms, such as shape, color, and texture, while effectively ignoring black pine nuts and other irrelevant objects that may interfere with classification, as show in Fig. 8. Meanwhile, adjusting the erase threshold can control the erase area in the image, simulate occlusion situations, enhance the model's attention to visible areas, and improve the model's robustness. This module determined the erased area by mining the correlation between each region in the training image and the label.



(a) Before erasing　　(b) After erasing

Fig. 8　Images before and after erasing

The effectiveness of erase modules lay in their ability to mitigate the impact of irrelevant or misleading background information, which often confounds traditional CNNs. This was achieved through the generation of attention maps that highlighted the most important regions of the image, effectively isolating the target object from its surrounding environment. By erasing non-informative or occluded parts of the image, the network can focus more on the visible features of the crop, improving the model's ability to make accurate predictions even under suboptimal conditions. The erase module was shown in Fig. 9.

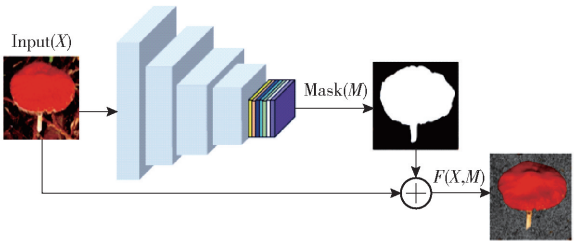


Fig. 9 Overall flow of erase module

The importance score of the rectangular region S was obtained as follows:

$$G(S,x,y)=\frac{\sum_{(i,j)\in S} g_{ij}(x,y)}{aa} \tag{2}$$

where (i,j) represented the image of the selected region. x was the input image, y was the label corresponding to image x, $g_{ij}(x,y)$ represented the given label y in image x. The activation mapping of pixel point $x(i,j)$ represented the area of the erased region, and $G(S,x,y)$ represented all images in that region. The weighted sum of prime values represented the importance of the region.

After obtaining the importance level of each region in the image based on the attention activation map, erasing can avoid the randomness of the erased regions, thereby alleviating the problems of over erasure and under erasure, and generating better training samples. Using a new training model can improve its generalization and robustness. By adjusting the erasure threshold and varying the size of the erased region, different portions of the mushroom image can be selectively erased. This ensured that each image retained an adequate receptive field and contextual information, addressing the issues of under-erasure and over-erasure, as shown in Fig. 10.

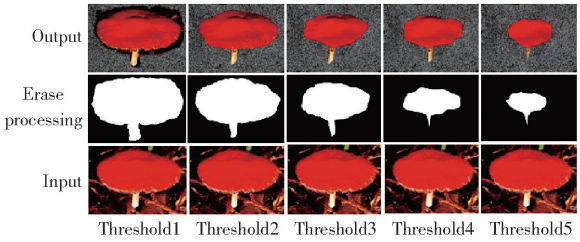


Fig. 10 Examples processed with different thresholds

2.3 Label Smoothing Strategy

In the 59 types of mushroom images, there were many types of mushrooms, and the shapes of some mushrooms with similar genera and families were unsuitable for distinguishing. Mushroom observers were prone to misrecognition when labeling, which may lead for mushroom identification to inaccurate actual labeling. Directly using such data may make the model overconfident and quickly lead to network overfitting. Therefore, the label smoothing regularization strategy was used to prevent the model from relying too much on the labels of the original data and convert the original hard labels into soft labels during the training process to improve the model's generalization ability[20-22]. The main formula was as follows:

$$\boldsymbol{y}_i(c)=\begin{cases}1-\sigma_i & (c=k)\\ \dfrac{\sigma_i}{c} & (c\neq k)\end{cases} \tag{3}$$

Where $\boldsymbol{y}_i(c)$ was a c-dimensional vector, and the element value of each dimension (0 or 1) indicated whether the current sample belonged to that class of data. k denoted the true class of y, and σ was the label smoothing hyperparameter, which empirically ranged as $\sigma\in[0,1]$.

