ISSN: 2508-7894
KJAI website: http://www.kjai.or.kr
doi: http://dx.doi.org/10.24225/kjai.2023.11.3.17

Multi-Scale Dilation Convolution Feature Fusion (MsDC-FF) Technique for CNN-Based Black Ice Detection

Sun-Kyoung KANG[1]

Received: June 27, 2023. Revised: July 05, 2023. Accepted: September 05, 2023.

Abstract

In this paper, we propose a black ice detection system using Convolutional Neural Networks (CNNs). Black ice poses a serious threat to road safety, particularly during winter conditions. To overcome this problem, we introduce a CNN-based architecture for real-time black ice detection with an encoder-decoder network, specifically designed for real-time black ice detection using thermal images. To train the network, we establish a specialized experimental platform to capture thermal images of various black ice formations on diverse road surfaces, including cement and asphalt. This enables us to curate a comprehensive dataset of thermal road black ice images for a training and evaluation purpose. Additionally, in order to enhance the accuracy of black ice detection, we propose a multi-scale dilation convolution feature fusion (MsDC-FF) technique. This proposed technique dynamically adjusts the dilation ratios based on the input image's resolution, improving the network's ability to capture fine-grained details. Experimental results demonstrate the superior performance of our proposed network model compared to conventional image segmentation models. Our model achieved an mIoU of 95.93%, while LinkNet achieved an mIoU of 95.39%. Therefore, it is concluded that the proposed model in this paper could offer a promising solution for real-time black ice detection, thereby enhancing road safety during winter conditions.

Keywords : Black ice detection, Convolutional Neural Network (CNN), Multi-scale dilation convolution, Feature fusion, Road safety

Major Classification Code: Artificial Intelligence

1. Introduction

Road safety is a critical concern worldwide, with millions of accidents occurring annually due to various factors such as adverse weather conditions. Among these, black ice is one of the most hazardous and unpredictable road conditions, particularly during winter seasons. Black ice refers to a transparent and highly slippery layer of ice that forms on road surfaces, with the underlying road surface visible through the layers of ice. On dark roads such as asphalt, black ice appears black, hence its name is originated from it. Black ice on roads visually resembles a wet surface and is not easily identifiable to the naked eye as shown in Fig. 1. This phenomenon poses a significant threat to motorists, leading to a higher risk of accidents, injuries, and even fatalities (Park et al., 2017, Smith et al., 2017).

Black ice forms when the temperature drops below 0℃, resulting in a thin layer of ice formed by moisture on the

* This paper was supported by Wonkwang University in 2023.
1 First Author & Corresponding Author. Professor, Department of Computer Engineering, Wonkwang University, Korea. Email: doctor10@wku.ac.kr

road surface. It primarily occurs in shaded areas such as mountainous regions, roads with many trees, tunnel entrances, and bridge road surfaces where sunlight does not reach (Korea Traffic Accident Analysis System). As 70% of the Korean land is mountainous, many roads are shaded, making it easier for black ice to form on the road surface during winter. According to statistics on winter traffic accidents from 2015 to 2019 released by the Korean National Police Agency, out of a total of 7,236 accidents, there were 186 fatalities due to snow-related accidents and 706 fatalities due to black ice, which is 3.8 times higher than the fatalities from snow-related accidents (Korea Traffic Accident Analysis System). Additionally, accidents caused by black ice often result in secondary and tertiary damages, leading to major disasters and an increased likelihood of chain collisions. While recent advancements in road condition detection-based early warning systems for traffic safety have gained attention, comprehensive research on black ice detection is limited.

