SSDC-MiniNet: 소형 강철 표면 결함 분류 모델

Adam Pradipta[1,2], David J. Richter[2], 김경백 [2]
[1]Dept. of Industrial and Systems Engineering, Institute of Technology Sepuluh Nopember, Indonesia
[2] 전남대학교, 인공지능융합학과

adampradipta92@gmail.com, david_richter@jnu.ac.kr, kyungbaekkim@jnu.ac.kr

SSDC-MiniNet: A Tiny CNN Steel Surface Defect Classifier

Adam Pradipta[1,2], David J. Richter[2], Kyungbaek Kim[2]
[1]Dept. of Industrial and Systems Engineering, Institute of Technology Sepuluh Nopember, Indonesia
[2]Dept. of Artificial Intelligence Convergence, Chonnam National University, South Korea

Abstract

Steel surface defect classification is important in producing an optimal product. However, worker errors while classifying the metal surface can lead to issues in production. So, recently researchers found a way to invest more in Industry 4.0 technology with methods such as Convolutional Neural Networks (CNNs). But models can be too large to be implemented in the factory. In response, we proposed a tiny model called SSDC-MiniNet with only 23,588 parameters and 0.105 GFLOPS. The model was trained on the North Eastern University (NEU) dataset, which consists of 1,800 images of six common metal surface defects. After testing the model, we achieved a competitive accuracy of 98.67% with only around 1% of the parameter count and flops.

1. Introduction

Quality control in metal surfaces is extremely important for ensuring that the final product is strong, and reliable [4]. Steel defects can be one of the main causes of an increase in production cost, monitoring the product during production is inevitable. The defect can, for example, be caused by inadequate facility conditions [3]. Manual human inspection systems can be suboptimal because they are time consuming and less automatic so they can create fatigue and stress levels that can lead to errors in the monitoring process [2,4]. As technology is growing fast the implementation of intelligent systems has gotten more attention in measuring accuracy and correctness. Surface quality inspection nowadays often uses computer vision-based technology [4]. Computer vision-based approaches are used to handle this kind of situation; however, most models require large datasets which can be laborious and financial costly [1]. However, defect classification datasets are often small [5]. CNNs are one of the best methods in solving this problem [2], as they can extract features from the image automatically. In order to train our model, we used the North Eastern University (NEU) dataset, which contains 1,800 images that classify into six types of defects [3,4] (with augmentation generation for training) and compare it to DenseNet121 [7], ResNet50 [9], and MobileNet (with different alpha scaling) [8].

2. Methodology

The proposed SSDC-MiniNet is based on the ideas of depthwise-separable (DS) convolutions [8] and squeeze and excitation (SE) attention [10]. DS convolutions offer a more parameter efficient alternative to normal 2D convolutions by combining and stacking depthwise and pointwise

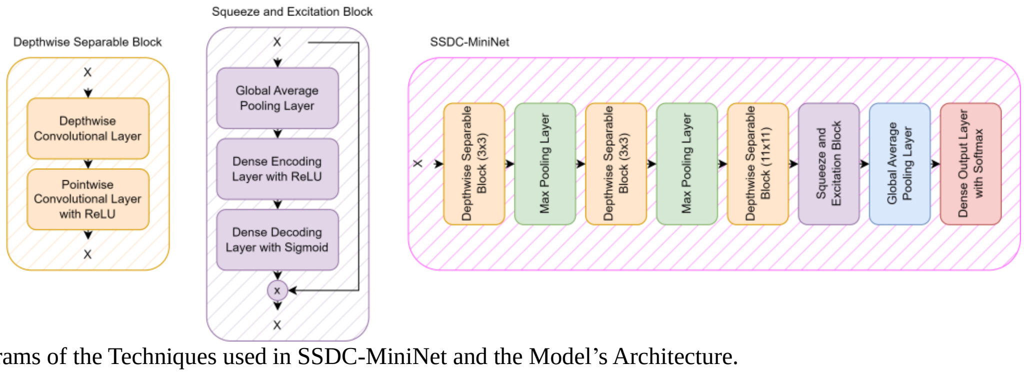

Figure 1: Diagrams of the Techniques used in SSDC-MiniNet and the Model's Architecture.

Table 1: Comparison of different model sizes and performances. Note that the inference speed was measured with a NVIDIA GeForce RTX 3090 and might be more significant on less powerful devices.