2.4 Improved EL-DenseNet Model

The DenseNet architecture promotes feature reuse through dense connections, facilitating feature propagation, mitigating the vanishing gradient problem, and reducing the number of parameters. However, when the target area in a training image is occluded, the DenseNet model lacks robustness and fails to accurately classify the target based on remaining features and background information. To address this limitation, the erase module was introduced into DenseNet. In the event of occlusion, this erase module became more sensitive to unobstructed information, enhancing the attention to the visible regions of the mushroom and enabling accurate target classification.

In the original DenseNet, if a Dense Block contains L layers, the input to each layer is the concatenation of the outputs from all previous layers, as described in eq. (1). The enhanced E-DenseNet incorporated an erase module before and after each non-linear activation function within every Dense Block of the original DenseNet. In terms of model performance and experimental hardware constraints, DenseNet121 was selected as the backbone network and optimized accordingly. Specifically, the erase module was integrated into each Dense Block (DB). This modification ensured that the feature maps processed by each non-linear function were selectively

erased by the erase module, based on the importance of different regions in the image. By removing irrelevant background information or occlusions, the model generated improved training samples, thereby enhancing both its robustness and classification accuracy. The proposed network architecture was referred to as EL－DenseNet, with the structure of the EL－DenseBlock shown in Fig. 11.

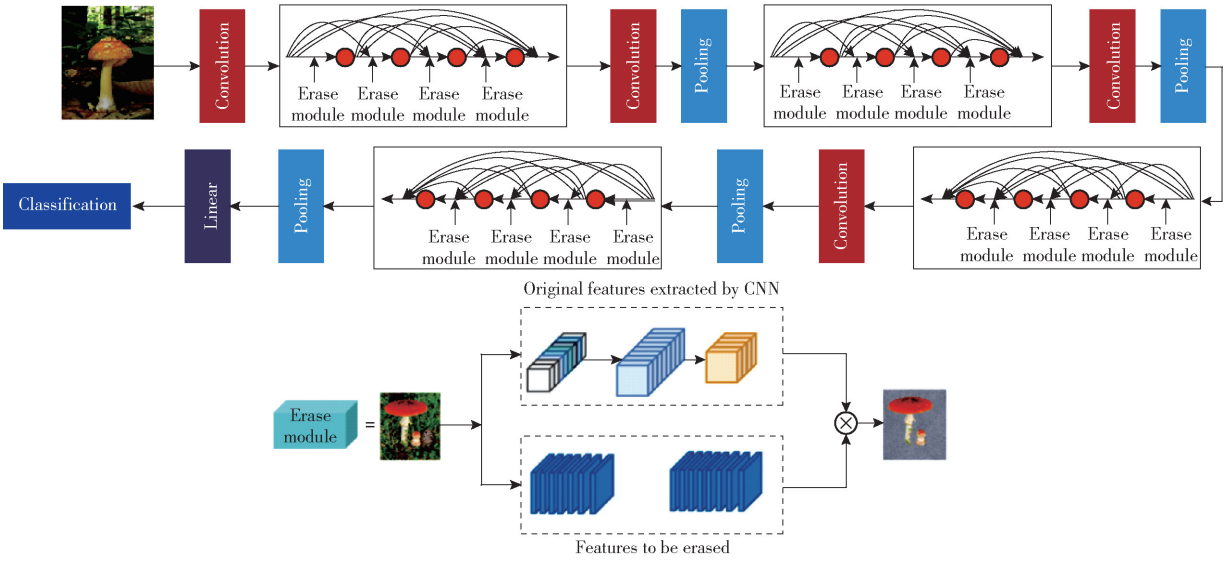


Fig. 11　Network structure of EL－DenseNet model

Totally 59 mushroom images were utilized, some of which depicted species with similar genera and families, making them difficult to distinguish. As a result, mushroom observers may mislabel these images, leading to inaccurate annotations in practice. Using such data directly could cause the model to become overly confident, potentially resulting in overfitting. To mitigate this, label smoothing regularization was applied to prevent the model from relying too heavily on the original labels. During training, this technique converted the original hard labels into soft labels, thereby improving the generalization ability. The network model, optimized with the erase module and incorporating label smoothing, was referred to as EL－DenseNet.

