



딥러닝 기반 LNGC 화물창 스캐닝 점군 데이터의 비계 시스템 객체 탐지 및 후처리

이동건¹·지승환^{2,†}·박본영²

목포해양대학교 조선해양공학과¹

목포해양대학교 대학원 해양시스템공학과²

Object Detection and Post-processing of LNGC CCS Scaffolding System using 3D Point Cloud Based on Deep Learning

Dong-Kun Lee¹·Seung-Hwan Ji^{2,†}·Bon-Yeong Park²

Department of Naval Architecture and Ocean Engineering, Mokpo National Maritime University¹

Department of Ocean System Engineering, Graduate School, Mokpo National Maritime University²

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, quality control of the Liquefied Natural Gas Carrier (LNGC) cargo hold and block-erection interference areas using 3D scanners have been performed, focusing on large shipyards and the international association of classification societies. In this study, as a part of the research on LNGC cargo hold quality management advancement, a study on deep-learning-based scaffolding system 3D point cloud object detection and post-processing were conducted using a LNGC cargo hold 3D point cloud. The scaffolding system point cloud object detection is based on the PointNet deep learning architecture that detects objects using point clouds, achieving 70% prediction accuracy. In addition, the possibility of improving the accuracy of object detection through parameter adjustment is confirmed, and the standard of Intersection over Union (IoU), an index for determining whether the object is the same, is achieved. To avoid the manual post-processing work, the object detection architecture allows automatic task performance and can achieve stable prediction accuracy through supplementation and improvement of learning data. In the future, an improved study will be conducted on not only the flat surface of the LNGC cargo hold but also complex systems such as curved surfaces, and the results are expected to be applicable in process progress automation rate monitoring and ship quality control.

Keywords : PointNet(포인트넷), Deep learning(딥러닝), Point cloud(스캐닝 점군 데이터), Object classification(객체 분류), Object segmentation(객체 분할)

1. 서론

변화하는 글로벌 시장 환경에서 국내 조선소의 생산경쟁력을 확보하기 위한 생산관리 고도화가 요구되고 있으며, 스마트 생산 고도화를 위한 연구로 시뮬레이션 기반 생산관리 (Woo et al., 2016), 생산계획 정확도 향상 (Lee et al., 2014) 등 생산성 향상을 위한 연구가 진행되고 있다. 최근 조선 산업에서 4차 산업기술 활용에 대한 관심이 높아짐에 따라 빅데이터 및 기계 학습 적용을 통한 조선 생산 정보 분석 (Ham et al., 2016), 기

계학습을 적용한 블록 탑재 시수 예측 (Kwon & Ruy, 2020) 등 다양한 사례가 개발되고 있다. 이에, 본 연구에서는 인공지능 기술의 일종인 딥러닝(deep learning)을 선박 생산의 정도 관리 공정 자동화에 적용하여 조선 산업의 경쟁력을 향상시키고자 하였다.

제조업 분야에서는 다양한 정도 관리 연구가 수행되고 있으며, 양질의 제품 생산을 위해 정도 계획은 매우 중요하다. 국내 대형 조선소 삼성중공업에서는 블록 탑재 간섭부위 사전 조치 용으로, 국제 선급 DNV(이전 DNV-GL)에서는 작업이 완료된 멤

브레인 LNGC(Liquefied Natural Gas Carrier) 화물창 정도 관리를 위해 3D 스캐너를 활용한 바 있다. 이처럼 정도 관리 분야에서는 고정밀 데이터 확보에 용이한 3D 라이더(LiDAR, Light Detection and Ranging) 스캐너를 활용해 계측을 수행한다.

본 연구에서는 3D 스캐너로 계측된 LNGC 화물창 데이터를 활용하여 비계 시스템(scaffolding system) 객체 탐지(object detection) 및 제거를 위한 연구를 수행하였다. 객체 탐지는 컴퓨터 비전과 이미지 처리와 관련된 컴퓨터 기술로서, 초기에 연구된 객체 탐지 기법들은 주로 디지털 이미지 또는 영상(2D)으로부터 특정한 계열의 객체 인스턴스(인간, 건물 등)를 감지하는 내용이 주를 이룬다. 본 연구에서는 기존의 2D 데이터(이미지, 영상)가 아닌 3D 데이터(스캐닝 점군 데이터) 기반 3D 객체 탐지 기법을 적용하였으며 기법의 효율성과 지속적인 활용가능성을 검토하였다. 2D 데이터가 아닌 3D 데이터를 활용하여 딥러닝 기법을 적용하는 연구는 '17년도부터 점차 활성화되고 있으나 데이터의 구조적인 복잡성으로 인해 산업에 적용 가능한 고수준의 정확도 도달을 위한 연구는 현재까지도 진행 중에 있다.

LNGC 화물창 공정 시 액화천연가스를 보온하기 위해 단열 패널(insulation panel)을 내부 선각(inner hull)에 접착하는 방식을 따르는데, 접착된 단열재의 품질은 내부 선각, 마스틱(mastic), 단열 패널의 편평도와 밀접한 관계를 가진다. Jeong et al. (2021)에 따르면 LNGC 화물창은 선박의 6자유도 운동에 따른 액화천연가스의 지속적인 슬로싱 하중으로 인한 인장응력을 받는데, 이는 단열 시스템의 손상을 야기하는 주 원인으로 작용하기 때문에 사전에 수행된 화물창 정도 관리의 중요성이 부각된다.