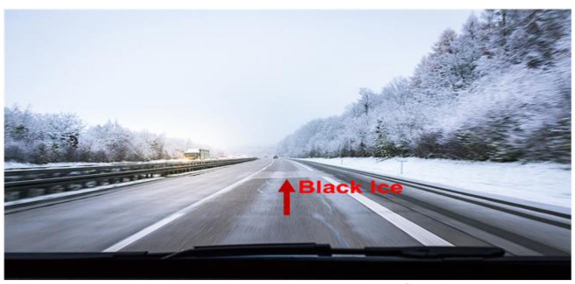

Figure 1: Black ice on road surfaces in real-world scenario

Detecting black ice in real-time is a challenging task due to its elusive nature and the potential dangers it presents to road users. Traditional methods for black ice detection rely heavily on weather forecasting systems and ground-based sensors, which may not provide accurate and timely information. Therefore, due to the difficulty in detecting black ice, there is a growing need for advanced technologies that can enable real-time black ice detection and assess the presence of black ice on the road, thereby providing timely warnings to drivers. Moreover, by implementing black ice detection devices, driver safety can be ensured, and the occurrence rate of traffic accidents can be reduced.

Black ice, a transparent ice layer on road surfaces, is notoriously difficult to detect due to its minimal visual cues. Traditional computer vision methods struggle to capture the subtle features associated with black ice, thus necessitating the need for more sophisticated techniques. In recent years, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable success in various computer vision tasks, including image recognition, object detection, and semantic segmentation. The ability of CNNs to automatically learn and extract relevant features from raw data makes them a promising approach for black ice detection on road surfaces (Krizhevsky et al., 2012).

In this paper, we leverage the capabilities of CNNs and introduce a multi-scale dilation convolution feature fusion (MsDC-FF) technique to enhance black ice detection accuracy. Thus, the proposed model can achieve real-time and accurate detection of black ice. The proposed technique incorporates multi-scale dilation convolution (MsDC), which involves applying dilation filters of varying sizes to capture features at different scales. By adjusting the spacing within the kernel, we increase the diversity of feature extraction. The fusion of these multi-scale dilation convolution features enhances the CNN's ability to identify black ice regions effectively.

In the following sections, we will describe the architecture of our black ice detection system in detail with methodology. The remainder of this paper is organized as follows: Section 2 provides an overview of related work on black ice detection and the application of CNNs in computer vision tasks. Section 3 presents the methodology and architectural details of our proposed in detail. In Section 4, we will discuss the experimental setup, including the construction of our thermal road black ice dataset. Furthermore, it is concluded that the results of extensive evaluations are presented comparing the performance of our system with existing image segmentation models.

2. Related Work

Many studies have explored black ice detection using image-based methods, Q. Lin et al. (2017) designed a road icing detection system based on OpenCV+Python, and a Support Vector Machine (SVM) classifier was employed to identify four types of road conditions: dry, wet, snowy, and icy. Lee et al. (2020) created a black ice detection dataset using Google image search and utilized CNN deep learning techniques to detect dry, wet, snowy, and black ice conditions, achieving a recognition rate of 96%. These research works highlight the effectiveness of using CNN-based methods for black ice detection.

With the widespread availability of cameras, images have become a convenient, fast, and cost-effective method for acquiring road information. Meanwhile, research on camera-based road surface black ice monitoring remains relatively limited. Therefore, this paper adopts an image-based road black ice detection with a deep neural network model. The following research works on image detection based on each proposed deep neural network are considered in this paper for performance comparison purpose to the proposed network model. These model are FCN, U-Net, DeepLabv3+, PSPNet, ENet, and LinkNet. The summary of approaches adopted in these models are as following: 1) FCN: Long et al. (2015) introduced a Fully Convolutional Network with imprecise edge segmentation due to information loss during down-sampling and up-sampling. 2) U-Net: Ronneberger et al. (2015) proposed an encoder-decoder model with concatenation of feature maps for improved segmentation. 3) DeepLabv3+: Chen et al. (2017) revisited DeepLabv3+ with dilated convolutions, ASPP, and CRF post-processing for accurate segmentation. 4) PSPNet: Zhao et al. (2017) combined ResNet and dilated networks with parallel pooling and up-sampling for scale-specific feature information and refined segmentation. 5) ENet: Paszke et al. (2016) introduced an efficient neural network with bottleneck modules and filter decomposition for precise segmentation and reduced complexity. 6) LinkNet: Chaurasia and Culurciello (2017) proposed a network with residual modules for enhanced feature extraction and optimized computation in real-time segmentation.

