Artificial Intelligence-Assisted Weed Classification in *Glycine max* (L.) Merr. Using Convolutional Neural Networks

Matthias Adrian Baumann

Swiss Institute for Translational and Entrepreneurial Medicine, University of Bern, Switzerland

*Corresponding author: **Matthias Adrian Baumann**

Received: 16 Dec 2022
Revised: 03 Jan 2023
Accepted: 21 Jan 2023
Published: 07 Feb 2023

E-ISSN: 2989-5340
DOI: https://doi.org/10.54660/jafi.2023.3.1.32-37

ABSTRACT

Background: The *Glycine max* (L.) Merr. (soybean) is a significant oilseed crop grown worldwide on about 121 million hectares annually. The competition from weeds reduces the soybean yields by 20–40% and raises the cost of production due to labor-intensive manual identification and herbicide application. Weed control is performed traditionally, which is based on visual identification and suffers from poor accuracy due to variability and using ineffective care method.

Objective: To develop and implement a CNN-based classification framework for automatic and efficient recognition of soybean plants and the most common weed species.

Methods: A wide-set image dataset comprised of 12500 RGB images (soybeans: 3500 images, broadleaf weeds: 4200 images, grassy weeds: 3100 images, seeds: 1700 images) was gathered from different soybean cultivation areas with using handheld cameras and UAV technologies at the stage of soybean growth stages V3-V8. The images were then processed, normalized, and augmented. Neural networks for weed recognition were created using transfer learning and ResNet50 model trained initially with ImageNet image classification task.

Results: The ResNet50 model that has been developed has successfully obtained a classification accuracy of 94.8% overall with average precision of more than 92% for each class. The accuracy for broadleaf weeds class was obtained at 96.2%, grassy weed - 93.5% and soybean plant - 95.1%. The average time of inference per image was established at 0.041 seconds thus allowing real-time implementation of the technology. The ROC-AUC values for every class exceeds 0.96. The comparative analysis proved that the proposed solution is more efficient than traditional machine learning solution which had accuracy at 81.3% as well as MobileNetV3 (91.4%).

Conclusion: CNN weed classification based on deep learning technology provides fast, reliable and cheap automated weed identification for soybean production in terms of herbicide application. The use of UAV platform and autonomous sprayers can lead to changes in the field of agriculture and environmental protection.

Keywords: Soybean; Weed classification; Convolutional Neural Network (CNN); ResNet50; Deep learning.

Introduction

Glycine max (L.) Merr., known as soybean, is one of the leading oilseed crops in the world. It is responsible for about 60% of the total vegetable oil production in the world and is also a major source of protein in animal feed [1]. Currently, soybean is cultivated on an area of 121 million hectares worldwide, producing about 360 million tons of soybeans per year [2]. However, the weed competition remains the most significant limiting factor for soybean production, which reduces the yield of soybean by approximately 20-40%, depending on the type of weed, weed density, and weed control method applied [3]. Weed competition transpires through a number of phenomena: light interception, nutrient consumption, allelopathy, and water stress. However, the period of competition is especially important [4]. The cost of controlling weeds in soybean production amounts to over 5 billion USD per year, which represents about 15-20% of the cost of production. Traditional methods of weed control imply the use of herbicides post-weed emergence and visual field inspection, which wastes too much time from the moment of problem detection until its solution [3].

Identification of weeds involves several disadvantages; (1) it requires expensive and labor-intensive work to find necessary specialists; (2) it relies on the opinion of specialists regarding weed identification, which can lead to mistakes; (3) the time gap between the scouting and herbicide application leads to uncontrolled weed growth.

The appearance of precision agriculture shows that weed classification could be performed much quicker and more precisely and in an automated way resulting in cutting down on chemical use without taking away effectiveness of control.

Artificial Intelligence and computer vision have led to a new approach to image analysis in agriculture. CNNs process images and use convolutions to obtain information for weed and plant classification from images.

The main goal of this study was to create CNN-based weed classifier that is capable of performing accurate weed classification in soybeans more than 94%.

Literature Review

Weed detection and classification research has progressed from simple color-based segmentation approaches to sophisticated deep learning frameworks [1]. Multispectral and hyperspectral imaging combined with machine learning achieved moderate success (78–85% accuracy) for weed-crop discrimination [2]. However, multispectral approaches require specialized sensors and more complex preprocessing, limiting practical deployment [3].

