Computers and Artificial Intelligence；Vol.2, No.3; 2025
E-ISSN: 3006 -3728 P-ISSN：3006-371X
DOI: https://doi.org/10.70267/cai.25v2n3.1019
Published by: Zeus Press

Meta-learning Driven Automatic Hyperparameter Optimization for Neural Networks in Computer Vision

Xin Zhang*

School of Landscape Architecture, Beijing Forestry University, Beijing 100083, China

**Corresponding author: Xin Zhang, E-mail: 1258498237@qq.com.*

Abstract

With the growing complexity of computer vision tasks and the explosive expansion of data scales, the optimization of neural network hyperparameters has become increasingly critical to model performance. Efficient hyperparameter optimization enables breakthroughs in tackling complex tasks and adapting to intricate scenarios. Traditional optimization methods rely heavily on manual experience and suffer from low efficiency. In contrast, meta-learning empowers hyperparameter optimization through its core logic of “learning to learn”: it extracts generalizable experience to build meta-cognition, thereby enhancing computational efficiency while achieving strong generalization capabilities. This paper first systematically introduces basic concepts such as neural networks, then elaborates on classic traditional optimization methods and meta-learning-based approaches, and then presents the experimental results of selected algorithms. Finally, it identifies unresolved issues and provides an outlook on future development trends.

Keywords

meta-learning, computer vision, hyperparameter optimization, MAML, reptile

1. Introduction

As artificial intelligence (AI) technology penetrates deeply into fields such as computer vision, natural language processing, autonomous driving, and medical imaging, deep learning—driven by the rapid advancement of computing power and the era of big data—has achieved revolutionary breakthroughs, transitioning from laboratory research to large-scale industrial applications. In recent years, model structures have become increasingly complex, and the scale of hyperparameters has expanded dramatically. The performance of these models depends not only on innovations in algorithm structures but also on key adjustment processes. Among these processes, the optimization of neural network hyperparameters directly affects the model’s generalization ability and training efficiency.

Tasks in the computer vision (CV) field require the extraction of hierarchical features from pixel data. These tasks share high commonality and transferability, while training on datasets incurs extremely high costs—and model performance is highly sensitive to hyperparameters. However, traditional manual hyperparameter tuning relies heavily on expert experience and lacks adaptability. To address this challenge, automated machine learning (AutoML) and hyperparameter optimization, which are based on meta-learning principles, have emerged. Snoek et al. (2012) successfully applied Gaussian processes to hyperparameter optimization, significantly reducing the number of training epochs, although the dimensionality of the hyperparameters remained high. On January 28, 2016, the British journal *Nature* reported that AlphaGo (Silver et al., 2016)—developed by DeepMind, a subsidiary of Google—defeated the European Go champion with a score of 5--0. AlphaGo adopted a deep neural network model, marking a major breakthrough in AI and verifying the great potential of deep learning. In 2017, Bello et al. (2017) demonstrated the feasibility of using neural networks to replace traditional optimization algorithms, providing an important source of inspiration. In the same year, Chen et al. (2017) reported that a recurrent network controller could be trained via reinforcement learning to learn automatic update rules, which offered significant insights for hyperparameter optimization. In industrial scenarios, computer vision applications have strict requirements for model performance; meta-learning-based optimization not only ensures model accuracy but also lowers technical barriers. Given this context, this paper briefly summarizes traditional hyperparameter optimization methods (including Bayesian optimization) and meta-optimization methods (including MAML), along with their experimental results.

