Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28277

A Look at the Use of the ResNet18 Model in Medical and Industrial Visual Checking

Zehou Liu

School of Automation, Central South University, Changsha, China
8207221825@csu.edu.cn

Abstract. As a basic deep residual network design, ResNet18 well handles the drop in performance during deep network training using its residual learning parts, and has become one of the key models widely used in medical imaging and industrial vision checking. This paper gives a full look at the use and tech changes of ResNet18 and its better versions in jobs like image sorting, object finding, and defect spotting. Looking at better ways for different use cases, researchers have done work from many sides, like adding attention tools (such as SE parts and SimAM), changing transfer learning plans, making data boost methods better, and model cutting and light design. The related results have shown good work in jobs like coronary artery narrowing sorting, road hole finding, metal surface defect checking, flame steadiness spotting, epilepsy EEG signal sorting, and coal gangue picking, with model correctness mostly staying between 94% and 99.5%. The paper also talks about real problems in current work, like issues with light setup, not enough few-shot use, limited cross-case fitting, and delays in making standard data sets. From this, it gives possible next steps like structure depth bettering, multimodal working together, and built-in use in more industrial cases, hoping to give a full tech guide for researchers in linked areas.

Keywords: ResNet18, deep learning, image recognition, attention system, transfer learning, simple network

1. Introduction

As AI technologies keep changing, deep learning is doing a more and more key job in image recognition and smart checking, causing big shifts in both medical diagnosis and factory automation. Of the many network designs out there, ResNet18 has found wide use as a base model for lots of vision jobs. Its skip link setup well handles gradient loss and performance drop often seen in deep network training, while its good depth and low cost make it very useful for real-world uses. In medical places, ResNet18-based image study helps doctors make choices by making diagnostic steps faster and more correct. In factory settings, vision check systems using this design have greatly boosted the automation and trust of finding flaws and judging quality. A full look at ResNet18 and its versions in these areas, with a study of improvement methods and performance results, gives helpful ideas for pushing the real-world use of AI technologies.

In recent years, researchers have changed the ResNet18 design to fit the needs of different use cases through many changes and improvements, getting good results. Using medical imaging as an example, Ye and colleagues [1] used transfer learning on ResNet18 for two-class sorting of coronary

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28277

artery stenosis images. By adjusting a ready-made model without big design changes, they got a sorting accuracy of 96.97%, much better than starting results and showing the model's ability to be used in medical areas. In industry uses, Ma et al. [2] dealt with the problems of road pothole sorting from laser images by adding a squeeze-and-excitation (SE) attention part into ResNet18, making an SE-ResNet18 design. This attention part changes channel-based features on the fly, making the network look at key areas. Their method reached 99.58% accuracy and an F1-score of 93.78% on pothole sorting jobs, showing the use of attention parts for making feature showing better. Dealing with the need for fast work in industry places, Xue and partners [3] looked into light setup plans. They used ResNet18 as the feature pulling base for the YOLOv3 finding system and used random cutting for model making smaller. This setup got 96.27% mean average precision with a per-picture work time of 45.5 milliseconds in coal gangue sorting uses, mixing high accuracy with better guess speed — a key need for use on limited embedded systems.

Together, these studies show the big ability of ResNet18 to change when used with other methods like transfer learning, attention parts, and model cutting. This paper looks at the steps forward of ResNet18 and its better versions in medical and work settings, checking the tech parts and how well different change ideas work. The goal is to give useful ideas for current studies and real use in similar areas.

2. Chapter 2 overview of related technologies

As an important part of the deep residual network family, ResNet18 has become a standard model in medical image analysis and industrial visual inspection because of its strong feature extraction ability and good computing speed. This chapter will look at different changed models based on ResNet18 and their use in many areas in recent years, focusing on their change methods, technical features, and results.

2.1. Application of ResNet18 model based on transfer learning

Transfer learning gives a useful way to handle the problem of small labeled data in specific areas, helping models work better when training examples are few. In a medical imaging study, Ye and colleagues looked at this method by using it for the classification of coronary artery stenosis with ResNet18. They adjusted a version of ResNet18 that was already trained on big datasets like ImageNet, changing its learned features to fit the medical image area. Their tests showed that this transfer learning method got a classification accuracy of 96.97%, doing better than a ResNet18 trained completely from zero by 8.08%. These findings show the worth of using pre-trained models in medical diagnostics, especially when dealing with small datasets, and suggest a real way to make reliable computer-aided diagnosis systems.