LNGC 화물창 내부 선각의 편평도 계측 검사를 위해 3D 스캐너를 활용하여 데이터를 측정할 시 Fig. 1과 같은 스캐닝 점군 데이터(3D point cloud)가 수집된다. 그러나 비계 시스템이 장애물 역할을 하여 Fig. 2와 같이 다수의 음영구역이 발생하며, 이러한 음영구역의 데이터를 보충하기 위해 화물창을 구역별로 복수 계측 후 정합(registration)하는 작업이 요구된다.

LNGC 화물창과 같은 거대 구조물의 경우 계측 시 수집되는 스캐닝 점군 데이터가 억 단위를 넘어가는 빅데이터(big data)가 생성되며 이에 복수 계측된 스캐닝 점군 데이터 정합 작업은 데이터 양에 따라 수십, 수백 시간이 소요된다. 정합은 2회 이상 계측된 스캐닝 점군 데이터를 통합하는 작업으로 이를 통해 음영구역이나 부족한 부분의 데이터 보충이 가능하다(Fig. 3). LNGC 화물창과 같은 빅데이터를 정합하고자 할 때에는 작업에서 발생하는 계산비용 절감을 위해 사전에 비계 시스템 객체를 탐지하여 지우고 정합하여 화물창의 품질 검사를 수행하는 것이 바람직하지만, 수작업으로 비계 시스템 객체 데이터를 제거할 시 품질 검사 대상인 화물창 내부 선각 데이터를 보존하며 작업해야 하기 때문에 작업자의 숙련도가 요구되며, 수억 개의 빅데이터를 실시간으로 검토하면서 처리하는 작업은 고사양의 하드웨어 시스템이 요구되기 때문에 작업자를 육성하고 고가의 시스템을 구축하는 것은 실무적인 입장에서 채택하기 어려운 선택이다.

따라서, 본 연구에서는 LNGC 화물창 품질관리 고도화 연구의 일환으로 LNGC 화물창 스캐닝 점군 데이터를 정합하는 작업의 계산비용을 줄이기 위해, 정합 작업의 선행 단계인 비계 시스템 객체 제거에 딥러닝 기법을 적용함으로써 LNGC 화물창 전체의 스캐닝 점군 데이터로부터 비계 시스템 객체를 자동으로 탐지하고 비계 시스템 객체로 예측된 데이터만 제거하였다.

대량의 스캐닝 점군 데이터를 정합하는 작업과 같은 빅데이터를 처리하는 작업은 원활한 수행을 위해 고사양의 하드웨어 시스템이 요구되기 때문에 기존 시스템이 구축된 산업에 진입하는 것이 어려운 상황이다. 따라서 본 연구와 같은 인공지능 기술을 빅데이터 처리에 적용하여 자동화가 부여된 데이터 처리 기술이 산업 진입에 핵심 요소로 작용할 것으로 사료된다.

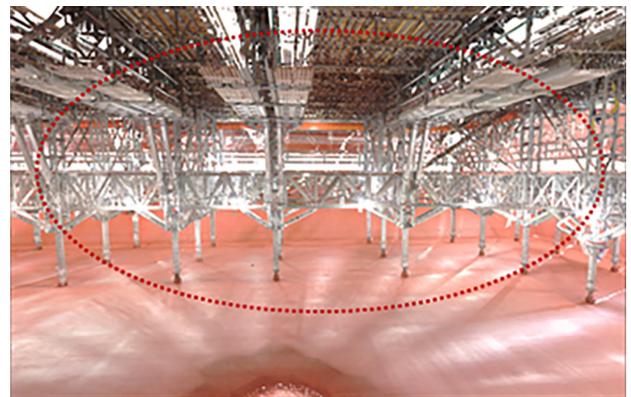


Fig. 1 First floor's scaffolding system in LNGC cargo containment system

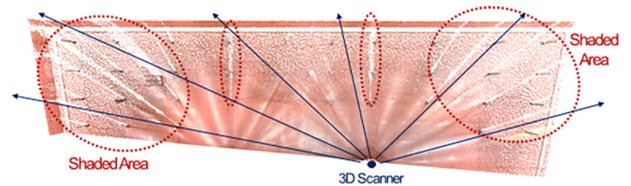


Fig. 2 Point cloud in first scan using 3D scanner

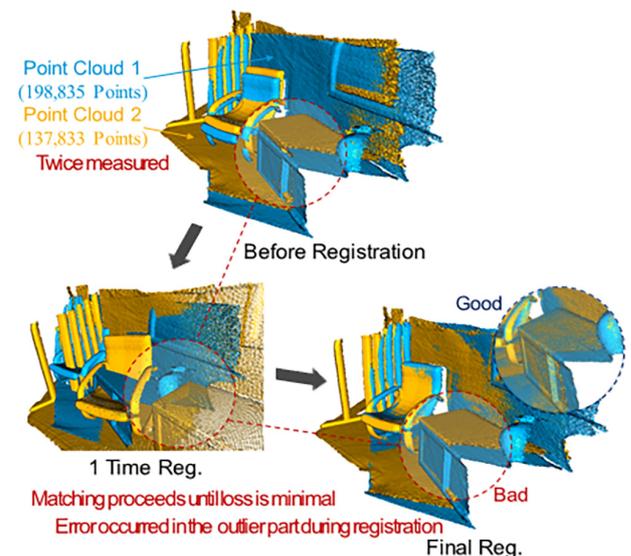


Fig. 3 Point cloud registration

2. 딥러닝 기반 객체 탐지

2.1 2D 분류 및 분할

기존 컴퓨터 공학 분야에서는 이미지 처리 기술을 통한 객체 탐지에 관한 지속적인 연구가 이루어지고 있다. 2D 데이터를 통한 딥러닝 기반 객체 탐지 기법은 크게 분류(classification)와 분할(segmentation)로 나눌 수 있으며, Fig. 4와 같이 스틸 드럼(steel drum)의 이미지에 대해 스틸 드럼이라는 하나의 라벨을 예측하는 행위를 분류라 칭한다. 이러한 예측은 심층 신경망(deep neural network)을 통해 학습되는데, 이는 입력층(input layer)과 출력층(output layer) 사이에 여러 개의 은닉층(hidden layer)들로 이루어진 인공 신경망(artificial neural network)이다(Kim et al., 2019).