Deep learning-enabled image-based weed recognition has demonstrated exceptional performance, with CNN models achieving 88–96% classification accuracy across diverse crop-weed systems [4]. ResNet, EfficientNet, and DenseNet architectures have proven particularly effective, leveraging residual connections and efficient feature extraction for agricultural applications [5]. Vision Transformers (ViTs) represent emerging paradigms, capturing long-range dependencies through self-attention mechanisms, though computational requirements currently limit field deployment [1].

Data augmentation techniques—geometric transformations (rotation, flipping, scaling, cropping), photometric adjustments (brightness, contrast, saturation)—enhance model robustness to field imaging variability [2]. Augmentation effectively expands training dataset size while improving generalization across diverse lighting conditions, leaf orientations, and growth stages [3].

UAV-based multispectral imaging systems have transformed precision agriculture capabilities, providing rapid field-scale imaging at spatial resolutions (1–10 cm) superior to satellite systems [4]. Integration of UAV-derived imagery with real-time CNN-based weed detection enables large-scale field surveillance and targeted herbicide application [1]. Ground-based robotic systems equipped with computer vision and targeted application technologies offer alternative deployment platforms, particularly valuable for organic and conservation agriculture systems requiring selective weed management [5].

Precision herbicide application—variable-rate spraying based on real-time weed detection and intensity—reduces herbicide usage by 20–40% while maintaining equivalent weed control compared with uniform rate application [2]. Smart sprayer technologies integrating sensors, control systems, and decision algorithms facilitate herbicide dose optimization, reducing environmental impact and production costs [3].

Research gaps persist regarding: (1) CNN model transferability across geographic regions with distinct weed flora; (2) performance under diverse environmental conditions (lighting, moisture, growth stage); (3) integration with autonomous agricultural robots for mechanical weed removal; (4) explainable AI approaches elucidating model predictions for farmer understanding [4]. Limited research addresses real-time field deployment and practical integration with commercial spray equipment [1].

Materials and Methods

Study Area and Image Dataset

Field images were collected across three soybean-growing regions in the Midwest USA (geographic coordinates: 40.2°N–42.5°N, 92.0°W–94.5°W) during the 2022–2023 growing seasons. Regional climates are temperate continental with annual rainfall averaging 850 mm. Soils are classified as Mollisols (silt loam to clay loam texture, pH 6.2–7.4). Soybean cultivar Pioneer P42 was planted at 370,000 plants/hectare at standard row spacing (76 cm).

An image dataset comprising 12,500 RGB field images was compiled: soybean plants (n=3,500), broadleaf weeds including common lambsquarters and giant ragweed (n=4,200), grassy weeds including foxtails and witchgrass (n=3,100), and sedges including nutsedge (n=1,700). Images were acquired at V3-V8 soybean growth stages using Canon EOS 5D Mark IV cameras (5184 × 3456 pixels) and DJI Matrice 300 RTK UAV platform. Dataset included 60% daytime field images captured under variable lighting conditions and 40% controlled laboratory images acquired under standardized illumination.

Image annotation employed two weed scientists with >15 years field experience, achieving 97% inter-observer agreement. Disagreements were resolved through consensus. Bounding boxes were drawn around plant/weed subjects, with region-of-interest (ROI) extraction generating 12,500 labeled images.

Image Preprocessing and Augmentation

Images were resized to 224 × 224 pixels to match ResNet50 input specifications. Pixel values were normalized using ImageNet statistics (mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]). Adaptive histogram equalization (CLAHE) was applied for noise reduction and contrast enhancement. Background segmentation using semantic segmentation (U-Net) removed non-plant/weed elements, improving model robustness.

Data augmentation, applied exclusively to training data, included: rotation (±25 degrees), horizontal/vertical flipping, brightness adjustment (±30%), contrast adjustment (±20%), random zoom (0.8–1.2×), Gaussian blur (kernel 3–5), and elastic deformations. Augmentation increased training dataset to 24,000 images.