3　Experiments and Analyses

3.1　Model Learning Rate Optimization

The learning rate is a critical factor influencing model performance and plays a pivotal role in achieving optimal results. A high learning rate can cause the network to fail to converge and oscillate around the optimal value, where as a low learning rate can lead to slow convergence and increase the time required to locate the optimal value. To determine the optimal learning rate for the improved EL－DenseNet model, experiments were conducted with varying learning rates. The dataset was randomly divided into training, validation, and test sets in a ratio of 7∶2∶1. The optimization algorithm used in the experiments was stochastic gradient descent (SGD).

To address this limitation, momentum was introduced to enhance the SGD algorithm. Momentum retained a portion of the previous update direction while incorporating the current gradient to fine-tune the final update, thereby improving model stability, accelerating convergence, and reducing the risk of getting stuck in local optima. Experiments were conducted with initial learning rates of 0.000 5, 0.005, 0.05, and 0.5. The learning rate decayed after each parameter update was set to 0.000 5, momentum was fixed at 0.9, the batch size was set to 128, and the number of epochs was set to 100. The training set was randomly shuffled before each iteration.

The experimental results, presented in Fig. 12, indicated that the EL－DenseNet model achieved the highest accuracy when the learning rate was set to 0.001.

3.2　Model Comparison Experiment

To verify the superiority of the improved DenseNet model structure incorporating the erase module and the enhancement of model performance achieved by the label smoothing loss function in complex backgrounds, comparative experiments were conducted using the

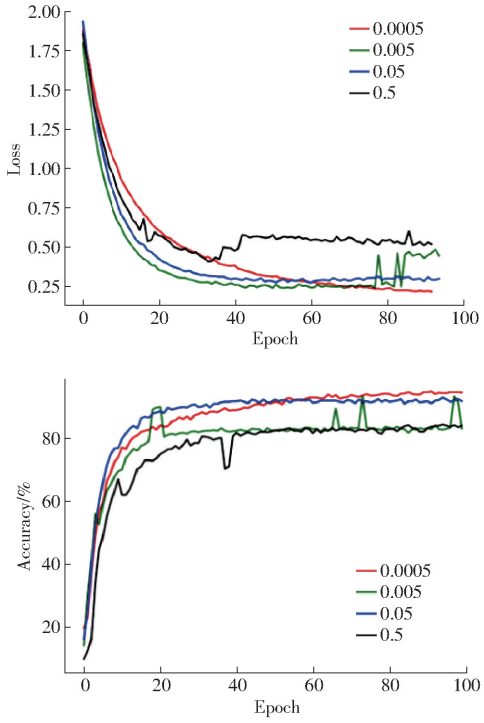


Fig. 12 Accuracy and loss of different learning rate

optimal learning rate identified in Section 3.1. The experiments evaluated three transfer learning-based network models: DenseNet, E－DenseNet, and EL－DenseNet. The experimental results were presented in Fig. 13. Each model achieved near-complete convergence after approximately 80 epochs. The highest accuracy achieved by the three models was 91.2% for DenseNet, 94.8% for E－DenseNet, and 96.7% for EL－DenseNet. Among these, the EL－DenseNet model proposed demonstrated the highest

