

딥러닝을 이용한 조선소에서 쓰러진 작업자의 검출에 관한 연구

박경민* · 김선덕** · 배철오***†

*, *** 목포해양대학교 해양경찰학부 교수, ** 목포해양대학교 해양경찰학부 박사과정

A Study on the Detection of Fallen Workers in Shipyard Using Deep Learning

Kyung-Min Park* · Seon-Deok Kim** · Cherl-O Bae***†

*, *** Professor, Division of Coast guard, Mokpo National Maritime University, Mokpo 58628, Korea

** PhD candidate Division of Coast guard, Mokpo National Maritime University, Mokpo 58628, Korea

요 약 : 선박은 크고, 복잡한 구조로 되어 있기 때문에 다른 작업자의 위치를 알아내기 어려우며, 특히 작업자가 쓰러진 경우에는 발견하기가 쉽지 않아 신속한 대처가 어렵다. 그리하여, 신체에 디바이스를 부착하는 방법이나 카메라를 이용하여 쓰러짐을 검출하기 위한 연구가 진행되고 있다. 기존의 영상기반 쓰러짐 검출은 사람의 신체부위를 검출하여 쓰러짐을 판단하였으나, 조선소에서는 다양한 복장과 자세로 작업으로 인해 검출하기가 어렵다. 본 논문에서는 쓰러짐 영역 전체를 추출하여 딥러닝 학습으로 선박 작업자의 쓰러짐을 이미지 기반으로 검출하였다. 학습에 필요한 데이터는 조선소의 건조중인 선박에서 쓰러진 모습을 연출하여 획득하였으며, 이미지를 좌우대칭, 크기조절, 회전하여 학습 데이터의 수를 증가하였다. 성능평가는 정밀도, 재현율, 정확도 그리고 오차율로 평가하였으며, 데이터의 수가 많을수록 정밀도가 향상되었다. 다양한 데이터를 보강하면 카메라를 이용한 쓰러짐 검출 모델의 실효성이 향상됨으로서 안전 분야에 기여할 수 있을 것으로 사료된다.

핵심용어 : 선박, 쓰러짐, 딥러닝, 이미지, 안전

Abstract : In large ships with complex structures, it is difficult to locate workers. In particular, it is not easy to detect when a worker falls down, making it difficult to respond quickly. Thus, research is being conducted to detect fallen workers using a camera or by attaching a device to the body. Existing image-based fall detection systems have been designed to detect a person's body parts; hence, it is difficult to detect them in various ships and postures. In this study, the entire fall area was extracted and deep learning was used to detect the fallen shipworker based on the image. The data necessary for learning were obtained by recording falling states at the shipyard. The amount of learning data was augmented by flipping, resizing, and rotating the image. Performance evaluation was conducted with precision, reproducibility, accuracy, and a low error rate. The larger the amount of data, the better the precision. In the future, reinforcing various data is expected to improve the effectiveness of camera-based fall detection models, and thus improve safety.

Key Words : Ship, Fallen, Deep learning, Image, Safety

1. 서 론

선박은 크기가 크고 복잡한 구조로 이루어져 있기 때문에 작업자 간 서로 통신이 되지 않을 시에는 위치를 알아내기

가 쉽지 않아 많은 시간이 소요된다. 기술의 발달로 인해 선박은 점점 대형화 되어가고 있으며 선원은 점점 줄어들고 있어 사람의 위치를 찾는 시스템이 필요하다. 해양의 사고에서 안전사고는 충돌, 전복, 화재, 폭발 등 항해관련 사고와 무관하게 사람이 사망, 실종, 부상한 것으로 우리나라 선박에서만 매해 100여건이 넘게 발생하고 있다(KMST, 2018; Joo et al., 2018). 선박의 건조나 정비 시 작업자에게 빈번하게 발

* First Author : allre@mmu.ac.kr, 061-240-7216

† Corresponding Author : baeco@mmu.ac.kr, 061-240-7223

생하는 쓰러짐은 신체에 손상을 입히게 되고, 기능장애로 인한 정신적인 피해까지 초래하게 된다. 때로는 작업자 혼자 이동하거나 작업하는 경우가 발생하게 되는데, 이 때 넘어지거나 쓰러지는 경우 다른 사람에게 알려거나 발견되기 어렵다. 쓰러짐 검출에 관한 연구는 주로 노약자나 장애인을 위해 이루어지고 있으며, 자이로스코프와 가속도 센서를 신체에 부착하여 신체의 움직이는 방향을 계산한 후 쓰러짐 상태 유무를 판단하는 것이다(Li et al., 2009). 이러한 디바이스를 이용한 방법은 계속 신체에 부착하고 있어야 하는 단점이 있으므로 이를 개선하기 위해 영상기반으로 쓰러짐을 검출하는 연구가 이루어지고 있다.