3. Methodology

This paper proposes following key construction background. 1) CNN-based black ice detection using thermal image: Thermal cameras are capable of detecting temperature variations on road surfaces, allowing us to distinguish regions with black ice from non-icy areas. By exploiting the thermal signatures of black ice, our system enhances the detection accuracy, even in challenging visual conditions (Breckon & Fisher, 2012). 2) Black ice dataset construction: Based on the research work on a comprehensive dataset of thermal road black ice images in (Wang et al., 2019; Kim et al., 2021), we construct black ice dataset that includes diverse formations of black ice on different road surfaces, such as cement and asphalt. This dataset serves as a valuable resource for training and benchmarking the performance of our proposed system. 3) Multi-scale dilation convolution feature fusion: To enhance the accuracy of black ice detection, we introduce a novel technique called multi-scale dilation convolution feature fusion (MsDC-FF). This technique adapts the dilation ratios of convolutional filters based on the resolution of input images. By incorporating multi-scale contextual information, our approach improves the system's ability to detect and segment black ice areas accurately (Huang et al. 2017).

3.1. Network Structure

This paper introduces a CNN-based architecture for real-time black ice detection with an encoder-decoder network for infrared images. Based on constructed infrared black ice road data set, our CNN-based model is trained for establishing a comprehensive dataset of thermal road black ice images for a training and evaluation purpose.

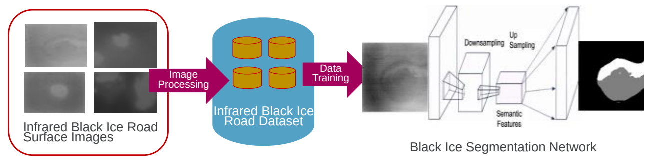

Figure 2: Illustration of black ice semantic segmentation training module

For enhancing the accuracy of black ice detection, a multi-scale dilation convolution feature fusion (MsDC-FF) technique is proposed. With adjustment of the dilation ratios based on the input image's resolution, this MsDC-FF technique can improve the network's ability. Figure 2 shows the illustration of black ice semantic segmentation training module.

3.2. Multi-Scale Dilated Convolution Feature Fusion (MsDC-FF) Network Framework

The multi-scale dilated convolutional feature fusion module is utilized to fully integrate low-level features and high-level features, and thereby it establishes a multi-scale convolutional attention module. A multi-scale information can help resolve ambiguous boundaries and produce more robust extraction results. In particular, thermal images have ambiguous boundaries and weak contrast features compared to visible light images, making it suitable to propose an encoder-decoder deep learning model based on multi-scale dilated convolution feature fusion.

The network architecture proposed in this paper is divided into two parts: an encoder and a decoder. The encoder consists of three stages of encoder blocks, while the decoder consists of four stages of decoder blocks, as shown in Figure 3. To reduce the model size, the early stages of the encoder block use two convolution layers to reduce the resolution to one-fourth and restore the original image size through convolution layers after passing through the decoder.

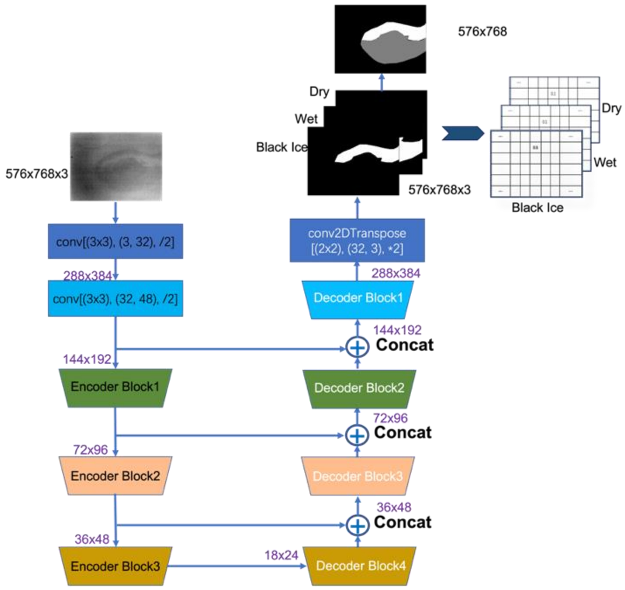

Figure 3: Network architecture diagram of the proposed CNN-based model

The encoder block expands the receptive field by a method that extends the feature map information, and the multi-scale dilated convolution extracts feature information from thermal road images by parallelly connecting dilated convolution layers with different dilation rates. In the decoder block, transpose convolutions are used instead of up-sampling to restore the image size and detailed feature information.