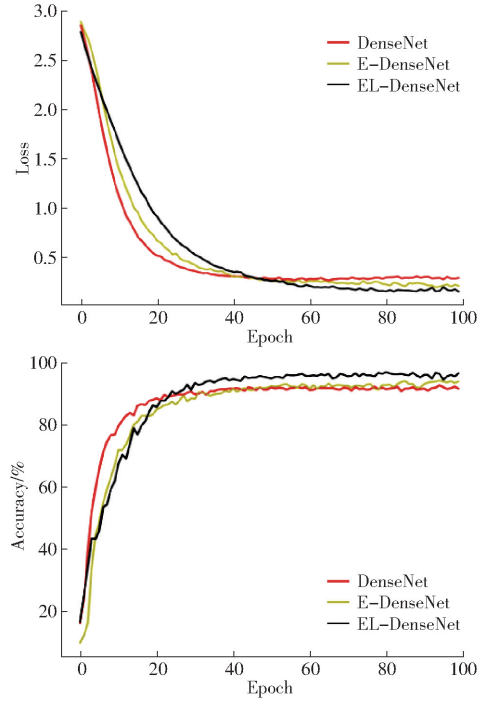


Fig. 13 Accuracy and loss of different models

recognition accuracy. Specifically, E－DenseNet improved the accuracy by 3.69 percentage points compared with DenseNet, while EL－DenseNet further enhanced the accuracy by 1.9 percentage points compared with E－DenseNet.

Table 1 showed experimental results of different ablation methods. Compared with Test 1, the simultaneous use of the erasure module and label smoothing improved the model's accuracy by 5.5 percentage points (from 91.2% to 96.7%). This indicated that the erasure module and label smoothing technique exhibited a synergistic effect, further enhancing the performance. The erase module strengthened the robustness to input data, while label smoothing improved its generalization.

Table 1 Comparison of experimental results of ablation methods

Test	Erase module	Label smoothing	Accuracy/%
1	×	×	91.2
2	×	√	93.1
3	√	×	94.8
4	√	√	96.7

Note: √ represented the adoption of this method, while × represented the non-adoption of this method.

Table 2 presented the performance of the EL－DenseNet model on different datasets. On the MO106 dataset, the model achieved high performance, with an accuracy of 92.9%, precision of 92.9%, recall of 92.3%, and an F1 score of 92.6%, indicating a well-balanced and stable performance in the classification task. On the Fungi dataset, the model demonstrated even more outstanding results, with accuracy, precision, recall, and F1 score all reached approximately 96.7%, further validating the robustness and generalization capability of EL－DenseNet across diverse datasets.

Table 2 Comparison of evaluation results between MO106 and Fungi %

Dataset	Accuracy	Precision	Recall	F1 score
MO106	92.9	92.9	92.3	92.6
Fungi	96.7	96.8	96.7	96.7

To evaluate the effectiveness of the proposed model on the mushroom dataset, comparisons were conducted with several classical classification models, including MobileNet V2, ResNet50, Efficient－B0, and Efficient－B7. The results were summarized in Table 3. MobileNet V2 exhibited the lowest recognition performance on mushroom images, achieving an

Table 3　Comparison of evaluation results of different models

Model	Accuracy/%	Precision/%	Recall/%	F1 score/%	FLOPs	Parameters
MobileNet V2	75.5	75.7	75.5	75.6	3.1×10^{8}	3.5×10^{6}
ResNet50	87.5	87.6	87.5	87.6	4.0×10^{9}	2.6×10^{9}
Efficient－B0	87.8	87.8	88.1	87.9	3.9×10^{8}	7.7×10^{9}
Efficient－B7	88.9	88.3	88.9	88.6	3.7×10^{10}	6.6×10^{10}
DenseNet	91.2	91.4	91.4	91.4	3.9×10^{9}	2.0×10^{10}
E－DenseNet	94.8	94.8	94.8	94.8	4.3×10^{9}	2.0×10^{10}
EL－DenseNet	96.7	96.8	96.7	96.7	4.3×10^{9}	2.0×10^{10}