영상기반의 쓰러짐 검출의 대부분은 배경과 사람을 먼저 구분하고 사람의 자세, 움직이는 형태의 변화, 머리의 위치, 실루엣 등으로 판단한다(Solbach and Tsotsos, 2017). RGB 카메라와 적외선 카메라를 이용하여 쓰러지는 상황에서의 특징값을 추출한 후 분류기로 분류하는 방법은 지팡이나 물건을 소유하였을 경우나 쓰러지는 유형에 따라 가속도, 거리, 자세 등이 달라 판별하는 기준치가 달라 검출에 어려움이 있다(Bian et al., 2015). 또한, 키넥트(Kinect) 센서를 이용하여 머리의 위치, 몸의 중심, 관절 등을 검출 후 좌표값을 이용하여 쓰러짐을 판단하였으며, 이는 제한된 공간에서 간단한 시나리오로 검출률을 평가하였다(Mastorakis and Dimitrios, 2014; Mundher and Zhong, 2014). 산업현장에서의 근로자는 작업환경에 따라 특수한 작업복, 마스크, 장갑 등을 착용하고 다양한 도구를 이용하여 여러 자세로 작업을 하므로 신체의 머리, 어깨, 관절 등의 구분이 어려워 쓰러짐 검출에 어려움이 발생한다(Lu et al., 2018; Cao et al., 2017).

딥러닝(Deep Learning)의 발전은 다양한 분야에 영향을 미치고 있으며, 특히 객체 검출 분야에 우수한 성능을 나타내고 있다(Szegedy et al., 2017). 딥러닝은 검출하고자 하는 객체의 다양한 이미지를 신경망 학습을 이용해 특징을 추출하여 하나의 클래스로 생성할 수 있는 점을 이용하여, 조선소 작업자의 모습과 작업 중인 모습을 연출하여 데이터를 획득한 후 신경망 학습을 통해 클래스를 생성하여 쓰러짐을 검출하였다. 본 논문에서는 딥러닝을 이용하여 쓰러진 선박 작업자를 검출함으로써 해운산업의 안전 분야에 적용가능한지 확인하고자 한다.

2. 쓰러짐 검출방법

컴퓨터 비전 분야의 발달로 인하여 안전 분야에 적용하고자 하는 연구가 이루어지고 있으며, 그 중 영상에서 쓰러짐을 검출하는 연구가 진행되고 있다(Park and Bac, 2019). 쓰러짐은 바닥에 누운 상태로서, 이를 검출하기 위해 관절의

위치를 검출한 후 자세를 추정하는 방법이 대부분이다. 하지만, 이러한 방법은 넘어지는 자세에 따라 판별하는 기준치가 다르며, 특수한 복장을 갖추는 조선소의 작업자들의 관절을 찾는 것은 어려움이 있다. 본 연구에서는 딥러닝을 이용한 객체 검출분야의 기술을 활용하여 쓰러진 선박 작업자를 검출하고자 한다. 객체검출에 사용되고 있는 합성곱 신경망(CNN, Convolution Neural Networks)은 검출률이 계속 향상되고 있으며, 이미지의 객체를 검출하는 2015년 ILSVRC(ImageNet Large Scale Visual Recognition Challenge) 대회에서 사람보다 정확한 알고리즘이 소개되었다(He et al., 2015). CNN 기반의 객체 검출방법에는 Faster_RCNN(Regions with CNN), YOLO(You Only Look Once), SSD(Single Shot Detector) 등의 모델이 있으며 계속 새로운 모델이 개발되고 있다(Ren et al., 2015; Redmon et al., 2016; Liu et al., 2016).