지역화(localization)는 2D 데이터에 대해 찾고자 하는 객체의 색상과 같은 특징들을 검토하여 위치를 파악하는 행위이다. 일반적으로 객체의 위치를 찾아 경계 상자(bounding box)를 예측하여 생성(Fig. 4)하며 예측정확도는 IoU, Accuracy 등의 관련 지표들을 통해 검토한다.

분할은 2D 데이터의 모든 픽셀들에 대해 사전에 부여된 라벨 정보를 학습하고 이를 바탕으로 예측하여 데이터 내의 객체별로 경계 상자를 생성하기 위한 사전 작업이라고 할 수 있다.

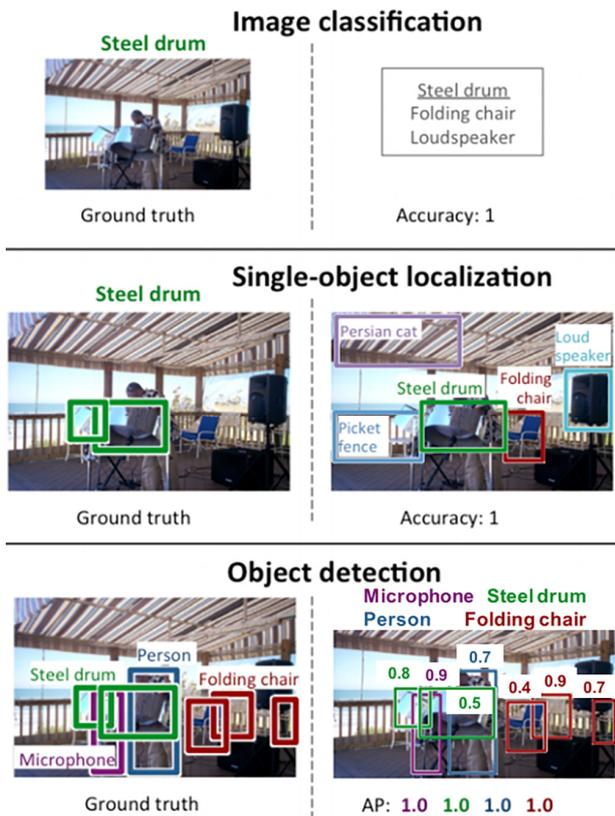


Fig. 4 Classification to object detection in 2D data (Russakovsky et al., 2015)

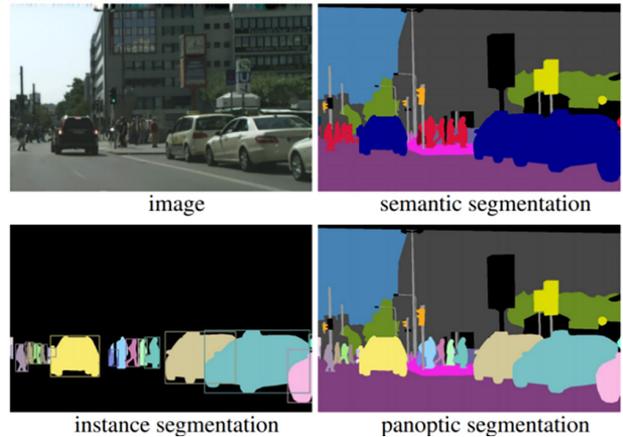


Fig. 5 An example of different types of image segmentation (Kirillov et al., 2019)

2D 분할을 위한 딥러닝 모델은 2D 데이터에 대해 찾고자 하는 객체별로 그 객체들의 라벨을 예측하며, 이는 모든 픽셀의 라벨을 예측하는 것과 같다(Fig. 5).

분류와 분할 모두 객체를 찾는 방법임에는 동일하지만 2D 데이터로부터 복수 객체를 찾고자 한다면 분할 작업은 필수적이라고 할 수 있다. 이러한 방법들은 초기 개발 당시, 보편적으로 활용되고 있는 2D 입력 데이터(이미지, 영상) 처리를 위해 연구되었으며, 최근에는 스캐닝 점군 데이터 형태인 3D 데이터에 적용하기 위해 확장된 객체 탐지 기법들이 다양한 방법으로 제시되고 있다. 이에, 딥러닝 모델에 대한 예측 정확도를 높이기 위해 이를 RGB-D 데이터와 혼합하거나 3D 복셀로 양자화하는 등 고도화된 연구가 지속적으로 수행되고 있으며 3.1절에서 이와 관련된 3D 객체 탐지 기법들을 검토하였다.