2. Fundamental Concepts

2.1 Computer Vision

Computer vision (CV) uses computer systems to simulate and achieve human visual functions, thereby enabling analysis and decision-making for images and videos (Hu & Wang, 2025). It is an interdisciplinary field involving AI, signal processing, statistics, information theory, and computational mathematics. The human visual process can be simplified into three steps: the eyes form an image on the retina, the brain interprets the image, and a response is generated. Computers simulate these steps through three corresponding processes: imaging by imaging devices, data processing via computer systems, and the output of results (Pan & Zhang, 2005). In the early stages, owing to the limited data volume and computing power, CV relied on manual feature classification to recognize characters and images. With the development of image processing technology, methods such as support vector machines (SVMs) for feature extraction, machine learning models, and deep learning (e.g., convolutional neural networks (CNNs)) have emerged. These methods enable better implementation of daily applications, including image classification, object detection, image semantic segmentation, image caption generation, pedestrian separation, facial recognition, motion estimation, and image inpainting (Liu, 2019). Recent studies indicate that more than 80% of the more than 250 PBs of data generated globally every day consists of images and videos (Liu, 2025). Faced with the task of processing such massive data, there is an urgent need to automate information extraction.

2.2 Neural Networks

2.2.1 Overview of Neural Networks

A neural network is a computational model that simulates the connection structure, information transmission, and processing patterns of neurons in the human brain. Its primary goal is to analyze data and make predictions. Like the human brain, a neural network contains many "neurons" (basic computing units), which are connected through an input layer, hidden layers, and an output layer to form a multilayer network architecture. Information is transmitted between neurons by simulating the "synaptic strength" of biological neurons through weights. Since the advent of neural networks in the 1940s, hundreds of variants have been developed. In computer vision tasks, the core objective of neural networks is to extract features from image and video data. The following sections focus on two typical network paradigms that support this objective: the convolutional paradigm represented by the CNN and the attention paradigm represented by the transformer.

2.2.2 Convolutional Neural Networks (CNNs)

Traditional methods for processing image data attempt to capture global information, which leads to redundant data and heavy computational burdens. The introduction of "convolution" shifted the approach to extracting local image features. Convolution uses a "convolution kernel" (a fixed-size matrix) that slides over the input image with a specific stride. At each sliding position, a convolution operation is performed: the convolution kernel and the corresponding pixels of the image are multiplied point-by-point and summed. This process preserves local features and generates a feature map.

Convolutional neural networks (CNNs) were among the earliest deep neural networks to be successfully trained and widely applied. In a typical CNN architecture, convolutional layers and sampling layers (pooling layers) alternate. Within a convolutional layer, the same convolution kernel is shared across the same feature map. After multiple rounds of "convolution" and "sampling", the fully connected layer finally connects to the output target. During pooling layer sampling, max pooling is commonly used: it outputs the maximum pixel value of a region to enhance local features and focus on prominent textures. Alternatively, average pooling can be applied: it outputs the arithmetic mean of a region to retain more background information and achieve a more holistic, global perspective (Zhang et al., 2019). CNNs significantly reduce the number of connections between neurons, lowering computational costs while effectively preventing overfitting (Ji et al., 2022). Currently, many optimized variants have been developed on the basis of CNNs. For example, ResNet (residual network), proposed innovatively by Huang et al. (2004), introduced residual connections into deep convolutional neural networks, alleviating the vanishing gradient problem caused by excessively deep CNN models (Eliasmith & Anderson, 2003). The YOLO (you only look once) neural network model (Yi, 2021) adopts an end-to-end approach: it scans an image once to complete both the localization and classification tasks, significantly improving the detection speed (up to 45 FPS).

Studies have applied CNNs to classify 4 common fungal infectious skin diseases, achieving an accuracy rate of up to 93.3% (Nigat et al., 2023). Lee et al. (2023) used a CNN model and nearly 600 periapical X-rays of single premolars to predict endodontic treatment outcomes, making endodontic treatment decisions more scientific. These cases all demonstrate the practicality of CNNs in medical image analysis.