Each convolution layer is activated by a ReLU layer and undergoes batch normalization for normalization. In Figure 3, conv[(3x3), (3, 32), /2] and conv[(3x3), (3, 32), *2] represents a convolution operation, where the first value (3x3) indicates a convolution kernel size of 3x3. The second element (3, 32) and (32, 3) indicate the number of input and output channels, respectively. The third element, /2, represents down-sampling with a stride of 2, while *2 represents up-sampling by a factor of 2.

Figure 4 shows the encoder block with multi-scale dilated convolutional feature fusion (MsDC-FF) module with scalable dilation ratio. The proposed encoder block combines dilated convolution layers with different dilation rates in parallel, generating more scale features from larger receptive fields. Through a series of feature concatenations, neurons in each intermediate feature map encode semantic information from multiple scales, while different intermediate feature maps encode multi-scale information from different receptive fields. Through a series of dilated convolutions, neurons in the layers further back can obtain progressively larger receptive fields without degradation issues.

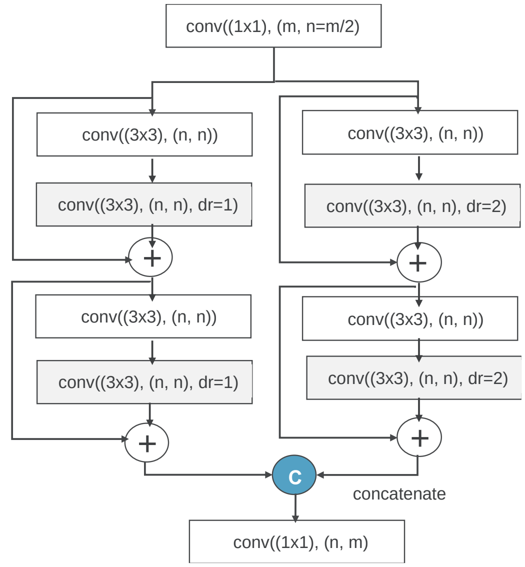

Figure 4: Encoder block diagram of the paralled execution of convoultion with multi-scale dilated convolutional feature fusion (MsDC-FF) module

Dilated convolution selects pixels to be used in the convolution operation based on the dilation rate, as shown by the colored pixels in Figure 5. It allows for a larger receptive field without increasing the number of parameters compared to standard convolutional operations. Dilated convolution can solve the trade-off problem between feature map resolution and receptive field size. The dilation rate (*dr*) in Figure 5 represents the dilation factor, and Equation (1) can be used to calculate the receptive field size (F_{dr}) in dilated convolution.

$$F_{dr} = (2dr + 1) \times (2dr + 1) \qquad (1)$$

where *dr* is the dilation ratio.

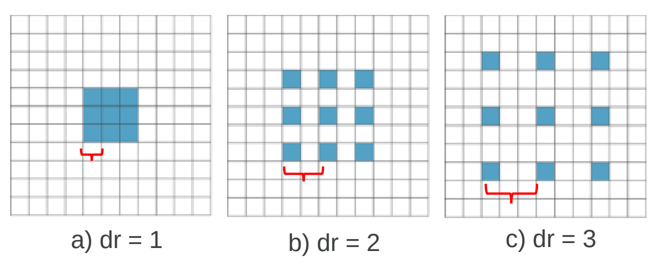

Figure 5: Dilation convolution with different dilation ratios

4. Simulation and Results

4.1. Image Dataset

The simulation setup used in this paper is as followings: TPV-IAHDR thermal cameras were used to capture the entire process of black ice formation from the beginning in a video with a resolution of 1280x720. Figure 6 illustrates an example of the gradual formation of black ice, starting from a wet road surface, in one specific case.