2.2 3D 데이터 활용을 위한 객체 탐지 기법

객체 탐지는 컴퓨터에게 시각적인 요소를 전달할 수 있는 매체로부터 데이터를 수집 및 처리하여 수행되는데 과거에는 주로 카메라와 같은 처리 장치로부터 데이터를 수집하였다. 이는 2D 형식의 데이터로, 동일한 크기의 픽셀들이 규칙적으로 배열된 형태이기 때문에 데이터를 딥러닝 기반 객체 탐지 기법이 수행되는 CNN(Convolution Neural Network)에 입력하여 병렬처리를 수행하는 것이 가능하였다. 그러나 스캐닝 점군 데이터는 3D 라이다 스캐너로부터 수집되는 3D 형식의 데이터로써 계속 이루어지는 스캐닝 공간에 비해 데이터의 분포된 정도가 희박(Sparse)하고 순차적인 수집이 안 되는(unordered) 특징을 가진다. 이에, 측정하고자 하는 객체가 스캐너로부터 멀어질수록 수집되는 데이터의 양이 현저히 감소하고 순차적인 데이터의 수집이 되지 않아 3D 객체 탐지 기법의 연구 초기에는 스캐닝 점군 데이터를 입력 데이터로 활용하여 기존의 2D 객체 탐지 아키텍처에 적용하는 것이 불가능하였다.

일반적으로 컴퓨팅 시스템을 기반하는 기술의 발전은 선행연

구 사례로부터 점진적으로 이루어지며 객체 탐지 기법 또한 같은 원리로 발전해왔기 때문에 3D 객체 탐지 기법은 기존의 2D 객체 탐지를 위한 딥러닝 아키텍처를 최대한 활용하고자, 동일한 구성의 기존 방법에 적용하는 것부터 시도되었다. 그러나 3D 스캐너로 수집된 스캐닝 점군 데이터의 위와 같은 특징들 때문에 2D 객체 탐지의 기존 기법들을 그대로 적용하는 것이 불가능하였고 이러한 문제점을 해결하고자 스캐닝 점군 데이터를 3D 복셀 그리드(3D voxel grid) 형태로 가공하여 데이터를 처리하는 기법인 VoxNet (Maturana and Scherer, 2015)이 등장하였으며 2017년, 스캐닝 점군 데이터 자체만을 활용하여 객체 탐지 아키텍처에 적용한 PointNet (Charles et al., 2017) 기법이 등장하였다. PointNet 기법은 큰 이슈를 몰고 현재 연구되고 있는 3D 입력 데이터를 활용한 딥러닝 기반 객체 탐지 기법들의 근간으로 자리매김하고 있다.

3. 딥러닝 기반 3D 객체 탐지 기법

3.1 스캐닝 점군 데이터와 색상, 깊이 데이터를 함께 활용한 Frustum PointNets

Fig. 6의 Frustum PointNets은 PointNet 기법을 개발한 주저자인 Charles, R.Q.에 의해 2018년 발표된 딥러닝 기반 3D 객체 탐지 기법이다. 주어진 RGB-D 데이터에 대해 먼저, CNN을 사용하여 RGB(색상) 데이터로부터 2D 객체 영역 제안(Proposal)을 생성한다. 각 2D 영역은 D(깊이) 데이터에서 스캐닝 점군 데이터를 얻는 3D 뷰 절두체(3D view frustum)로 돌출된다. 마지막으로 절두체의 점에서 객체에 대한 3D 경계 상자를 예측한다. 이전 방법들은 이미지 또는 3D 복셀에 초점을 맞추고 있으며 Frustum PointNets은 종종 자연적인 3D 패턴과 3D 데이터의 불변성을 가리지만 RGB-D 스캔을 팝업(Popup)하여 스캐닝 점군 데이터 자체 데이터에서 직접 작업한다는 것이 특징이다. 특히, 2D 객체 탐지기와 객체 위치 파악을 위해 고도화된 3D 딥러닝을 모두 활용하여 작은 객체에 대해서도 높은 효율성과 재현율(Recall)을 달성하며 스캐닝 점군 데이터 자체 데이터를 직접 학습함으로써 매우 희소한 포인트에서도 3D 경계 상자를 정확하게 추정할 수 있다는 장점을 가진다.

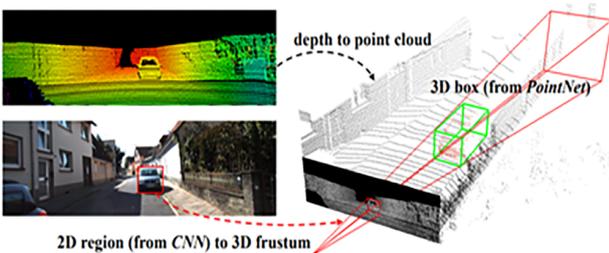


Fig. 6 3D object detection using rgb-d data of Frustum PointNets (Charles et al., 2018)

3.2 End-to-End 학습을 위한 VoxelNet

3D 객체 탐지 기법들은 연구 당시 3D 스캐너를 통해 수집한 데이터를 3D 객체 탐지 아키텍처에 적용하기 위해 2D 객체 탐지 아키텍처와 비슷한 형태를 취했으며 이에, 핸드크래프트 피처(Hand-craft feature)를 어떻게 만들 것인가에 집중했다. 핸드크래프트 피처란, 가장자리, 꼭짓점 등 객체를 파악할 수 있는 직관적인 특징들을 예로 들 수 있는데, 이러한 특징들을 통해 객체의 위치를 예측하여 경계 상자를 생성한다. 그러나 Fig. 7의 VoxelNet(Yin and Tuzel, 2017)은 핸드크래프트 피처를 제거하고 특징 추출과 경계 상자 예측을 통합하는 end-to-end 학습이 가능한 네트워크를 제안했다. 이에, 스캐닝 점군 데이터를 같은 간격으로 구성된 3D 복셀(2D의 픽셀을 3D화 한 것)로 쪼개고 VFE 레이어(Voxel Feature Encoding layer)를 통해 각 복셀 내의 포인트들을 이용하여 복셀의 특징(feature)을 만든다. 그리고 특징을 보유한 복셀들을 3D 컨볼루션(convolution) 연산을 통해 로컬 복셀 피처(local voxel feature)들을 통합하고 이를 RPN(Region Proposal Network)에 넣어 경계 상자를 예측하여 생성한다. 일반적인 딥러닝 기반 객체 탐지 기법은 수많은 로컬 데이터에 라벨 정보를 매핑하고 이를 통합하는 글로벌 레지스트레이션(global registration) 과정을 걸쳐 RPN 알고리즘을 통해 가장 동일한 클래스(class)라고 판단되는 경계 상자를 생성시킨다.