2.2.3 Transformer Neural Networks

The attention mechanism is inspired by human visual and cognitive processes—humans tend to focus on the most critical parts when making decisions (Ren & Wang, 2021). A CNN processes data via fixed weights and cannot flexibly adapt to task requirements. In contrast, the transformer—proposed by Vaswani et al.—uses an attention mechanism to dynamically adjust the degree of focus on different parts: it calculates the importance of each part to adjust the focus, assigning higher weights to key information. In essence, this is a form of dynamic weighting. Each layer of the transformer's encoder and decoder includes multihead self-attention (MHSA) (Miao, 2025). Through multiple parallel attention heads, information is captured from different perspectives, avoiding one-sidedness in information acquisition.

In recent years, the fusion of CNNs and transformers has become the mainstream framework for medical image segmentation. The CNN compensates for the transformer's high computational cost, whereas the transformer addresses the CNN's low flexibility. One study proposed a transformer-based cardiac MR image segmentation method with multiscale convolutional attention (TMCANet) (Li, 2025). Compared with existing network models in terms of Dice scores on the ACDC dataset, the TMCANet model achieved an accuracy of over 90% in left ventricle (LV) and right ventricle (RV) segmentation, indicating significantly improved segmentation performance. This finding verifies the value of their complementarity in complex medical image analysis.

2.3 Hyperparameters

Hyperparameters are configuration variables of a model or algorithm. They are set manually rather than learned during training and define the properties of the model (Li et al., 2022). Hyperparameters typically need to be determined in advance on the basis of the problem scenario, data characteristics, and other factors.

The learning rate is a key general hyperparameter that determines the step size of parameter updates during model training.

The hyperparameters of convolutional neural networks include the convolution kernel size, number of convolution kernels, convolution stride (Tong et al., 2018), number of network layers, and number of neurons per layer. The hyperparameters of transformer neural networks include the number of attention heads and the number of encoder layers.

2.4 Meta-Learning

The core of meta-learning is "learning to learn", first proposed by Maud. Metearning imitates humans' ability to learn quickly by referencing historical experience, enabling machines to make independent decisions. It is a set of algorithm frameworks (Y. Liu, 2022). A dataset consists of a training set and a test set. Meta-learning trains a model via the training set, evaluates the model via the test set, adjusts the model on the basis of the test results, and repeats the training and testing process. Finally, the model with the best performance on the test set is selected. Vilalta et al. proposed that meta-learning is an algorithm that dynamically improves bias through the accumulation of meta-knowledge, proving that the essence of meta-learning is to enable machines to learn to adjust parameters independently by existing experience (Vilalta & Drissi, 2002).

In recent years, emerging humanoid robots have faced complex and unpredictable situations in real life. Traditional machine learning methods require enormous computational costs to fully simulate all the scenarios. The emergence of meta-learning has endowed AI with self-learning and strong generalization capabilities (Li et al., 2021); for example, few-shot learning enables rapid adaptation to new tasks on the basis of limited experience. In the field of computer vision, meta-learning-based hyperparameter optimization methods have significant advantages in terms of autonomy over traditional methods.

3. Traditional Hyperparameter Optimization (HPO) Methods

3.1 Manual Tuning

Traditional manual hyperparameter tuning relies on the developer’s experience, knowledge, and task requirements to manually adjust the hyperparameters of a neural network for improved model performance. For example, in a CNN-based cat-and-dog image classification task, developers might manually set the first convolutional layer to 32 3×3 convolution kernels before training. If the model is underfit after training, the number of convolution kernels can be manually increased to 64 to enhance feature extraction capabilities. If cats or dogs in the images are large scale, they can manually adjust the convolution kernel size to 5×5 to expand the receptive field and extract information more efficiently. However, when the number of hyperparameters is large and they are interrelated (in complex scenarios), it is difficult to find the globally optimal parameter combination. The following sections introduce several algorithms that are more suitable for hyperparameter optimization: brute-force algorithms (grid search, random search) and Bayesian optimization algorithms.