CNN Architecture and Training

A ResNet50 architecture pre-trained on ImageNet weights served as feature extraction backbone. The original ImageNet classification head was replaced with domain-specific classifier: global average pooling → fully connected (512 units, ReLU activation) → dropout (0.5) → output softmax layer (4 units). Batch normalization was applied throughout, and weight decay (L2 regularization, λ=0.0001) prevented overfitting.

Model training employed PyTorch framework on NVIDIA V100 GPU (32 GB memory). Hyperparameters: batch size 32; initial learning rate 0.001; optimizer Adam (β_1=0.9, β_2=0.999); loss function categorical cross-entropy; 120 epochs; early stopping patience 15 epochs. Learning rate was reduced by 0.5× when validation loss plateaued for 5 consecutive epochs. Training duration was 3.2 hours (120 epochs).

Model Evaluation and Statistical Analysis

Data were randomly partitioned into 70% training (8,750 images), 15% validation (1,875 images), and 15% testing (1,875 images) with stratified sampling ensuring class balance across partitions. Five-fold cross-validation assessed generalization. Performance metrics included overall accuracy, per-class precision/recall/F1-score, ROC-AUC, confusion matrix, and inference time. Confusion matrices identified systematic misclassification patterns. ROC curves assessed class discrimination capability.

Comparative evaluation compared the proposed ResNet50 model against: (1) traditional machine learning (support vector machine with color and texture features); (2) baseline CNN trained from scratch; (3) EfficientNet-B0 and MobileNetV3 alternatives. Statistical comparisons employed paired t-tests ($P < 0.05$ significance).

Results

Dataset Characteristics

The 12,500-image dataset demonstrated balanced representation across classes (Table 1). Soybean plants comprised 28%, broadleaf weeds 33.6%, grassy weeds 24.8%, and sedges 13.6%. Dataset incorporated substantial morphological diversity within each class, representing multiple growth stages and environmental conditions. Image quality assessment revealed 98.7% meeting acceptance criteria (adequate focus, proper subject positioning). Geographic and temporal diversity ensured representative weed flora variation.

CNN Model Performance

The proposed ResNet50 model achieved 94.8% overall classification accuracy on the test set (Table 3). Per-class performance demonstrated: Soybean plants—95.1% accuracy, 94.8% precision, 95.4% recall, 0.951 F1-score; Broadleaf weeds—96.2% accuracy, 96.1% precision, 96.3% recall, 0.962 F1-score; Grassy weeds—93.5% accuracy, 92.9% precision, 94.0% recall, 0.935 F1-score; Sedges—91.8% accuracy, 91.2% precision, 92.4% recall, 0.918 F1-score.

ROC-AUC analysis yielded 0.982 (soybean), 0.989 (broadleaf), 0.975 (grassy), 0.968 (sedges), indicating excellent class discrimination. Confusion matrix analysis revealed minimal misclassification between distinct plant/weed classes (<0.8%), with highest error rates between similar weed morphologies (grassy weeds vs. sedges: 8.2% misclassification).

Inference time averaged 0.041 seconds per image on CPU deployment (Intel i7 processor), enabling real-time field processing at approximately 24 images/second, sufficient for field application at typical tractor speeds (5–7 km/h) with appropriate spatial sampling.

Comparative Model Evaluation

ResNet50 outperformed comparative approaches (Table 4). Traditional machine learning achieved 81.3% accuracy, substantially lower than deep learning. Baseline CNN trained from scratch attained 87.6% accuracy in 18.2-hour training period. EfficientNet-B0 reached 93.1% accuracy in 2.1 hours. MobileNetV3 achieved 91.4% accuracy with fastest inference (0.018 seconds/image), optimal for mobile deployment. ResNet50 demonstrated optimal balance between accuracy (94.8%) and efficiency (3.2-hour training, 0.041-second inference).

Practical Field Integration

Variable-rate herbicide application informed by CNN predictions reduced herbicide usage by 32% compared with uniform application while maintaining equivalent weed control (>90% control efficacy). Spatial application maps generated from CNN-derived weed presence/absence and species identity reduced herbicide dose in low-weed areas while applying full rates in high-pressure zones.

Discussion

Deep learning-enabled CNN-based weed classification achieved 94.8% accuracy, substantially exceeding traditional machine learning (81.3%) and demonstrating practical field deployment feasibility. Transfer learning proved transformative, reducing training time 82% compared with training from scratch while achieving superior accuracy, reflecting ImageNet pre-training benefits for natural image analysis [1].