이처럼 컴퓨터 공학 분야에서 연구 중인 고도화된 딥러닝 기반 3D 객체 탐지 기법들은 고수준의 예측 정확도 확보를 위해 스캐닝 점군 데이터와 RGB-D 데이터를 함께 활용하거나 신속한 병렬 처리를 위해 3D 복셀로 공간을 쪼개어 복셀 내의 포인트들을 분석 및 예측하는 기법들이 개발되었다. 이러한 3D 객체 탐지 기법들의 연구에 활용된 데이터들은 실시간 차량 운행을 통해 수집된 데이터를 가공하여 제공하는 공개된 데이터 셋을 입력 데이터로 활용하였다. 본 연구에서는 비계 시스템 객체를 탐지하기 위한 데이터 셋이 없기 때문에 고수준의 객체 탐지 정확도를 위한 고도화된 연구 수행을 우선 보류하였다.

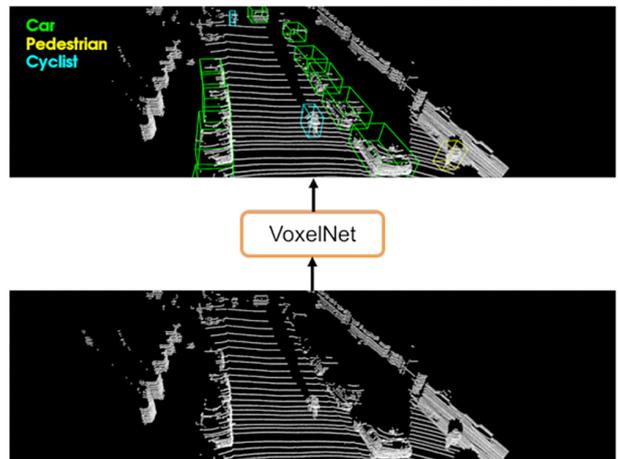


Fig. 7 Split point cloud into equally spaced 3D voxels of VoxelNet (Yin and Tuzel, 2017)

3.3 순서 불변 & 변환 불변

PointNet 기법은 스캐닝 점군 데이터만을 활용한 대표적인 객체 탐지 기법으로 기존 연구에서 활용되었던 2D 데이터(이미지 또는 영상)를 활용하지 않는다는 점에서 차이가 있다. 2D 데이터는 픽셀 단위로 구성되어 있고 픽셀은 가로, 세로 데이터 영역의 크기가 동일하기 때문에 2D 데이터에 대한 컨볼루션 연산이 쉽게 이루어져 CNN에 적용시키는데 문제가 없었지만 입력 데이터가 스캐닝 점군 데이터로 확장되면서 문제가 제기되었다. 3D 스캐너를 통해 데이터를 수집할 때에는 스캐너로부터 객체 데이터까지의 거리가 멀어질수록 수집되는 데이터의 양이 현저하게 줄어들고 데이터를 수집하는 순서가 순차적으로 이루어지지 않다는 특징이 있다. 이러한 특징들로 인해 어떤 위치의 점(Point)이 어떤 객체를 나타내는지, 라벨 정보를 몇 번째 점 데이터에 부여하고 이들을 입력 데이터로 활용하여 컨볼루션 연산을 수행해야 하는지에 대한 연구는 곧바로 이루어지지 않았다. 이후, 스캐닝 점군 데이터를 3D 복셀 그리드로 변환시킨 뒤 복셀 그리드를 컨볼루션 연산하는 기법들이 개발되었으나 이는 입력 데이터를 가공한다는 점에서 스캐닝 점군 데이터 자체 데이터를 활용한다는 것과 거리감이 있었다. 하지만 스캐닝 점군 데이터 자체 데이터를 활용한 딥러닝 기반 3D 객체 탐지 기법인 PointNet이 등장하면서 스캐닝 점군 데이터 자체 데이터만을 활용한 컨볼루션 연산이 가능함을 입증하였는데 PointNet은 크게 순서 불변(permutation invariance), 변환 불변(transformation invariance)을 따른다.

이 두 가지 법칙은 스캐닝 점군 데이터를 그대로 활용하기 위한 전제조건으로 우선, 3D 스캐너를 통해 수집된 데이터는 Fig. 8과 같은 형태로 수집된다. 좌측 인덱스(index)는 점 각각의 순번을 나타내고 우측에는 그와 대응되는 점 좌표를 나타낸다. 순서 불변은 심층 신경망(deep neural network) 학습 도중 데이터가 뒤섞일 수 있으나 인덱스에 따른 행 데이터는 고정되어야 한다는 법칙이고 PointNet은 이를 따른다.