3.2 Grid Search

Grid search finds the optimal hyperparameter combination by iterating through all predefined hyperparameter combinations and aligning them with the task objectives (Li et al., 2024). However, grid search has several drawbacks: as the number of hyperparameters increases, the number of hyperparameter combinations grows exponentially, leading to high computational costs from iteration; for continuous hyperparameters, predefinition fails to cover the entire search space, resulting in inaccurate results; and it cannot support parallel computing, increasing time costs.

3.3 Random Search

Randomly sample hyperparameter combinations from the hyperparameter search space and perform multiple searches to obtain combinations with good performance (Chen et al., 2020). Random search is very simple to implement and has low computational costs. However, this method has poor stability, lacks directionality, and may miss the globally optimal solution.

3.4 Bayesian Optimization

The hyperparameter optimization process can be regarded as a black-box optimization problem for generalization performance (Deng, 2019). Bayesian optimization builds a probabilistic model for the loss function. The Gaussian process (GP) serves as the basis of the classic probabilistic model for Bayesian optimization. Its essence is to learn information from loss function evaluations to predict the shape of the objective function, reduce the number of invalid attempts and samples, and ultimately find the minimum value of the loss function.

Recent studies have shown that the acquisition function (i.e., expected improvement) can control new possibilities on the basis of the potential improvement of parameter points to find the optimal result; to accelerate the startup speed of the Bayesian optimization algorithm, prior knowledge learned through meta-learning can be used to initialize Bayesian optimization (Peng, 2025). The following sections focus on methods that use meta-learning to empower hyperparameter optimization and improve the autonomy of machine learning.

4. Meta-Learning-Enabled Hyperparameter Optimization Methods

4.1 Gradient-based Meta-Learning Methods

4.1.1 MAML

Model-agnostic meta-learning (MAML) is a classic gradient descent-based meta-learning framework proposed by Finn et al. (2017). Its core goal is to use gradient information from meta-training tasks to aggregate gradient updates and initialize model parameters, enabling the model to adapt quickly to new tasks.

The core parameter optimized by the MAML framework is the initial parameter θ of the model. In the MAML framework, the dataset is divided into a meta-training set and a meta-test set. Small samples are selected from the meta-training set as tasks, and all tasks share the parameter θ—these tasks are used for pretraining sampling in MAML. The samples in one task are further divided into a support set and a query set (Deng & Zhang, 2024). The training process of MAML includes inner-loop optimization and outer-loop optimization.

In inner-loop optimization, a randomly initialized parameter θ is first used. A batch of tasks is sampled from the support set to train θ, and gradient descent is applied to minimize the loss function (Pang, 2023), resulting in the optimal parameter θ' for each task. This process can be expressed as Equation (1):

$$\theta' = \theta - \alpha \nabla_\theta L_T(f_\theta) \tag{1}$$

where α represents the inner-loop optimization learning rate and where L_T represents the loss of the support set data.

In outer-loop optimization, historical tasks and their optimized parameters θ' are sampled. The loss is calculated on the basis of the query set, gradient information is aggregated, and the parameter θ is updated. This process can be expressed as Equation (2):

$$\theta = \theta - \beta \nabla_\theta \sum_{T_i \sim T} L_{T_i}(f_{\theta'_i}) \tag{2}$$

where β represents the outer-loop optimization learning rate; the summation part represents the total loss of the historical tasks on the query set after obtaining θ'; and ∇_θ represents the gradient operator of the parameter θ.

Through the alternation of inner-loop and outer-loop optimization, a meta-parameter θ with good performance is obtained—i.e., an initial parameter that is as adaptable as possible to all tasks. When new tasks are solved, it is easier to fine-tune the parameters (Pang et al., 2023), improving the ability to quickly learn new tasks, reducing the number of update iterations, and greatly enhancing the optimization efficiency of new tasks.