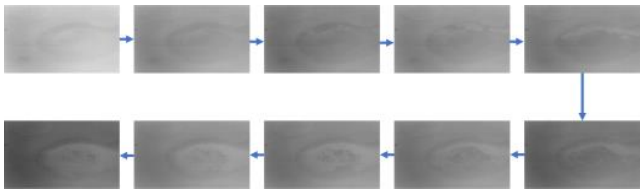

Figure 6: An example of the gradual formation of black ic e on a wet road surface.

These thermal camera images are used for training image dataset, by sampling and cropping frames at intervals of 200ms. This established total 1,156 black ice road images for 10 different cases and then, these images were divided into training, validation, and test datasets according to a ratio of 6:2:2. Therefore, the thermal road black ice dataset constructed in this paper is as shown in Table 1. The image dataset of the thermal road black ice was generated on asphalt roads and cement roads for different cases. These images are labeled by open-source image annotation tool, Labelme as shown in Figure 7. This displays some of the original images used in the paper along with their corresponding mask images.

Table 1: Number of infrared black ice road furface images

Dataset Type	No. of Images
Train Dataset	697
Validation Dataset	229
Test Dataset	229
Total Dataset	1156

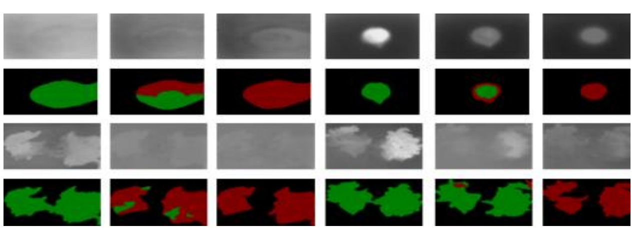

Figure 7: Example of original images and labeled or masked images of black ice generated on asphalt roads and cement roads for different cases

4.2. Simulation Setup and Results

The simulation setup was conducted on the following platform: to use Ubuntu 18.04 LTS as OS, to use GPU of four NVIDIA GeForce RTX 2080 Ti with 11GB of memory each. The deep learning frameworks used were Keras and TensorFlow. In the simulation, the number of epochs was set to 100, and batch sizes of 1, 2, 4, 8, and 16 were tested. During the training process, the cross-entropy loss function was utilized, and the Adam optimizer was employed with a learning rate of 0.001. Learning rate decay was applied to expedite the learning process.

Since random selection was used in parameter initialization and data selection during the network training, the network parameters and performance varied slightly with each training session. Therefore, in this paper, the training and testing processes were repeated ten times, and the average results were used to obtain more stable and reliable outcomes.

The performance comparison metric of the accuracy of black ice area detection used in this paper is mIoU (Mean Intersection over Union) metric. Meanwhile IoU is a measure of the overlap between the segmentation result and the ground truth, mIoU is the ratio of the intersection and union of quantized results and ground truth values as in Eq. (2).

$$m\mathrm{IoU} = \frac{1}{N} \sum_{i=1}^{N} \frac{X_{ii}}{T_i + \sum_{j=1}^{N}(X_{ji} - X_{ii})} \quad (2)$$

where N represents the number of pixel classes in the image, Ti denotes the total number of pixels for class i, X_{ii} represents the number of pixels where the true class is i and the predicted class is also i, and X_{ji} represents the number of pixels where the true class is i but the predicted class is j. mIoU is a simple and representative metric widely used to evaluate the segmentation results of networks in most image segmentation tasks. Moreover, the computational complexity of the proposed image semantic segmentation model was measured by comparing the size of the model parameters in kilobytes (KB).

The comparison evaluation result is shown in Table 2, where the proposed model has 2 parallel executions of convolutions in the encoder block. Encoder_block1, Encoder_block2, and Encoder_block3 have dilation rates with intervals of 4, 2, and 1, respectively. Table 2 shows the performance of mIoU and black ice IoU for conventional image segmentation networks, with batch sizes 8 and 16. The proposed CNN based MsDC-FF network model.