이미지의 문자를 식별하는 딥러닝 기법은 어떠한 방향에서 쓰여진 문자도 올바르게 예측하였다. 그것이 가능한 이유는 공간 변형 네트워크(STN, Spatial Transformer Networks) 때문인데 이는, 입력 데이터의 좌표축을 변환시키는 변환 행렬을 곱해 어떤 방향에서 쓰여진 문자들도 한 방향에서 바라보는 효과를 발생시키는 네트워크이다 (Jaderberg et al., 2015). PointNet은 공간 변형 네트워크를 변형시켜 3D 입력 데이터인 스캐닝 점군 데이터에 적용하였으며 이를 T-Net이라 칭하였다 (Charles et al., 2017). Fig. 9와 같이 입력 데이터의 차원을 확장시킨 뒤 최대 풀링(max pooling)하고, 차원을 축소시켜 모든 스캐닝 점군 데이터를 동일 방향에서 바라보고 식별하기 위한 변환행렬을 도출한다. 이 과정을 T-Net이라고 하며 도출된 변환행렬을 입력 데이터에 곱해 줌으로써 스캐닝 점군 데이터는 각기 다른 좌표축을 기준으로 생성된 형상의 데이터일지라도 기준 좌표축을 하나로 통일시켜 한 방향에서 바라보도록 변환시켜준다.

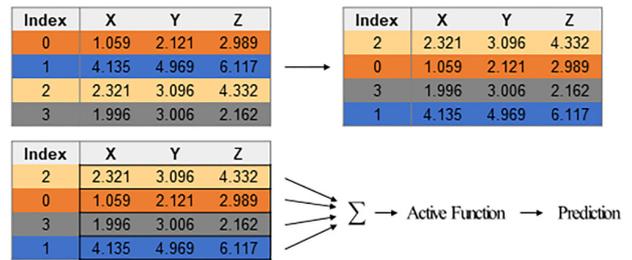


Fig. 8 Permutation invariance

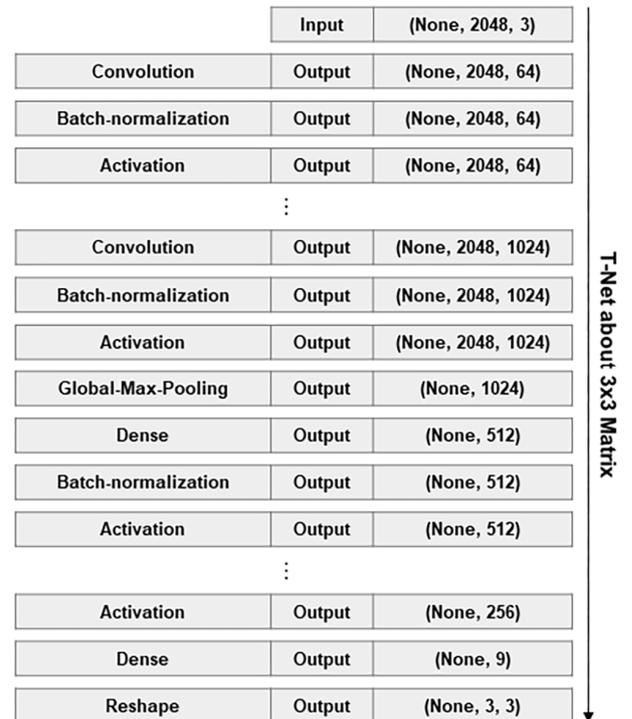
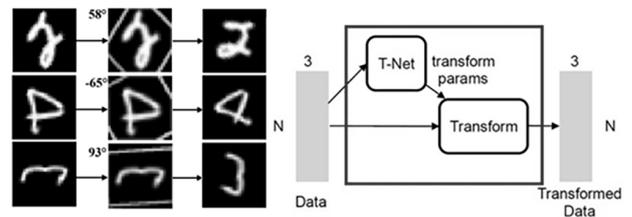


Fig. 9 Transformation network of PointNet architecture

4. PointNet 기반 딥러닝 아키텍처

4.1 스캐닝 점군 데이터를 활용한 객체 탐지

본 연구 목표는 LNGC 화물창 스캐닝 점군 데이터로부터 비계 시스템을 구성하는 스캐닝 점군 데이터 객체를 탐지하여 제거하는 것이다. 따라서 화물창의 모든 스캐닝 점군 데이터 중 비계 시스템을 구성하는 점을 예측하는 것이 1차 목표이다. 딥러닝 아키텍처는 x, y, z 3개의 차원을 가진 데이터가 입력되고 이들을 라벨링한 뒤, T-Net, MLP(Multi-Layer Perceptron) 반