The MAML algorithm obtains highly generalizable initial parameters through few-shot training, enabling rapid adaptation to new scenarios. In the field of computer vision, few-shot problems and the need to distinguish complex visual scenarios (e.g., object recognition under different lighting, angle, and environment conditions) are common challenges. Thus, the characteristics of MAML make it well suited for computer vision applications. Gao (2023) slightly improved MAML to classify skin lesions via the ISIC2019 skin lesion dataset, which contains 25,332 unique benign and malignant skin lesion images from over 2,000 patients, covering 8 different types of skin diseases (including melanoma and melanocytic nevi). The experimental results show that the improved MAML achieves an accuracy of 84.9%, a precision of 83.65%, and a recall of 84.1%, which far exceeds those of other meta-learning models. Dou (2022) addressed the few-shot dataset of ransomware by converting byte streams into grayscale images using their static features, extracting features with a convolutional neural network, and building an MAML model for virus classification. Under the 5-way, 5-shot task setting, both the test accuracy and severe accuracy reached over 90%.

4.1.2 FOMAML

To address the low computational efficiency of direct differentiation (second-order differentiation) of the meta-parameter θ in the outer-loop optimization of the MAML algorithm, the FOMAML (first-order MAML) algorithm was proposed (X. Liu, 2022). In outer-loop optimization, FOMAML differentiates the optimized parameter θ' of each historical task (first-order differentiation), thereby improving the computational speed.

4.1.3 Meta-SGD

Meta-SGD (meta-stochastic gradient descent), proposed by Li et al. (2017) from Huawei Noah's Ark Lab, is a meta-learning algorithm. Its core idea is to design hyperparameters such as the learning rate as learnable meta-parameters. The update method of meta-SGD is shown in Equation (3):

$$\theta_i{}' = \theta - \alpha \odot \nabla_\theta L_{D_{T_i}^{\mathrm{tr}}}(\theta) \tag{3}$$

where (Xie, 2024): ⊙ represents elementwise multiplication, indicating that the learning rate α is a vector—each parameter is multiplied by a separate learning rate, meaning that the learning rate for each parameter is learned. Obviously, Meta-SGD is more expressive than MAML is.

4.2 Reptile

Reptile is a lightweight and efficient meta-learning method. It addresses MAML's reliance on second-order derivatives by requiring only first-order gradient information (Xia & Cui, 2022). Like MAML, Reptile also uses a meta-training set and a meta-test set. It first initializes the parameter Φ, selects tasks from the meta-training set, performs K-step gradient descent on each task to fine-tune the parameter, and obtains a temporary parameter Φ'. To enhance the generalization of the parameter across all tasks, the difference between Φ' and Φ is calculated, the average of all differences is computed, and the global meta-parameter Φ is updated via this average difference.

Deng (2023) proposed a meta-learning-based open-world knowledge graph completion model via the reptile algorithm. Experiments were conducted on the open-world knowledge graph completion problem with K=4. The CNN was set to 3 layers, with a convolution kernel size of 3, 100 convolution kernels, and a pooling stride of 2. The experimental data revealed that the model achieved significant performance in terms of the MRR metric: it outperformed TCVAE by 20.9% on the WDtext dataset and by 10.9% on the DBPtext dataset.

4.3 Multitask Learning (MTL)

When the target task is influenced by numerous factors, multitask learning (MTL) (Zhang & Yang, 2022) shares features across related tasks, uses loss functions of multiple tasks to train multiple visual tasks simultaneously, and explores higher-level information between tasks (Shang et al., 2019). This enables the network model to more accurately capture the essence of the original task. There are two common MTL methods: hard parameter sharing and soft parameter sharing. In the hard parameter sharing mechanism, multiple tasks share the initial backbone network (e.g., CNN convolutional layers, transformer encoders) through hidden layers, and only the output layers are designed with separate parameters—this reduces the risk of overfitting (Long & Wang, 2015). In soft parameter sharing, an independent model is designed for each task, and constraints and corrections are applied to the parameters of different models (Fan, 2021).