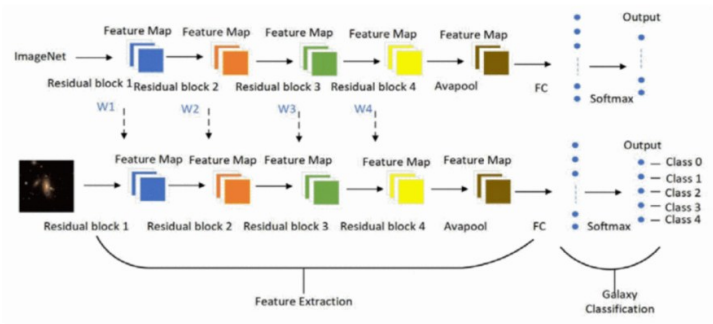


Figure 1. ResNet18 network model based on transfer learning [4]

DOI: 10.54254/2755-2721/2025.LD28277

2.2. Improved model with integrated attention mechanism

To make the model better at finding important parts, the attention method has been widely used to make the original ResNet18 design better.

2.2.1. SE-ResNet18 model

When dealing with road pothole detection from laser images, a common problem is that the targets are often small and the backgrounds very messy. To handle this, Ma et al. added a squeeze-and-excitation (SE) attention module into the ResNet18 design, making what they called SE-ResNet18. The thinking behind this plan is simple: by showing the links between different feature channels, the network learns to highlight those channels that have more helpful information while pushing down less important ones. This helped the model to pay attention more directly on pothole-related features. Their tests showed this — SE-ResNet18 got 99.58% accuracy and an F1-score of 93.78% on the classification job, doing better than several popular models. The work gives a strong example for fine-grained classification in tough visual settings [2].

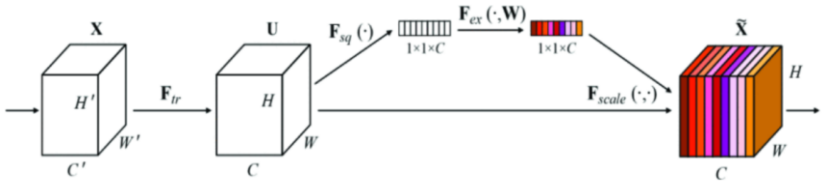


Figure 2. Channel attention module network module [5]

2.2.2. ResNet18-SimAM modle

In their study on boiler flame stability detection for power plant uses, Wang et al. put in a simple attention module called SimAM (Simple Parameter-Free Attention Module) into the lower parts of ResNet18. What makes SimAM important is that it gives value weights to single neurons without adding any extra learnable parts, which helps the model point out useful features while removing unimportant ones. The changed design, called ResNet18-SimAM, showed clear gains in tests: it learned more quickly during training and gave better results, getting 97.54% on the training set and 94.01% on the test set. This work gives a useful tool for watching equipment state and allowing early warnings in industrial settings [3].

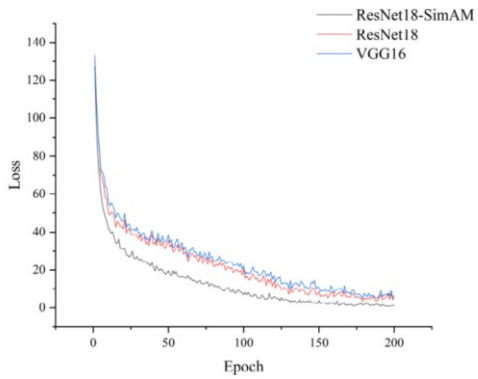


Figure 3. Loss curves of ResNet18-SimAM, ResNet18, and VGG16 during training [6]

DOI: 10.54254/2755-2721/2025.LD28277

2.3. A lightweight improved model for real-time detection

When putting models into use in industrial settings, things like inference speed and model size often matter as much as accuracy. With this in mind, Xue and his team made a simple detection model for coal gangue sorting robots, where real-time work is needed. They changed the first Darknet-53 backbone in the YOLOv3 detector to ResNet18, which quickly lowered the parameter number. To go even further, they used unstructured pruning to cut away extra connections, making a much smaller network. The final model found a good balance between speed and accuracy: it gave a mean average precision of 96.27%, handled each frame in only 45.5 milliseconds, and kept everything in a 65.34 MB size. Next to the basic YOLOv3, detection speed got better by 39.3%, while the parameter number fell by 71.4% — clear proof that the design fits the strict needs of real-time industrial uses and embedded deployment [7].