본 실험에서는 복잡한 조선소 환경을 고려하여 겹치는 상황에서 정확도가 높은 Faster_RCNN과 여러 모듈을 병렬로 연결한 구조로서 다양한 응용 모델이 많은 인셉션(Inception)-V2를 적용하였다(Szegedy et al., 2016).

본 실험에서는 복잡한 조선소 환경을 고려하여 속도는 느리지만 겹치는 상황에서 정확도가 높은 Faster_RCNN과 GoogLeNet을 적용하였다(Szegedy et al., 2016). GoogLeNet은 Fig. 1의 모듈을 쌓아 구성한 신경망으로 크기가 다른 여러 모듈을 병렬로 연결한 구조이다. 또한, 1×1 합성곱을 사용하여 파라미터(Parameter) 수가 감소하므로, 연산량이 줄어들도록 하였다.

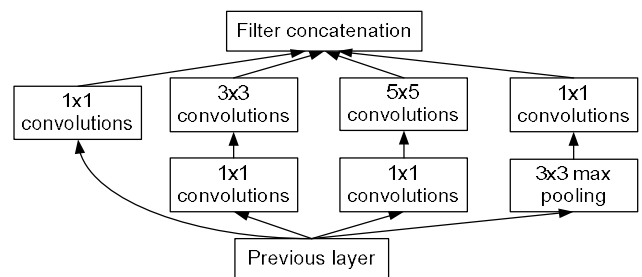


Fig. 1. Architecture of Inception module.

3. 실험

3.1 실험환경

모델 학습에서 필요한 데이터를 수집하기 위해, A 조선소의 야드와 건조중인 선박의 거주구역, 기관실, 갑판 등에서 쓰러지거나 일하고 있는 모습을 연출하여 이미지를 획득하였다. 이미지의 획득은 고정되지 않은 핸드폰 카메라를 이용하여 촬영을 하였으며, 별도의 조명은 설치하지 않았다. 학습환경은 Inter(R) Xeon CPU E5-1620 v4, 3.5 GHz, 16 GB

RAM, Windows10 Pro, NVIDIA GeForce RTX2080 8 GB, Tensorflow(GPU)1.13.1, python3.6, CUDA10.1, cudnn7.6.2의 환경에서 실험을 수행하였다.

3.2 데이터 학습

조선소에서 획득한 이미지는 총 965장으로, 쓰러진 이미지는 592장이고, 쓰러지지 않은 이미지는 373장이다. 랜덤 모듈을 이용하여 7대 3의 비율로 676장의 학습이미지와 289장의 테스트 이미지로 나누었다. 학습 데이터의 수를 증가시키기 위해, 좌우대칭, 크기 변경(0.8, 1.2배), 회전(-5°, 5°)을 적용하였다. Table 1은 5가지의 실험분류로서, 이미지 증가방법에 따른 이미지 수를 나타내고 있다.

학습에 사용된 모델은 COCO 데이터로 사전 학습된 모델인 Faster_rcnn_inception_v2이고, 파라미터는 23M, 클래스는 쓰러짐(Fallen)과 작업(Working) 2개, 학습률(Learning rate) 2e-4, 가속도(Momentum) 0.9로 설정하였다. 쓰러짐 영역과 작업 중인 영역을 외각 상자 추출 후 159,300번 학습하였으며, 학습 시간은 평균적으로 약 7시간이 소요되었다.

Table 1. Confusion matrix of the rotated image

	Augmentation	Number of image
Experiment_1	None	676
Experiment_2	Flip	1352
Experiment_3	Flip, Scale	4056
Experiment_4	Flip, Rotation	4056
Experiment_5	Flip, Scale, Rotation	12168

3.3 검출 성능 평가

학습한 데이터를 기반으로 쓰러진 이미지와 쓰러지지 않은 이미지로 나누어 검출성능을 평가하였다. Table 2는 검출 성능을 평가하기 위한 혼동행렬로, 실제 쓰러진 이미지에 대해 쓰러짐으로 판별한 경우는 True Positive, 실제 쓰러진 이미지에 대해 쓰러지지 않은 것으로 판별한 경우는 False Negative, 실제 쓰러지지 않은 이미지를 쓰러짐으로 판별한 경우는 False Positive, 실제 쓰러지지 않은 이미지를 쓰러지지 않은 것으로 판별한 것은 True Negative로 정의하였다.