In the field of computer vision, Playout et al. (2019) proposed integrating a detection task based on MTL to assist in the segmentation of retinal lesions. Ren and Lee (2018) used MTL to learn task-agnostic computer vision representations: a multitask learning framework based on AlexNet was trained to learn representations that capture high-level visual features by first training on three tasks (surface normal detection, depth estimation, and instance contour detection) and then transferring the learned representations to the VOC2007 and VOC2012 datasets for classification and detection tasks. Ye (2022) developed an improved MTL-based model that effectively combines the feature selection function of the SE module, achieving clearer contour segmentation results and estimation outcomes.

5. Conclusion and Outlook

5.1 Conclusion

In summary, this paper summarizes neural network-based hyperparameter optimization methods. First, it provides the relevant background of hyperparameter optimization and the meta-learning field. Second, it introduces concepts such as computer vision, neural networks, hyperparameters, and meta-learning. Third, traditional hyperparameter optimization methods are presented, and their advantages and disadvantages are summarized. Fourth, it elaborates on meta-learning-based hyperparameter optimization methods, including MAML and its variants, Reptile, and MTL, and provides experimental data to prove that meta-learning empowers a significant improvement in hyperparameter optimization performance, which is highly important for task processing.

5.2 Outlook

Although meta-learning-based automatic hyperparameter optimization methods have achieved remarkable progress, there are still issues to be optimized. For example, MAML suffers from unstable training—gradients are prone to explosion or vanishing after passing through multiple convolutional layers, the effect of batch normalization needs improvement, and the learning rate setting is rigid. MTL, on the other hand, is complex to implement: it requires simultaneous use of data from multiple tasks, and different loss functions need to be designed for different tasks to update gradients, resulting in low efficiency. Moreover, task transfer capability is limited, and exploring visual information to establish connections between tasks deserves further research.

Overall, despite the many unresolved issues, meta-learning-based hyperparameter optimization has remained a research hotspot for a long time. New methods and technologies will enable it to be more widely applied to object detection and recognition in complex scenarios (e.g., pedestrian detection) and real-time tasks requiring a balance between computing power and energy efficiency (e.g., autonomous driving, real-time video analysis).

References

Bello, I., Zoph, B., Vasudevan, V., & Le, Q. V. (2017). *Neural optimizer search with reinforcement learning* [Paper presentation]. Proceedings of the 34 th International Conference on Machine Learning, Sydney, Australia.

Chen, S. P., Wu, J., & Chen, X. Y. (2020). Hyperparameter optimization methods based on reinforcement learning. *Mini-micro Computer Systems, 41*(4), 679-684.

Chen, Y., Hoffman, M. W., Colmenarejo, S. G., Denil, M., Lillicrap, T. P., Botvinick, M., & Freitas, N. (2017). *Learning to learn without gradient descent by gradient descent* [Paper presentation]. Proceedings of the 34 th International Conference on Machine Learning, Sydney, Australia.

Deng, J. (2023). *Research on open-world knowledge graph completion based on meta-learning* [Master's thesis, Jilin University]. CNKI. https://doi.org/10.27162/d.cnki.gjlin.2023.005867.

Deng, S. (2019). Hyperparameter optimization method of CNN based on improved Bayesian optimization algorithm. *Application Research of Computers, 36*(7), 1984-1987. https://doi.org/10.19734/j.issn.1001-3695.2018.01.0021

Deng, T. Y., & Zhang, G. P. (2024). Research on improving small sample model generalization performance based on meta-learning and data augmentation. *Modern Information Technology, 8*(8), 93-96. https://doi.org/10.19850/j.cnki.2096-4706.2024.08.021

Dou, J. (2022). *Application research on computer virus classification based on adaptive N-gram algorithm and MAML model* [Master's thesis, Hainan Normal University]. CNKI. https://doi.org/10.27719/d.cnki.ghnsf.2022.000288

Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT Press.