accuracy of only 75.5%. In contrast, ResNet50, Efficient－B0, and Efficient－B7 all demonstrated accuracies around 90%, specifically 87.5% for ResNet50, 87.8% for Efficient－B0, and 88.9% for Efficient－B7. The EL－DenseNet model proposed achieved the highest recognition accuracy of 96.7%. EL－DenseNet demonstrated the best performance, achieving the highest values in accuracy, precision, recall, and F1 score (96.7%, 96.8%, 96.7%, and 96.7%, respectively), indicating exceptional classification performance and balance. E－DenseNet followed closely, with its metrics being nearly comparable to those of EL－DenseNet, achieving an accuracy and F1 score of 94.8%, which suggested high accuracy and robustness in the classification task. This superior performance can be attributed to the erase module, which reduced the impact of complex background and interference information, allowing the model to focus more on key features. Additionally, by introducing a label smoothing loss function, the model mitigated over-reliance on hard labels, converting them into soft labels during training to enhance generalization ability and improve accuracy in complex backgrounds. The EL－DenseNet achieved the highest accuracy and F1 score while maintaining a low number of FLOPs and parameters, indicating that it achieved the best balance between performance and efficiency. This made EL－DenseNet a significant advantage in practical applications, especially in environments where computing resources were limited.

Experiments were conducted on the test set, comprising 59 types of mushrooms in the Fungi, and the confusion matrix was generated accordingly. The results were illustrated in Fig. 14.

Grad－CAM[23] visualization technology effectively represented the feature information that each module in a network emphasized within an image, using a heatmap format. Fig. 15 illustrated the partial feature map visualization results of the EL－DenseNet mushroom recognition model during the feature extraction process, generated using Grad－CAM technology. The earlier E－DenseBlock primarily focused on non-mushroom features. However, as the network depth was increased, the extraction of mushroom-related features became progressively more pronounced. Due to the characteristics of dense networks, the deeper layers' focus on mushroom features was influenced by the feature information from all preceding layers. In the EL－DenseBlock, all Dense Blocks were optimized by using the erasure attention module, which strengthened the extraction of mushroom features while suppressing the extraction of non-mushroom features, thereby enhancing the overall performance of the model.

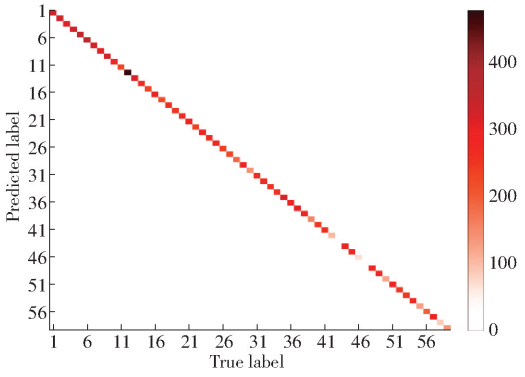


Fig. 14　Confusion matrix based on EL－DenseNet

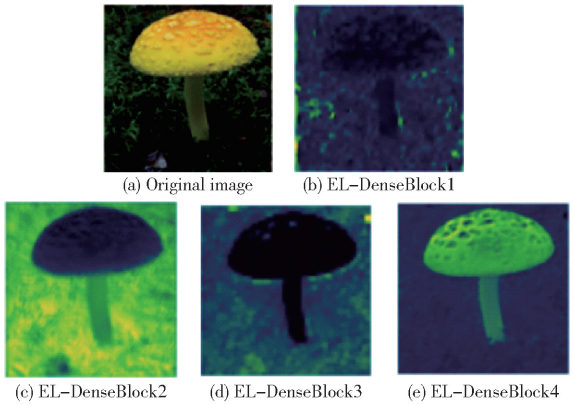


Fig. 15　Feature maps of feature extraction

4 Conclusion

(1) An EL - DenseNet model was proposed based on the erase module to address the poor performance of DenseNet in complex backgrounds. This not only improved accuracy but also enhanced robustness. A series of experiments were conducted using the publicly available MO106 as the source dataset and a self-constructed mushroom dataset, Fungi, consisting of 59 categories, as the target dataset.