Table 2. Confusion matrix of the fallen detection

	Actual fallen	Actual non-fallen
Predicted fallen	True Positive	False Positive
Predicted non-fallen	False Negative	True Negative

식(1) - (4)는 각각 정밀도(Precision), 재현율(Recall), 정확도(Accuracy) 그리고 오차율(Error rate)를 나타내며, 성능을 평가하기 위해 사용하였다. 쓰러짐 검출을 위한 연산시간은 800 × 640 화소에서 0.109s가 소요되었다.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (3)$$

$$Error\ rate = \frac{FN + FP}{TP + FN + FP + TN} \quad (4)$$

Table 3-7은 5가지의 학습데이터에 따른 실험 결과를 혼동행렬로 각각 나타내고 있으며, Table 8은 검출 성능을 나타내고 있다. 실험_1과 실험_2는 쓰러지지 않은 이미지를 쓰러진 것으로 판별한 경우가 많아 오차율이 크다. 데이터 수를 증가시킨 실험_3과 실험_4에서 검출성능이 향상되었으며, 좌우반전, 크기변경, 회전을 적용한 실험_5에서는 쓰러지지 않은 이미지를 쓰러진 것으로 판별한 경우가 없어 정밀도가 향상되었으나, 쓰러진 이미지를 검출하지 못한 경우가 있어 재현율이 낮았다.

Table 3. Confusion matrix of the Experiment_1

	Actual fallen	Actual non-fallen
Predicted fallen	172	9
Predicted non-fallen	3	105

Table 4. Confusion matrix of the Experiment_2

	Actual fallen	Actual non-fallen
Predicted fallen	174	11
Predicted non-fallen	1	103

Table 5. Confusion matrix of the Experiment_3

	Actual fallen	Actual non-fallen
Predicted fallen	177	5
Predicted non-fallen	2	105

Table 6. Confusion matrix of the Experiment_4

	Actual fallen	Actual non-fallen
Predicted fallen	175	3
Predicted non-fallen	3	108

Table 7. Confusion matrix of the Experiment_5

	Actual fallen	Actual non-fallen
Predicted fallen	174	0
Predicted non-fallen	5	110

Table 8. Detection performance of the experiment

	Precision	Recall	Accuracy	Error rate
Experiment_1	0.950	0.983	0.958	0.042
Experiment_2	0.941	0.994	0.958	0.042
Experiment_3	0.973	0.989	0.976	0.024
Experiment_4	0.983	0.983	0.979	0.021
Experiment_5	1.000	0.972	0.983	0.017