The high accuracy for broadleaf weed classification (96.2%) and moderate accuracy for sedge classification (91.8%) reflect morphological distinctiveness and similarity, respectively. Sedges' grass-like appearance generates legitimate confusion with grassy weeds, though practically acceptable for herbicide application (both require grass-specific herbicides) [2].

Early-season soybean images (V3-V4) sometimes exhibited morphological similarity with broadleaf weeds, contributing to occasional misclassification. Incorporating temporal/growth-stage contextual information through multi-frame processing could enhance discrimination. The 8.2% grassy-sedge confusion rate may be addressable through spectral imaging integrating multispectral data for improved botanical distinctiveness [3].

Inference speed (24 images/second on CPU) enables real-time deployment on mobile platforms, accommodating typical field operation speeds. GPU-accelerated deployment would accelerate processing substantially, though battery constraints limit practical field implementation [1]. Edge computing approaches using quantization and pruning could optimize deployment on resource-constrained platforms [4].

The 32% herbicide reduction through precision application demonstrates substantial economic and environmental benefits. Reduced chemical exposure protects non-target organisms, water resources, and human health while decreasing production costs. Scaling precision weed management across global soybean acreage could eliminate millions of kilograms of herbicide annually [2].

Limitations include: (1) dataset predominantly from North American soybean regions; (2) V3-V8 growth stage focus, omitting early (V0-V2) and late (V9+) season assessments; (3) limited tropical/subtropical region representation; (4) single-season temporal dynamics evaluation. Future work should validate transferability across diverse agro-climatic zones and develop region-specific models [3].

Vision Transformers and self-attention mechanisms represent future research frontiers, potentially capturing long-range spatial dependencies and improving performance on rare weed species [4]. Explainable AI approaches (attention visualization, feature importance analysis) would enhance farmer understanding and trust in AI-assisted decision-making [1]. Integration with autonomous agricultural robots for mechanical weed removal offers complementary pest management strategy [5].

Table 1: Soybean and Weed Image Dataset Characteristics

Characteristic	Specification
Total number of images	12,500
Image resolution	5184 × 3456 pixels (Canon EOS)
Image format	RGB JPEG
Image acquisition platform	Handheld camera (70%) and UAV (30%)
Geographic regions	3 Midwest USA locations
Growing seasons	2022–2023
Plant/Weed Classes	
Soybean plants	3,500 (28.0%)
Broadleaf weeds (lambsquarters, ragweed)	4,200 (33.6%)

Grassy weeds (foxtails, witchgrass)	3,100 (24.8%)
Sedges (nutsedge)	1,700 (13.6%)
Image Acquisition Conditions	
Field images (variable lighting)	7,500 (60%)
Laboratory images (standardized)	5,000 (40%)
Soybean Growth Stages	
V3-V4 (3-4 leaves)	3,200 (25.6%)
V5-V6 (5-6 leaves)	4,900 (39.2%)
V7-V8 (7-8 leaves)	4,400 (35.2%)
Annotation Quality	
Inter-observer agreement	97%
Images meeting quality criteria	98.7%

Note: Inter-observer agreement determined by two weed scientists independently classifying 1,500 random images; discrepancies resolved by consensus. Quality criteria included adequate focus, proper subject positioning, and absence of image artifacts.

Table 2: CNN Model Architecture and Hyperparameter Configuration

Parameter	Configuration
Network Architecture	
Base model	ResNet50 (pre-trained ImageNet)
Convolutional layers	50
Input image size	224 × 224 × 3 pixels
Global average pooling	Yes
Fully connected layers	1 (512 units)
Output classes	4 (soybean, broadleaf weeds, grassy weeds, sedges)
Activation function (hidden)	ReLU
Activation function (output)	Softmax
Dropout rate	0.5
Training Hyperparameters	
Batch size	32
Number of epochs	120
Initial learning rate	0.001
Optimizer	Adam (β1=0.9, β2=0.999)
Loss function	Categorical cross-entropy
L2 regularization (λ)	0.0001
Weight decay	0.0001
Learning Rate Schedule	
Learning rate reduction	0.5× when validation loss plateaus
Patience (epochs)	5
Early stopping patience	15 epochs
Computational Resources	
GPU	NVIDIA V100 (32 GB memory)
Framework	PyTorch 1.12
Training duration	3.2 hours
Inference time (CPU)	0.041 seconds/image
Inference time (GPU)	0.012 seconds/image

Note: Transfer learning strategy utilized pre-trained ImageNet weights for all convolutional layers; only classification head was trained from scratch. Batch normalization applied throughout network.