(2) Comparative experiments among three models demonstrated that E - DenseNet improved accuracy by 3.6 percentage points over the original DenseNet, while EL - DenseNet further improved accuracy by 1.9 percentage points over E -DenseNet. The accuracy of the EL - DenseNet model reached 96.7%, indicating that the erase module effectively removed complex background and interference information, thereby enhancing the model's focus on key features. Additionally, the label smoothing helped mitigate excessive reliance on the original data labels, improving the model's generalization ability.

Reference

[1] ADITY R, PURWONO S H. Jamur: info lengkap dan kiat sukses agribisnis[J]. Depok, Indonesia/West Java: Agriflo, 2012, 2: 195 - 201.

[2] HAMZA A, MYLARAPU A, KRISHNA K V, et al. An insight into the nutritional and medicinal value of edible mushrooms: a natural treasury for human health[J]. Journal of Biotechnology, 2024, 381: 86 - 99.

[3] MIAO Quanlong, ZHOU Yang, LI Jiantao, et al. Extraction method of phenotypic parameters of *Agaricus bisporus* based on YOLO v8 - ABSeg[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 158 - 168. (in Chinese)

[4] YANG Shuzhen, HUANG Jie, YUAN Jin. Spatiotemporal prediction algorithm for mushroom growth status based on improved LSTM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 221 - 230. (in Chinese)

[5] CHEN Z, ZHANG P, ZHANG Z. Investigation and analysis of 102 mushroom poisoning cases in Southern China from 1994 to 2012[J]. Fungal Diversity, 2014, 64: 123 - 131.

[6] YAN Z, LIU H, LI J, et al. Application of identification and evaluation techniques for edible mushrooms: a review[J]. Critical Reviews in Analytical Chemistry, 2023, 53(3): 634 - 654.

[7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 - 778.

[8] FENG J, TAN H, LI W, et al. Conv2NeXt: reconsidering ConvNeXt network design for image recognition[C]// 2022 International Conference on Computers and Artificial Intelligence Technologies (CAIT). IEEE, 2022: 53 - 60.

[9] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115: 211 - 252.

[10] FENG J, LING X, WANG Y, et al. Mushroom classification based on deep residual network[C]// 2023 IEEE 4th International Conference on Pattern Recognition and Machine Learning (PRML). IEEE, 2023: 128 - 133.

[11] ZAHAN N, HASAN M Z, MALEK M A, et al. A deep learning-based approach for edible, inedible and poisonous mushroom classification[C]//2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE, 2021: 440 - 444.

[12] WU L, CHEN Y. Mushroom recognition and classification based on convolutional neural network[C]// 2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). IEEE, 2022, 5: 1430 - 1433.

[13] KHASANAH H M, AMINUDDIN A, ABDULLOH F F, et al. Optimizing mushroom classification through machine learning and hyperparameter tuning[J]. Engineering and Applied Science Research, 2024, 51(5): 651 - 660.

[14] CHEN S, TAN X, WANG B, et al. Reverse attention for salient object detection[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 234 - 250.

[15] HOU Q, JIANG P T, WEI Y, et al. Self-erasing network for integral object attention[J]. Advances in Neural Information Processing Systems, 2018, 31: 1 - 11.

[16] KISS N, CZÙNI L. Mushroom image classification with CNNs: a case-study of different learning strategies[C]//2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA). IEEE, 2021: 165 - 170.

[17] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 - 4708.

[18] ZHANG K, GUO Y, WANG X, et al. Multiple feature reweight densenet for image classification[J]. IEEE Access, 2019, 7: 9872 - 9880.