Fan, X. (2021). *Research on pedestrian re-identification based on multi-task learning* [Master's thesis, Nanjing University of Posts and Telecommunications]. CNKI. https://doi.org/10.27251/d.cnki.gnjdc.2021.000407

Finn, C., Abbeel, P., & Levine, S. (2017). *Model-agnostic meta-learning for fast adaptation of deep networks* [Paper presentation]. Proceedings of the 34 th International Conference on Machine Learning, Sydney, Australia.

Gao, Z. (2023). *Research on automatic classification method of skin lesions based on deep meta-learning* [Master's thesis, Shandong Normal University]. https://doi.org/10.27280/d.cnki.gsdsu.2023.000695

Hu, H., & Wang, J. (2025). Countermeasures for the application of computer vision image processing technology. *Digital Design, 1*, 42-44.

Huang, G., Zhu, Q., & Siew, C.-K. (2004, 25-29 July 2004). *Extreme learning machine: a new learning scheme of feedforward neural networks* [Paper presentation]. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), Budapest, Hungary.

Ji, C., Gao, Z., & Qin, J. (2022). A review of image classification algorithms based on convolutional neural networks. *Computer Applications, 42*(4), 1044-1049. https://doi.org/10.11772/j.issn.1001-9081.2021071273

Lee, J., Seo, H., Choi, Y. J., Lee, C., Kim, S., Lee, Y. S., Lee, S., & Kim, E. (2023). An endodontic forecasting model based on the analysis of preoperative dental radiographs: A pilot study on an endodontic predictive deep neural network. *Journal of Endodontics, 49*(6), 710-719. https://doi.org/10.1016/J.JOEN.2023.03.015

Li, F. C., Liu, Y., Wu, P. X., 董方, 蔡奇, & 王哲. (2021). A review of meta-learning research. *Journal of Computer Research and Development, 44*(2), 422-446.

Li, H. X., Song, D. L., Kong, J. N., & et al. (2024). Evaluation of hyperparameter optimization techniques for traditional machine learning models. *Computer Science, 52*(8), 242-255.

Li, J. (2025). *Research on cardiac medical image segmentation method based on CNN-transformer hybrid network* [Master's thesis, Qilu University of Technology]. CNKI. https://doi.org/10.27278/d.cnki.gsdqc.2025.000715

Li, Y. R., Zhang, Y. L., & Wang, J. C. (2022). A review of Bayesian optimization methods for hyperparameter estimation. *Computer Science, 49*(S1), 86-92.

Li, Z., Zhou, F., Chen, F., & Li, H. (2017). Meta-SGD: Learning to learn quickly for few-shot learning. *arXiv preprint*, arXiv:1707.09835. https://doi.org/10.48550/arXiv.1707.09835

Liu, B. (2025). Development and application of artificial intelligence computer vision technology in the era of big data. *Information and Computer, 31*(5), 34-36.

Liu, X. (2022). *Research on automatic hyperparameter optimization based on reinforcement learning and meta-learning* [Master's thesis, University of Electronic Science and Technology of China]. CNKI. https://doi.org/10.27005/d.cnki.gdzku.2022.001575

Liu, Y. (2022). *Research on automatic hyperparameter optimization methods based on cognitive meta-learning* [Master's thesis, Nanjing University of Aeronautics and Astronautics]. CNKI. https://doi.org/10.27239/d.cnki.gnhhu.2022.001508

Liu, Z. (2019). On computer vision technology. *Digital User, 25*(8), 159.

Long, M., & Wang, J. (2015). Learning multiple tasks with deep relationship networks. *arXiv preprint*, arXiv:1506.02117. https://doi.org/10.48550/arXiv.1506.02117

Miao, J. (2025). *Research on convolutional neural network optimization and application based on transformer and collaborative attention mechanism* [Master's thesis, University of Electronic Science and Technology of China]. CNKI. https://doi.org/10.27005/d.cnki.gdzku.2025.002766

Nigat, T. D., Sitote, T. M., & Gedefaw, B. M. (2023). Fungal skin disease classification using the convolutional neural network. *Journal of Healthcare Engineering, 2023*, Article 6370416. https://doi.org/10.1155/2023/6370416

Pan, C., & Zhang, C. (2005). Brief introduction to computer vision. *Automation Exposition, 22*(5), 92-93,99.