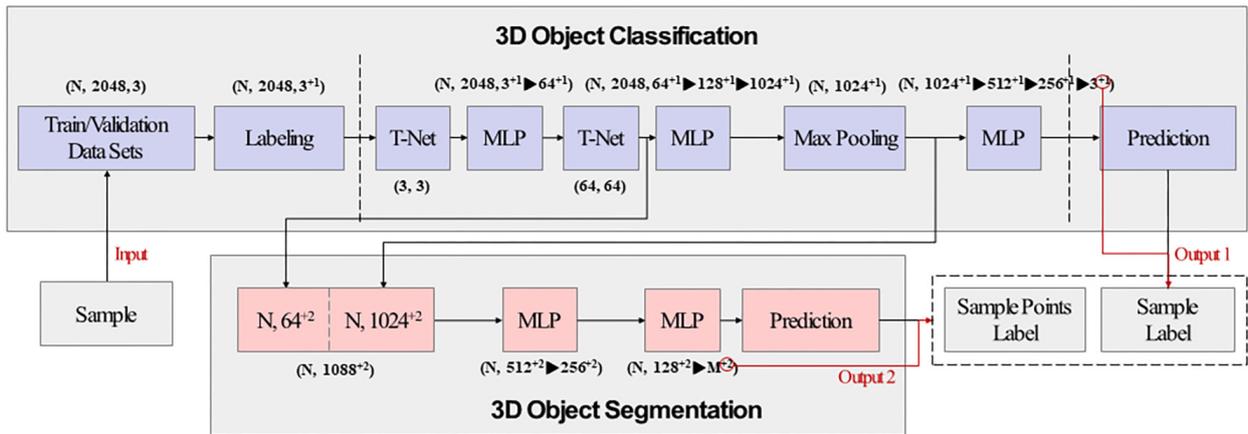


Fig. 10 PointNet architecture for 3D object detection

복, 최대 풀링, 다시 MLP, 예측(prediction) 순으로 진행한다 (Fig. 10). 이는 분류(classification)에 대한 아키텍처로 입력 데이터에 대해서 하나의 리벨만을 예측할 수 있기 때문에 분할(segmentation)을 통해 점 각각에 대한 리벨 예측을 수행한다.

분할 아키텍처는 분류 아키텍처와 크게 다르지 않지만 분류 아키텍처에서 최대 풀링 및 글로벌 레지스트레이션이 수행된 데이터 셋을 가져옴으로써 차이가 발생한다. 글로벌 레지스트레이션이 수행되면 딥러닝 아키텍처에 입력되는 각각의 점 데이터들을 하나로 합치게 되는데 이때 점 데이터들이란, 포인트 각각에 대한 리벨 정보를 가진 데이터들을 뜻한다. 글로벌 레지스트레이션이 수행된 데이터를 기존의 T-Net까지 수행된 데이터 셋에 붙여줌으로써 순서 불변, 변환 불변이 적용된 스캐닝 점군 데이터의 점 각각에 대한 리벨 정보를 맵핑시키는 것과 같다. 이로써 학습하고자 하는 모든 점 데이터와 맵핑된 리벨 정보를 학습하고 객체 탐지를 수행하기 위한 입력 데이터가 완성된다.

스캐닝 점군 데이터를 활용한 딥러닝 기반 3D 객체 탐지 기법은 PointNet이 가장 대표적으로, PointNet은 하드웨어 성능을 고려하여 특정 수준(3,000 Points) 미만의 데이터 예측을 초기 목표로 설정하였다. 본 연구에서는 딥러닝 모델 생성 시 발생할 수 있는 오류 최소화를 위해 PointNet에 사용된 데이터 파라미터를

레퍼런스로 활용하였으며, 입력 데이터 크기에 제한이 있어 화물창 데이터를 구역으로 쪼개는 전처리 작업을 수행하였다 (Fig. 11). 이에, 기존 4억 개 가량의 원본 스캐닝 점군 데이터에 구역을 지정하여 약 250만 개로 감소시켰고, 데이터를 샘플링하여 2,960개(약 1.2%)로 감소시켰다. 그 후, 감소된 스캐닝 점군 데이터로부터 C-wall, Leg support, Side support 3개로 파트로 분할하고자 하였다(Fig. 12). 3D 객체 탐지를 위해 감소된 스캐닝 점군 데이터를 라벨링 하였다(Fig. 13). 라벨링은 가상의 스캐닝 점군 데이터를 사용자가 바라보는 시각으로부터 평면으로 관통하고, 선택한 모든 영역의 점에 대해서 리벨을 지정하는 라쏘 선택(lasso selection)을 통해 수행하였으며

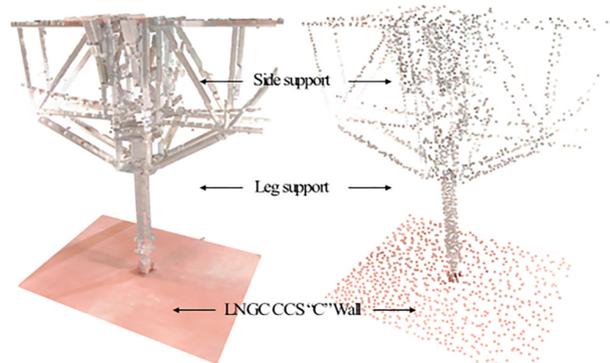


Fig. 12 Cropped and down sampled data (2,515,470 to 2,960 points)

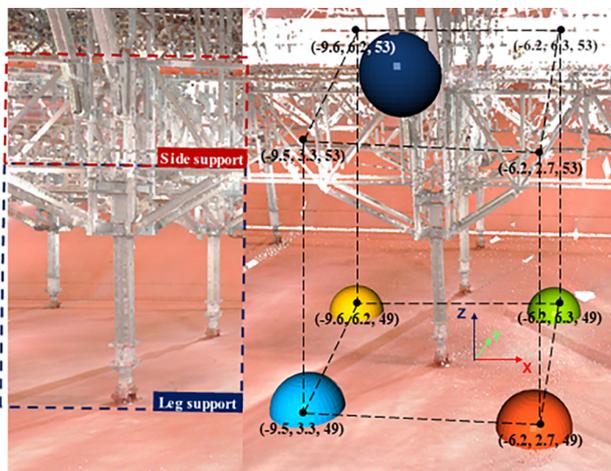


Fig. 11 Prediction area division (426,943,172 points)