Model	GFLOPS	Param.	Memory	Train Acc.	Test Acc.
DenseNet121	5.7	7,043,654	26.87 MB	**96.47%**	**98.67%**
ResNet50	7.75	23,600,006	90.03 MB	81.80%	83.33%
MobileNet	1.15	3,235,014	12.34 MB	87.67%	**98.67%**
MobileNet α=0.1	**0.013**	40,499	00.16 MB	93.13%	96.00%
MobileNet α=0.25	0.064	220,086	00.80 MB	95.93%	94.67%
SSDC-MiniNet	0.105	**23,588**	**00.09 MB**	96.00%	**98.67%**

convolutions in order. SE assesses the importance of each channel and weights them accordingly, leading to improved performance. Additionally, in the last layer we utilize a large 11x11 kernel for a better receptive field in the rather large last layer in terms of width and height. These methods result in a model with only just over 23,500 parameters and only around 0.1 Giga FLOPS, making SSDC-MiniNet significantly more lightweight than other CNN. See Figure 1 for the model architecture and Table 1 for model size and comparison to the baseline models. The NEU dataset, comes with 6 classes of 300 images each (see Figure 2). To ensure no overfitting we apply normalization to all data and augmentations (rotation, width & height shift, horizontal shift) via data generator to the train set only, the test and validation sets remains un-augmented. Train test split is set to 1,500 images training and 150 each for validation and testing. Validation data was not used in these experiments. All parameters used for training can be seen in Table 2.

Table 2: Training configuration.

Parameter	Value
Number of Classes	6
Training Data	1,500 (250 per class + augmentation generator)
Validation Data	150 (25 per class, no augmentation)
Test Data	150 (25 per class, no augmentation)
Optimizer	Adam
Loss	Categorical Crossentropy
Number of Epochs	50
Input Size	200x200x3
Batch Size	32
Callback	Checkpoint (Test. Acc. Best Model Evaluation)
Machine	RTX 3090, Ryzen 7 5800X, 64GB RAM

3. Results

When looking at the accuracy scores (Table 1) obtained by our model as well as the baseline models used for comparison one can easily observe that SSDC-MiniNet achieves accuracy scores on the unseen and un-augmented test dataset in line with DenseNet121 and full MobileNet and even outperforms the ResNet50 model by over 15%. It does all that while only being around 1% of the parameter size of full MobileNet architecture (and even less than that when compared to DenseNet121 and ResNet50). Additionally, when comparing the GFLOPS and the memory footprint generated by the model architecture, it becomes obvious how much more parameter efficient SSDC-MiniNet is compared to the other full sized models. When compared with the small alpha MobileNets, parameter count still is lower, but GFLOPs are a little higher, but SSDC-MiniNet manages to outperform both small MobileNets. Our proposed model is not out-performed by any models used in this work and manages to outperform three. This shows that our model can learn robust features maps even on relatively small datasets and operate on them at a high performance level in terms of accuracy while doing so very efficiently.

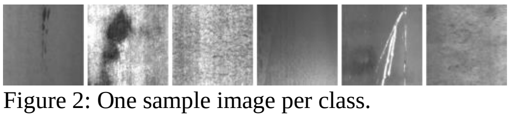

Figure 2: One sample image per class.

4. Conclusion

In this paper we propose the parameter and computationally efficient SSDC-MiniNet with only 23,588 parameters and 0.105 GFLOPS. This model achieved 98.67% accuracy on the NEU dataset, falling in line with the much larger DenseNet121, ful MobileNet and even outperforming the ResNet50 and smaller MobileNet alpha models. This proves that out model can reach high level performance in line with large scale models while being much smaller, even in scenarios with smaller datasets.

Acknowledgements

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2025-RS-2022-00156287, 33%). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629, 34%) grant funded by the Korea government(MSIT). This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry(IPET) through the Agriculture and Food Convergence Technologies Program for Research Manpower development, funded by Ministry of Agriculture, Food and Rural Affairs(MAFRA)(project no. RS-2024-00397026, 33%).

Referenecs

[1] Liu, Yang, et al. "A light-weight deep-learning model with multi-scale features for steel surface defect classification." *Materials* 13.20 (2020): 4629.

[2] Lee, Soo Young, et al. "Steel surface defect diagnostics using deep convolutional neural network and class activation map." *Applied Sciences* 9.24 (2019): 5449.

[3] Song, Kechen, and Yunhui Yan. "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects." *Applied Surface Science* 285 (2013): 858-864.

[4] Islam, M. F., & Rahman, M. M. (2018). Metal surface defect inspection through deep neural network. In *2018 International Conference on Mechanical, Industrial and Energy Engineering, ICMIEE* (p. 258).

[5] Natarajan, Vidhya, et al. "Convolutional networks for voting-based anomaly classification in metal surface inspection." *2017 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2017.

[7] Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[8] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).

[9] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[10] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.