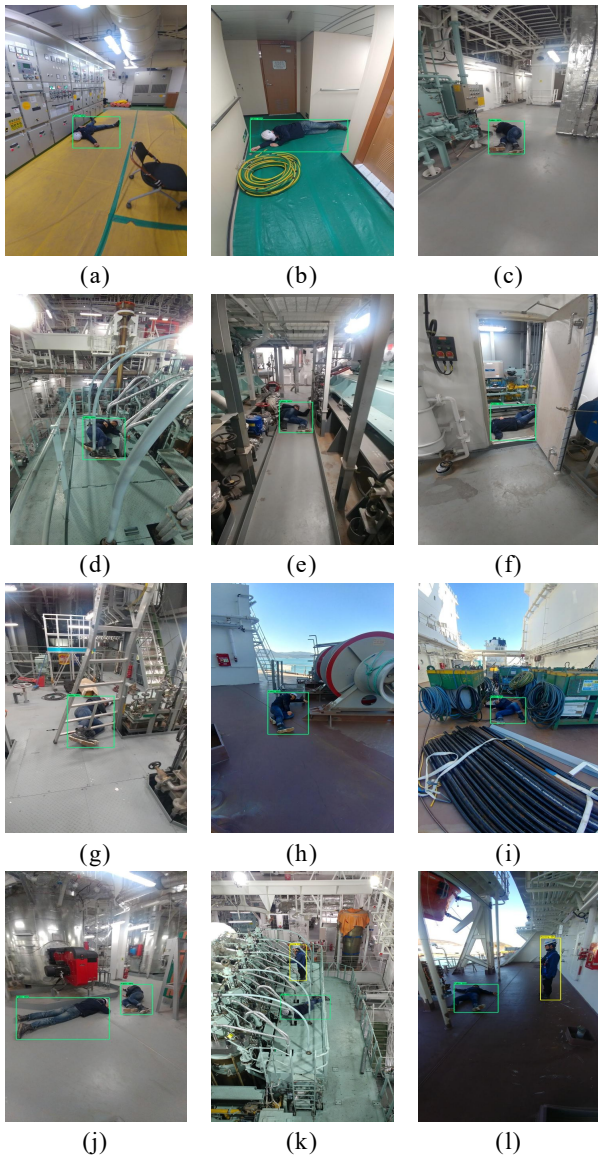


Fig. 2. Example of True Positive.

Fig. 2는 실제 쓰러진 이미지를 쓰러진 이미지로 판별한 경우로 녹색박스는 쓰러짐 영역, 노란색 박스는 쓰러지지 않은 영역을 나타내고 있다. 선박의 배전반실, 거주구역, 기관실, 메인엔진, 갑판 영역에서 검출된 쓰러진 이미지를 나타내고 있으며, (d)는 복잡한 파이프가 있는 환경에서도 검출되었으며, (f)는 쓰러진 작업자의 전체 모습이 아닌 일부분의 영역도 검출하였다. (g)는 사다리 밑에 쓰러진 작업자를 검출한 것이고, (j)는 두 명의 쓰러짐 검출, (k)와 (l)은 쓰러진 이미지와 쓰러지지 않은 이미지 둘 다 검출되었음을 나타낸다. Fig. 3은 쓰러진 영역이 아닌 이미지를 쓰러진 것으로 판별한 경우를 나타내고 있다. (a)는 배관과 전선 등이 쌓여있는 영역으로 실험_3과 실험_4에서 쓰러짐으로 판별하였다. (b)는 실험_3에서 계단에 오르고 있는 작업자를 쓰러짐으로 판별한 것이고, (c)는 실험_4에서 의자를 쓰러짐으로 판별한 것이다. Fig. 4는 실험_5에서 쓰러진 작업자를 검출하지 못한 것으로 나타낸 것으로, 신체의 일부분만 보이거나 쓰러짐 방향에 따라 검출하지 못하였다.

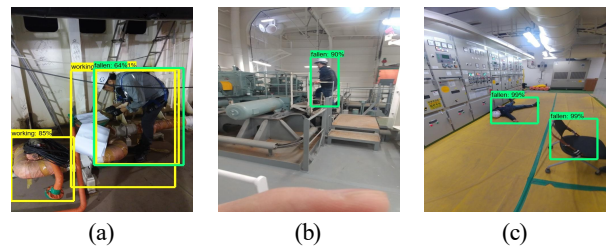


Fig. 3. Example of False Positive.

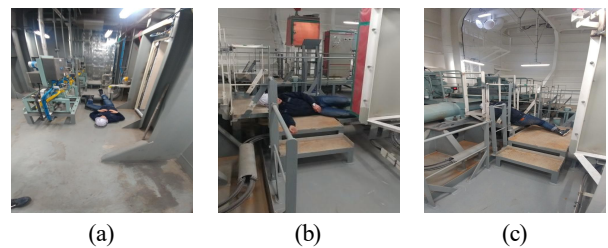


Fig. 4. Example of False Negative.

4. 결 론

본 논문에서는 합성곱 신경망인 Faster_rcnn_inception_v2을 적용하여 조선소에서 발생하는 작업자의 쓰러짐 검출을 이미지 기반으로 검출하였다. 쓰러진 이미지와 쓰러지지 않은 이미지 데이터를 획득한 후 반전, 크기변경, 회전을 통해 5가지의 학습데이터로 나누었다. 테스트 이미지로 쓰러진 이미지와 쓰러지지 않은 이미지를 검출하여 혼동행렬로 분류한 후 정밀도, 재현율, 정확도 그리고 오차율로 성능을 평가하였다. 학습 데이터의 수가 증가함에 따라 정밀도는 향상되었으나, 재현율, 정확도, 그리고 오차율은 비슷하였다. 이는 본 실험에서 사용한 학습 데이터의 수가 부족한 것으로 판단된다.