In Table 2, due to the larger sizes of U-Net, PSPNet, and DeepLabV3+, they could not be executed with a batch size of 16 due to memory limitations. Among the conventional image segmentation networks, LinkNet achieved the highest segmentation performance with an mIoU value of 95.48% for batch size 16, while it achieved the best Black Ice IoU value of 94.37% for batch size 8. However, for batch size 8, the proposed network model outperformed LinkNet with an mIoU of 95.93% and a Black Ice IoU of 94.87%.

Table 2: Peroformance comparison results of conventional networks models ans the proposed MsCD-FF network model

Network Model	Parameters (KB)	*m*IoU(%)		Black Ice IoU(%)	
		8	16	8	16
U-Net	31,055	69.69	-	61.38	-
FCN8	65,810	89.09	93.40	86.05	92.58
PSPnet	134,325	85.85	-	83.27	-
DeepLabV3+	41,253	93.65	-	88.20	-
ENet	371	94.35	94.36	93.81	93.10
LinkNet	11,555	95.39	95.48	94.37	94.33
Proposed MsCD-FF Net	492	95.93	95.61	94.87	94.48

5. Conclusion

In this paper, we presented a novel approach for black ice detection using a CNN with a multi-scale dilation convolution feature fusion (MsDC-FF) technique. The proposed technique achieved significantly improved accuracy in black ice detection, outperforming the baseline models by a substantial margin. The results indicate that the incorporation of multi-scale dilation convolution features enhances the CNN's ability to detect black ice accurately. Our findings suggest that this technique has great potential for improving road safety during winter seasons by enabling more effective identification of black ice hazards.

References

Breckon, T., & Fisher, R. B. (2012). A novel thermal-based approach to black ice detection. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR) (pp. 1492-1495). IEEE.

Chaurasia, A., & Culurciello, E. (2017, December). Linknet: Exploiting encoder representations for efficient semantic segmentation. In 2017 IEEE visual communications and image processing (VCIP) (pp. 1-4). IEEE.

Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.

Huang, G., Chen, D., Li, T., Wu, F., Van Der Maaten, L., & Weinberger, K. Q. (2017). Multi-scale dense networks for resource efficient image classification. arXiv preprint arXiv:1703.09844.

Kim, S.-J., Yoon, W.-S., & Kim, Y.-K. (2021). Characteristics of Black Ice Using Thermal Imaging Camera. Journal of the Korean Society of Industry Convergence, *24*(6_2), 873–882. https://doi.org/10.21289/KSIC.2021.24.6.873

Korea Traffic Accident Analysis System [Internet]. Available: http://taas.koroad.or.kr/.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, *60*(6), 84-90.

Lee, H., Hwang, K., Kang, M., & Song, J. (2020, December). Black ice detection using CNN for the Prevention of Accidents in Automated Vehicle. In 2020 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1189-1192). IEEE.

Li, Q., Ji, Y. W., Wang, Z. P., & Dou, X. (2017). Design of Road Icing Detection System Based on Opencv+ Python. Journal of Shaanxi University of Science & Technology (Natural Science Edition), *35*(2), 158-164.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).

Park, G. Y., Lee, S. H., Kim, E. J., & Yun, B. Y. (2017). A case study on meteorological analysis of freezing rain and black ice formation on the load at winter. Journal of Environmental Science International, *26*(7), 827-836.

Paszke, A., Chaurasia, A., Kim, S., & Culurciello, E. (2016). Enet: A deep neural network architecture for real-time semantic segmentation. arXiv preprint arXiv:1606.02147.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

Smith, S., Williams, B. L., & Prato, C. G. (2017). Black ice detection. In Encyclopedia of Traffic Science (pp. 1-7). Springer.

Wang, Q., Zhang, X., Chen, C., & Li, P. (2019). Black ice detection method based on the temperature field characteristic of thermal images. Journal of Advanced Transportation, 1-14.

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881-2890).