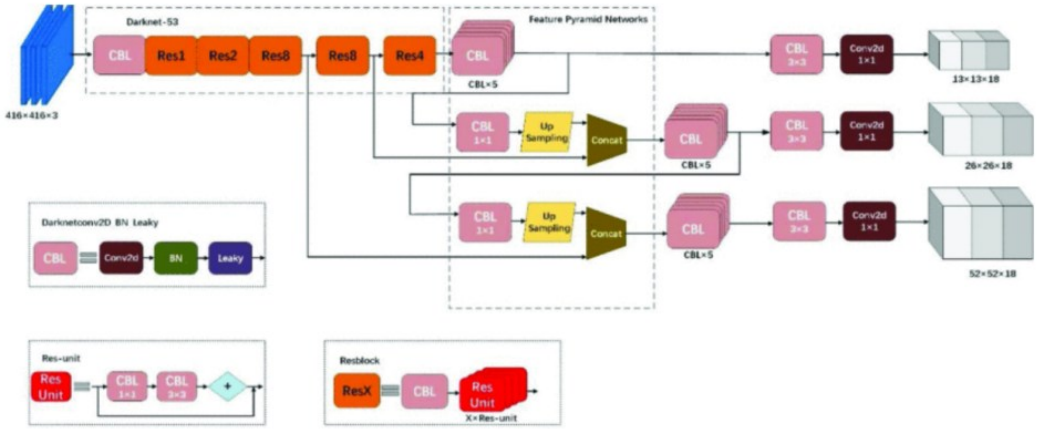


Figure 4. YOLOv3 network structure [8]

2.4. Application of special data preprocessing techniques

Not all raw data is ready for convolutional networks to use — especially in special fields where the original signals look nothing like natural images. In a study on classifying epileptic seizure phases from EEG signals, Dubey and colleagues worked on this by changing how the data was made ready. They began by cutting and putting together the raw EEG sequences to make the dataset bigger, then used continuous wavelet transform (CWT) to turn those one-dimensional time-series signals into two-dimensional time-frequency spectrograms. With that step, what started as a signal classification problem became an image classification job, good for standard CNN setups. The spectrograms were then given to a normal ResNet-18 model, which gave accuracies of 98.4%, 99.1%, and 98.2% across three classification cases (ictal versus interictal, normal versus ictal, and normal versus interictal). These results show how ResNet18, together with careful preprocessing, can work with data from very different types [9].

Taken together, the studies looked at here show that by putting together methods like transfer learning, attention parts, model cutting, and special data work, ResNet18 has shown it can be used for many different problems in medicine and industry. Whether the aim is better accuracy, quicker results, or a smaller model size, the design finds a useful middle ground — and in this way, gives helpful ideas for making vision systems that work well in real life.

DOI: 10.54254/2755-2721/2025.LD28277

Table 1. Comparison of applications of various models of ResNet18 model

Serial number	Model name	Improvement way	Application area	Accuracy/performance measures
1	Transfer Learning with ResNet18	Join transfer learning	Classification of coronary artery narrowing images	Accuracy 96.97%
2	SE-ResNet18	Introduce channel attention mechanism	Laser picture sorting of road surface holes	Accuracy 99.58%, F1 score 93.78%
3	Improved ResNet18	The convolutional layer is changed by three convolutions with a kernel size of two.	Looking for Flaw Depth in Metal Plates	Aluminum plate defect depth accuracy 98.1%, stainless steel 95.4%.
4	ResNet18-SimAM	Add the SimAM attention method	Testing the Stability of Power Plant Boiler Flames	Training set 97.54%, test set 94.01%
5	ResNet18	Data enhancement Continuous wave change	Classification of Epilepsy Seizure Stages (EEG)	Three groups classification accuracy: 98.4%, 99.1%, 98.2%
6	ResNet18-YOLO	Unstructured pruning YOLOv3 improvement	Finding and picking out coal gangue	mAP 96.27%, find time 45.5ms, model size 65.34MB

3. Introduction to application scenarios

ResNet18 and its versions work well for pulling out features and doing math fast. This has led to their use in many real jobs in medicine and industry. In both areas, they give useful tech answers made for particular problems.