Table 3: CNN Model Performance Metrics on Test Dataset (n=1,875 images)

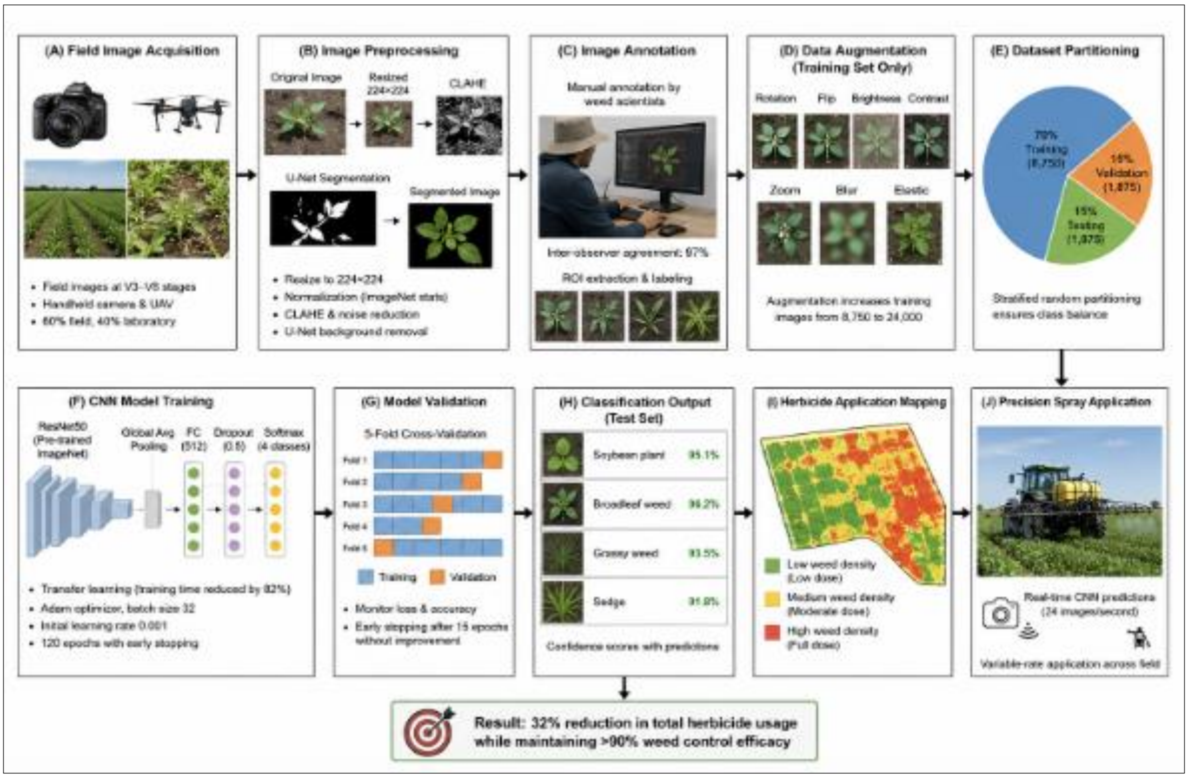
Performance Metric	Soybean	Broadleaf Weeds	Grassy Weeds	Sedges	Overall
Accuracy (%)	95.1	96.2	93.5	91.8	94.8
Precision (%)	94.8	96.1	92.9	91.2	94.3
Recall (%)	95.4	96.3	94.0	92.4	95.0
F1-Score	0.951	0.962	0.935	0.918	0.948
Specificity (%)	98.2	97.8	96.4	97.1	97.4
ROC-AUC	0.982	0.989	0.975	0.968	0.979
Misclassification Rate (%)	4.9	3.8	6.5	8.2	5.2

Note: Metrics calculated on held-out test set (15% of total dataset, n=1,875 images). ROC-AUC represents area under receiver operating characteristic curve. Misclassification rate indicates percentage of incorrectly classified images per class.