[19] KÖRSCHENS M, BODESHEIM P, DENZLER J. Occlusion-robustness of convolutional neural networks via inverted cutout[C]//2022 26th International Conference on Pattern Recognition (ICPR). IEEE, 2022: 2829 - 2835.

[20] MÜLLER R, KORNBLITH S, HINTON G E. When does label smoothing help? [J]. Advances in Neural Information Processing Systems, 2019, 32:1－10.
[21] ZHANG C B, JIANG P T, HOU Q, et al. Delving deep into label smoothing[J]. IEEE Transactions on Image Processing, 2021, 30: 5984－5996.
[22] LUKASIK M, BHOJANAPALLI S, MENON A, et al. Does label smoothing mitigate label noise? [C] // International Conference on Machine Learning. PMLR, 2020: 6448－6458.
[23] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad－CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128: 336－359.

EL－DenseNet：基于擦除模块和 DenseNet 的蘑菇识别

王耀君　赵韦婷　别宇辉　贾　璐
（中国农业大学信息与电气工程学院，北京 100083）

摘要： 目标遮挡是计算机视觉中的一个重要挑战，尤其是在农业应用中，作物遮挡会模糊关键特征，影响模型的识别性能。为了解决该问题，提出了一种基于 EL－DenseNet 模型的蘑菇识别方法。EL－DenseNet 是 DenseNet 的扩展，它包含一个擦除注意模块，用于增强对可见特征的敏感性。擦除模块有助于消除复杂的背景和不相关的信息，使蘑菇关键特征得以保留，并提高模型在杂乱环境中的识别精度。考虑到相似蘑菇物种识别的困难性，采用标签平滑正则化方法来减少由观察者引起的错误标签。该策略在训练过程中将硬标签转换为软标签，减少了模型对噪声标签的过度依赖，提高了模型的泛化能力。实验结果表明，结合迁移学习的 EL－DenseNet 算法对有遮挡复杂背景下的蘑菇识别率为 96.7%。与原有的 DenseNet 等经典模型相比，该方法具有更高的识别精度和鲁棒性，为智能蘑菇识别提供了一种很有前景的解决方案。

关键词： 多蘑菇识别；擦除模块；标签平滑；DenseNet

基金项目： 国家重点研发计划项目（2023YFD1701000）

（上接第 584 页）

[19] 叶志成，张莉莉，廖凯华，等. 基于 Morris 和 Sobol 法的茶园地上部生物量与氮磷流失模拟敏感性分析[J]. 水土保持研究，2024，31(2)：43－51，59.
YE Zhicheng, ZHANG Lili, LIAO Kaihua, et al. Sensitivity analysis of simulation of aboveground biomass and nitrogen and phosphorus losses in tea plantation based on Morris and Sobol methods[J]. Research of Soil and Water Conservation, 2024, 31(2):43－51, 59. (in Chinese)
[20] 姜传礼，赵健赟，丁圆圆，等. SPA 算法与机器学习的黄河源土壤水分反演[J]. 光谱学与光谱分析，2023，43(6)：1961－1967.
JIANG Chuanli, ZHAO Jianyun, DING Yuanyuan, et al. Study on soil water retrieval technology of Yellow River source based on SPA algorithm and machine learning[J]. Spectroscopy and Spectral Analysis, 2023, 43(6):1961－1967. (in Chinese)
[21] 宁京，邹滨，涂宇龙，等. 土壤 As 含量光谱指数反演方法评估[J]. 光谱学与光谱分析，2024，44(5)：1472－1481.
NING Jing, ZOU Bin, TU Yulong, et al. Evaluation of soil As concentration estimation method based on spectral indices[J]. Spectroscopy and Spectral Analysis, 2024, 44(5):1472－1481. (in Chinese)
[22] 李婷婷，胡钧铭，韦彩会，等. 水稻叶片营养吸收机制及专用叶面肥发展趋势[J]. 江苏农业科学，2016，44(12)：12－16.