3.1. Medical imaging assisted diagnosis

In healthcare, ResNet18 has been used to help diagnostic work in different imaging types, with the aim of improving both correctness and speed while making the job easier for doctors. In one paper, Ye and coworkers used the model for heart diagnostics. By changing a ready-made ResNet18, they made a system that could automatically sort coronary angiograms to find blood vessel narrowing, getting a correctness of 96.97% [1]. This kind of tool could help doctors in big screenings, helping find early disease signs and maybe making action faster. In another study, Dubey et al. took ResNet18's use to brain science [3]. They turned EEG signals into time-frequency pictures using continuous wavelet transform, then put these into a normal ResNet18 to sort seizure states — ictal, interictal, and normal. The model did well in three sorting jobs, with correctness hitting 99.1%. This shows a more clear, automatic way for epilepsy diagnosis, which could have real medical good.

3.2. Industrial vision and automated inspection

In industrial settings, ResNet18 and its improved versions work much like automated inspectors on the factory floor, allowing smart detection of product flaws and equipment status.

DOI: 10.54254/2755-2721/2025.LD28277

3.2.1. Infrastructure maintenance

Ma and colleagues worked on road condition monitoring by adding a channel attention system into ResNet18, leading to an SE-ResNet18 model [2]. Made to sort laser-scanned road surface images, the model showed very good at spotting potholes, getting 99.58% accuracy. This type of automatic check can help preventive maintenance plans — pointing out possible dangers early to make traffic safety better and cut repair costs.

3.2.2. Smart manufacturing and non-destructive testing

In making things, She et al. used a better ResNet18 with a four-coil eddy current sensor to check defect depth in metal sheets [3]. Their method could measure defect depth in aluminum and stainless steel plates with good precision (98.1% and 95.4% accuracy, each). As a non-contact testing way, it looks good for fields like aerospace and car making, where material soundness is key.

3.2.3. Energy production safety monitoring

Wang and colleagues used a ResNet18-SimAM model to watch flame stability inside power plant boiler furnaces [6]. By looking at flame images in real time, the model checked burning conditions and got 94.01% accuracy on test data. Uses like this help make plant work safer and more effective, helping stop burning-related problems while meeting energy-saving aims.

3.2.4. Resource sorting and industrial robotics

In the area of getting back resources, Xue et al. made a simple detection method using ResNet18 and YOLO for coal gangue sorting [7]. Built into robot sorters, the system finds and splits gangue from coal as it goes, getting a mean average precision of 96.27% with a per-frame processing time of only 45.5 milliseconds. Besides making sorting better and faster while lowering the need for workers, this work also helps more smart and cleaner ways in coal mining.

4. Challenge analysis

4.1. Balancing model efficiency with real-time demands

The main problem: Many factory uses — like sorting robots or inline quality checks — must react in milliseconds. Fast models, though, often have big designs and many parameters [7]. That means high computing needs and slower results, which makes them difficult to use on small or edge devices where resources are low.

What helps: Model compression methods like pruning [6], quantization, or knowledge distillation can cut down both size and compute. Also, designing simpler networks — say, using depthwise separable convolutions — and using hardware acceleration can make inference speeds faster. The goal is to find a good balance between accuracy and speed.

4.2. Handling scarce domain data and skewed class distributions

The main problem: In medicine and some special industrial areas, getting lots of good labeled data is hard and costly. Also, normal samples are usually common — faulty ones are few. This unevenness often makes the model favor the bigger group, making it tough to catch the small signs of the smaller group (like real flaws). The outcome? Poorer generalization and worse detection accuracy [1,9].

What helps: Data augmentation — simple steps like turning, changing size, or adding noise to current images — can make a small dataset go further. To handle class imbalance, there are some

DOI: 10.54254/2755-2721/2025.LD28277

ways: resampling (either adding more rare samples or reducing common ones), changing the loss function to treat classes in a different way, or even using generative models (GANs) to make real-looking examples of the rare class.

4.3. Improving generalization and adaptability to new settings

The main problem: A model trained under certain conditions — for example, fixed lighting or a certain material type — can lose its ability when moved to a new setting [2,3]. For example, a defect detector set for one kind of metal plate might fail on another. This lack of flexibility limits how much the model can be used.

What helps: Training on more different sets of data is a start. After that, area change methods and moving learning can help the model get used to new target situations with less trouble. The idea is to build models that stay strong when the real world gives surprises.

4.4. Managing computational and deployment costs

The main problem: Improved models — especially those with attention parts — usually need more from hardware. That means higher computer power and memory, which can push up costs [6,7]. For many small and medium-sized businesses, this becomes a true block to use.