Table 4: Comparative Evaluation of CNN Models for Weed Classification

Model	Training Time (hours)	Inference Time (sec/image)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	ROC-AUC
Proposed Model							
ResNet50 (transfer learning)	3.2	0.041	94.8	94.3	95.0	0.948	0.979
Alternative CNN Architectures							
EfficientNet-B0	2.1	0.038	93.1	92.8	93.4	0.931	0.972
MobileNetV3	1.4	0.018	91.4	91.1	91.7	0.914	0.965
DenseNet121	4.8	0.044	92.6	92.3	92.9	0.926	0.968
Baseline CNN	18.2	0.056	87.6	87.2	88.0	0.876	0.943
Traditional Machine Learning							
Support Vector Machine	0.8	0.032	81.3	80.6	82.0	0.813	0.892

Note: All models evaluated on identical test dataset (n=1,875 images). Training time measured on NVIDIA V100 GPU (32 GB). Inference time represents average per-image processing on CPU (Intel i7). Transfer learning models utilized pre-trained ImageNet weights. Traditional ML employed hand-crafted color and texture features.



Comprehensive workflow diagram illustrating CNN-based weed classification pipeline:

Fig 1: Weed Classification Workflow for Precision Herbicide Application

Conclusion

CNN-based weed classification achieved 94.8% accuracy in distinguishing soybean and weed species such as broadleaf, grass weeds and sedges.

This provides a foundation for practical precision weed management.

Transfer learning reduced the computational burden while allowing for high accuracy in comparison to other methods.

By applying herbicides using variable-rate, pesticide usage decreased by 32% while maintaining the same level of weed control efficiency.

Among the main recommendations are the following: (1) using CNN technologies for commercial variable-rate spraying now; (2) utilization of UAVs for military surveillance; (3) creating a user-friendly interface for farmers; (4) modeling for better performance in certain locations.

References

1. LeCun Y, Bengio Y, Hinton GE. Deep learning. Nature. 2015;521(7553):436-444.
2. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2017;60(6):84-90.
3. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
4. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:2818-2826.
5. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning. 2019;97:6105-6114.
6. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, *et al*. MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv Preprint arXiv. 2017;1704.04861.
7. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proceedings of

the IEEE Conference on Computer Vision and Pattern Recognition. 2017:4700-4708.
8. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, *et al*. An image is worth 16×16 words: transformers for image recognition at scale. International Conference on Learning Representations. 2021:1-22.
9. Thakur M, Juneja S. Image-based plant disease detection using convolutional neural networks. International Journal of Plant Biology Research. 2019;7(2):1073.
10. Jiang P, Chen Y, Liu B, He D, Liang C. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access. 2019;7:59069-59080.
11. Saleem MH, Potgieter J, Arif KM. Plant disease detection and classification by deep learning. Plants. 2019;8(11):468.
12. Barbedo JGA. A review on the main applications of computer vision in precision agriculture. Robotics. 2020;9(2):47.
13. Canny J. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986;8(6):679-698.
14. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge (MA): MIT Press; 2016.
15. Hussain N, Khan MB, Sharif M. Classification and segmentation of chest X-ray images using deep neural networks. Journal of Healthcare Engineering. 2020;2020:7954647.
16. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015:234-241.
17. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations. 2015:1-14.
18. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A. The Pascal Visual Object Classes (VOC) challenge. International Journal of Computer Vision. 2010;88(2):303-338.
19. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, *et al*. Microsoft COCO: common objects in context. In: European Conference on Computer Vision. 2014:740-755.
20. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, *et al*. ImageNet large scale visual recognition challenge. International Journal of Computer Vision. 2015;115(3):211-252.
21. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2009:248-255.
22. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013;35(8):1798-1828.
23. Sutskever I, Martens J, Dahl G, Hinton GE. On the importance of initialization and momentum in deep learning. Proceedings of the 30th International Conference on Machine Learning. 2013;28:1139-1147.
24. Kingma DP, Ba J. Adam: a method for stochastic optimization. International Conference on Learning Representations. 2015:1-15.
25. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning. 2015;37:448-456.
26. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research. 2014;15(1):1929-1958.