기존의 영상기반 쓰러짐 검출은 사람의 신체부위나 관절 등을 검출하여 쓰러짐을 검출하였으나, 조선소에서는 두꺼운 작업복과 안전장비, 도구 등을 소지하며 작업하고, 복잡한 배관과 장비가 섞여있어 특정 신체부위를 검출하기가 어렵다. 그리하여, 본 논문에서는 작업자의 쓰러진 영역 전체를 추출하여 학습하여 쓰러짐 검출을 한 결과, 검출 가능성을 확인하였다. 획득한 영상으로만 실험을 진행하였기 때문에 산업에 적용되는 것은 한계가 있으나, 다양한 데이터를 획득하고 다른 모델에도 적용함으로써 연구의 실효성이 향상될 수 있도록 하고자 한다.

References

- [1] Bian, Z. P., J. Hou, L. P. Chau, and M. T. Nadia(2015), IEEE Journal of Biomedical and Health Informatics, 19, pp. 430-439.
- [2] Cao, Z., T. Simon, S. E. Wei, and Y. Sheikh(2017), Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, arXiv:1611.08050.
- [3] He, K., X. Zhang, S. Ren, and J. Sun(2015), Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778.
- [4] Joo, M., M. Choo, Y. Baek, N. Kim, A. Choi, S. Im, J. Lee, H. Kim, and H. Lee(2018), Korea Coast Guard's Human Biological Materials Storage Project for Identifying Bodies Recovered from the Sea: A Model Suggestion, Journal of the Korean Society of Marine Environment & Safety, Vol. 24, No. 2, pp. 171-178.
- [5] KMST(2018), <https://www.kmst.go.kr/>.
- [6] Li, Q., J. A. Stankovic, M. Hanson, A. Barth, and J. Lach (2009), Accurate, Fast Fall Detection Using Gyroscope and Accelerometer-Derived Posture Information, In Proceedings of the 6th International Workshop on Wearable and Implantable Body Sensor Networks, pp. 138-143.
- [7] Lu, N., Y. Wu, L. Feng, and J. Song(2018), Deep Learning for Fall Detection: 3D-CNN Combined with LSTM on Video Kinematic Data, Vol. 23, No. 1, pp. 314-323.
- [8] Liu, W., D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg(2016), SSD: Single Shot MultiBox Detector, Proceedings of the European Conference on Computer Vision, pp. 21-37.
- [9] Mastorakis, G. and M. Dimitrios(2014), Fall detection system using Kinect's infrared sensors, Journal of Real-Time Image Processing, Vol. 9, No. 4, pp. 635-646.
- [10] Mundher, Z. A. and J. Zhong(2014), A Real-Time Fall Detection System in Elderly Care Using Mobile Robot and Kinect Sensor, International Journal of Materials, Mechanics and Manufacturing, Vol. 2, No. 2, pp. 133-138.
- [11] Park, K. M. and C. O. Bae(2019), A Study on Fire Detection in Ship Engine Rooms Using Convolutional Neural Network, Journal of the Korean Society of Marine Environment & Safety, Vol. 25, No. 4, pp. 476-481.
- [12] Redmon, J., S. Divvala, R. Girshick, and A. Farhadi(2016), You Only Look Once: Unified, Real-Time Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788.
- [13] Ren, S., K. He, R. Girshick, and J. Sun(2015), Faster R-CNN: Towards real-time object detection with region proposal networks, In Neural Information Processing Systems.
- [14] Solbach, M. D. and J. K. Tsotsos(2017), Vision-based Fallen Person Detection for the Elderly, The IEEE International Conference on Computer Vision (ICCV) workshops, pp. 1433-1422.
- [15] Szegedy, C., S. Ioffe, V. Vanhouche, and A. A. Alemi(2017), Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, Proceeding of the Thirty-First AAAI Conference on artificial Intelligence, pp. 4278-4284.
- [16] Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna (2016), Rethinking the Inception Architecture for Computer Vision, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818-2826.

Received : 2020. 03. 18.

Revised : 2020. 05. 14. (1st)

: 2020. 06. 19. (2nd)

Accepted : 2020. 10. 28.