What helps: One way is to keep pushing algorithm improvements to make changes simpler. Another is to rethink how things are used — like splitting work between the cloud and edge devices. Heavy thinking can run in the cloud, while the edge does data gathering and easier work. That kind of mixed setup can help keep costs low.

5. Conclusion

This paper systematically reviews the application and technical improvements of ResNet18 and its variants in key fields such as medical imaging and industrial vision. As a classic residual network architecture, ResNet18 effectively alleviates the degradation problem in deep network training through its skip connection design, while maintaining a favorable balance between model depth and computational cost. By integrating techniques such as transfer learning, attention mechanisms, model compression, and data preprocessing, researchers have significantly enhanced the performance of ResNet18 across a range of tasks, including coronary artery stenosis classification, road pothole detection, metal surface defect inspection, flame stability monitoring, epileptic EEG signal classification, and coal gangue recognition. Experimental results show that these improved models consistently achieve accuracy rates between 94% and 99.5%, demonstrating strong adaptability and practical value.

Despite the encouraging progress, current research still faces several challenges, including difficulties in balancing model complexity with real-time deployment, limited generalization under scarce or imbalanced data conditions, and high computational costs that hinder adoption in resource-constrained scenarios. Future work may focus on developing more efficient architectures, exploring multimodal learning strategies, strengthening the adaptability of models across different application scenarios, and advancing the deployment of lightweight models in real-world industrial and clinical settings. Overall, ResNet18 remains a highly versatile baseline, and its continued evolution will play an important role in promoting intelligent vision systems toward broader and more practical applications.

DOI: 10.54254/2755-2721/2025.LD28277

References

[1] X. Ye, Y. Jin, Z. Wang, N. Feng, L. Mu and G. Xie, "Image Binary Classification of Coronary Artery Stenosis Based on Resnet18 and Transfer Learning," in 2023 35th Chinese Control and Decision Conference (CCDC), Yichang, China, 2023, pp. 2721-2726.

[2] R. Ma, M. Xu, Z. Li and X. Sun, "Classification of Pothole Laser Images Based on the Improved ResNet18 Network," in 2024 3rd International Conference on Electronics and Information Technology (EIT), Chengdu, China, 2024, pp. 619-623.

[3] S. She et al., "Evaluation of Defects Depth for Metal Sheets Using Four-Coil Excitation Array Eddy Current Sensor and Improved ResNet18 Network," in IEEE Sensors Journal, vol. 24, no. 12, pp. 18955-18967, 15 June15, 2024.

[4] Y. Guo, P. Liu, X. Mei and Z. Yang, "The Research on Galaxy Image Classification Model Based on ResNet18 and Transfer Learning," 2024 8th International Workshop on Control Engineering and Advanced Algorithms (IWCEAA), Nanjing, China, 2024, pp. 139-142, doi: 10.1109/IWCEAA63616.2024.10823737.

[5] N. Shahadat, "Lung Image Analysis using Squeeze and Excitation Based Convolutional Networks," 2023 26th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2023, pp. 1-6, doi: 10.1109/ICCIT60459.2023.10441501.

[6] Z. Wang, Y. Bian, M. Yang and G. Liu, "Power Plant Furnace Flame Stability Detection Based on ResNet18-SimAM," in 2024 IEEE Sustainable Power and Energy Conference (iSPEC), Kuching, Sarawak, Malaysia, 2024, pp. 321-325, doi: 10.1109/iSPEC59716.2024.10892397.

[7] G. Xue, S. Li, P. Hou, S. Gao and R. Tan, "Study on a Resnet18-Based Lightweight Recognition ALgorithm of Coal and Gangue," in 2024 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI), Kusatsu, Japan, 2024, pp. 83-87, doi: 10.1109/IIKI65561.2024.00024.

[8] Y. Yu, S. Yan and X. Hao, "Pedestrian Detection Based on Improved YOLOv3 Network," 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT), Jilin, China, 2023, pp. 297-301, doi: 10.1109/ICCECT57938.2023.10141270. keywords: Roads;Feature extraction;Real-time systems;Periodic structures;Pedestrian detection;YOLOv3;Convolutional Block Attention Module;SPP Structure,

[9] V. K. Dubey, S. Sarkar, R. Shukla, G. Singh and A. Sahani, "Epileptic Seizure Stage Classification from EEG Signal Using ResNet18 Model and Data Augmentation," in 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-5, doi: 10.1109/DELCON54057.2022.9753352.