Pang, H. Y. (2023). *Research on personalized recommendation algorithm based on meta-learning* [Master's thesis, Jilin University]. CNKI. https://doi.org/10.27162/d.cnki.gjlin.2023.002492

Pang, Y., Xu, H., Zhang, Y., Zhu, H. L., & Peng, X. (2023). Modulation recognition algorithm based on transfer meta-learning. *Acta Armamentarii, 44*(10), 2954-2963. https://doi.org/10.12382/bgxb.2022.0583

Peng, J. J. (2025). Hyperparameter optimization research on Bayesian optimization algorithm based on GP. *China New Technologies and Products,* (11), 38-40. https://doi.org/10.13612/j.cnki.cntp.2025.11.047

Playout, C., Duval, R., & Cheriet, F. (2019). A Novel Weakly Supervised Multitask Architecture for Retinal Lesions Segmentation on Fundus Images. *IEEE Transactions on Medical Imaging, 38*(10), 2434-2444. https://doi.org/10.1109/TMI.2019.2906319

Ren, H., & Wang, X. (2021). Review of attention mechanisms. *Computer Applications, 41*(S1), 1-6.

Ren, Z., & Lee, Y. J. (2018, 18-23 June 2018). *Cross-domain self-supervised multi-task feature learning using synthetic imagery* [Paper presentation]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA.

Shang, Q., Zhou, L., & Feng, L. (2019). Multi-task optimization algorithm based on denoising autoencoder. *Journal of Dalian University of Technology, 59*(4), 417-426. https://doi.org/10.7511/dllgxb201904013

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484-489. https://doi.org/10.1038/NATURE16961

Snoek, J., Larochelle, H., & Adams, R. P. (2012). *Practical bayesian optimization of machine learning algorithms* [Paper presentation]. Advances in neural information processing systems, Lake Tahoe, NE, USA.

Tong, W. G., Li, M. X., & Zhang, Y. K. (2018). Research on deep learning optimization algorithms. *Computer Science, 45*(S2), 155-159.

Vilalta, R., & Drissi, Y. (2002). A perspective view and survey of meta-learning. *Artificial Intelligence Review, 18*(2), 77-95. https://doi.org/10.1023/A:1019956318069

Xia, Y., & Cui, W. (2022). Personalized federated learning algorithm based on Reptile. *Computer Systems Applications, 31*(12), 294-300. https://doi.org/10.15888/j.cnki.csa.008875

Xie, C. (2024). *Research on few-shot learning image classification algorithms* [Master's thesis, Guizhou University of Finance and Economics]. CNKI. https://doi.org/10.27731/d.cnki.ggzcj.2024.000162

Ye, J. (2022). *Research on image pixel-level multi-vision task learning* [Master's thesis, Jiangxi University of Science and Technology]. CNKI. https://doi.org/10.27176/d.cnki.gnfyc.2022.000531.

Yi, D. (2021). *Research on YOLO-based target detection optimization algorithm* [Master's thesis, Nanjing University of Posts and Telecommunications]. CNKI. https://doi.org/10.27251/d.cnki.gnjdc.2021.000149

Zhang, S., Gong, Y., & Wang, J. (2019). Development of deep convolutional neural networks and their applications in computer vision. *Journal of Computer Science, 42*(3), 453-482. https://doi.org/10.11897/SP.J.1016.2019.00453

Zhang, Y., & Yang, Q. (2022). A survey on multi-task learning. *IEEE Transactions on Knowledge and Data Engineering, 34*(12), 5586-5609. https://doi.org/10.1109/TKDE.2021.3070203

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.