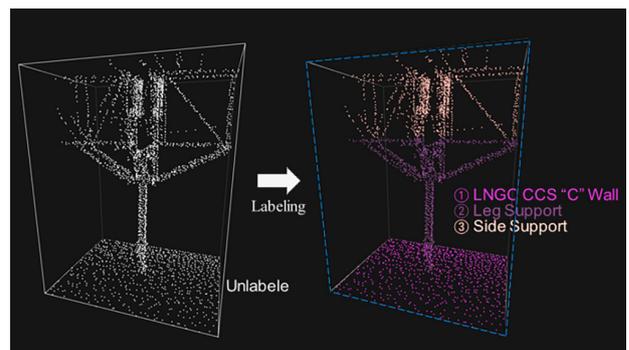


Fig. 13 Labeling for validation of prediction

라벨링된 데이터는 4.4절에서 예측 정확도를 검증할 때 활용하였다. 객체 탐지 기법의 성능 검증은 사전에 지정한 예측시퀀스 데이터의 라벨을 기준으로 삼고 수행하는데, 딥러닝 모델은 라벨 정보가 없는 데이터를 사전에 지정한 라벨과 같이 예측하도록 만드는 것이 최우선적인 목표이다. 이에, 모델 생성을 위해 데이터를 학습할 때에는 예측시퀀스 데이터의 라벨 정보를 활용하지 않는다.

4.2 학습/검증 데이터 생성

직전 수행한 예측 데이터를 모사하여 학습/검증용 데이터를 모델링 하였고 그 결과는 Fig. 14에 나타난다. PointNet에 사용된 데이터는 ModelNet10 데이터 셋으로, 이는 PointNet 이전의 선행연구에서 수행된 3D 모델 라벨링 자동화 기술이 활용된 데이터 셋이다. ModelNet10 데이터 셋은 16개 카테고리, 16,881개 데이터로 구성되어 있고 카테고리 별 데이터 개수와 지정된 라벨 개수(분할된 파트의 개수)의 차이가 있다. 이는 16종류의 특정 형상에 대한 정보만 저장된 데이터 셋이고 종류별로 예측 데이터를 위한 학습/검증 데이터가 구성되어 있기 때문에 비계 시스템 객체 탐지를 위한 입력 데이터로 활용이 불가하다.

본 연구 수행에 있어 학습/검증용 데이터 셋을 자동으로 라벨링하는 기술을 적용하는 것에 제약이 있었고 수동으로 라벨링하여 학습/검증용 데이터 셋을 생성시켜야 하므로 데이터의 크기를 최소화하면서 예측 정확도를 높이기 위한 방법을 탐색하였다.

PointNet에 사용된 ModelNet10 데이터 셋의 16개의 형상을 가진 종류들 중 데이터의 크기는 오름차순으로 모자, 로켓, 이어폰 순이었으나 이와 대응하는 예측 정확도 mIoU(mean Intersection over Union)는 82.5%, 57.9%, 73.0% 수준으로, 데이터의 크기가 가장 작지만 형상이 단순하고 2개의 라벨로 분할된 모자 형상 데이터 셋이 가장 높은 예측 정확도를 나타내었다 (Charles et al., 2017).

본 연구에서는 예측시퀀스 데이터를 모사하여 학습/검증용 데이터 셋을 생성함과 더불어 수작업으로 학습/검증용 데이터 셋을 라벨링하기 때문에 예측 정확도의 수준 가능성이 불가능하였다.

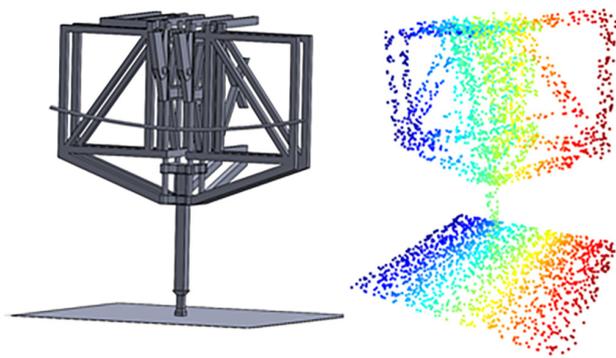


Fig. 14 Modeling by prediction data (left) Create a point cloud for learning (right)

따라서 예측정확도가 가장 낮은 로켓 형상 데이터 셋을 레퍼런스로 삼고 그의 결과 지표를 최소 목표수준으로 설정하였다. 로켓 형상 데이터는 80여 개로, 본 연구에서는 학습/검증 데이터를 80개 생성 및 라벨링하여 입력 데이터로 적용하였다(Fig. 14). 4.4절에서는 80개의 라벨링된 데이터 셋을 딥러닝 아키텍처에 적용하여 모델을 생성한 뒤 데이터를 예측시켜 산출되는 결과를 분석한다.

4.3 미니배치, 에포크 및 IoU

배치 크기(batch size)와 에포크(epoch)는 개발자가 설정할 수 있는 파라미터(parameter)로 두 개념은 컴퓨팅 시스템 하드웨어 한계로 인해 만들어졌다. 딥러닝은 초기 설정한 배치 크기를 기반으로 가중치와 편향으로 인한 손실을 줄여나가는 경사하강법(gradient descent)을 따르고, 업데이트하는 역전파 알고리즘(back-propagation algorithm)으로 수행되는데, 배치 크기가 너무 작으면 딥러닝 모델은 학습/검증 데이터에 과대적합(overfitting)될 수 있고, 배치 크기가 너무 크면 과소적합(underfitting)될 가능성이 크다. 적절한 배치 크기와 에포크 설정은 모델을 학습할 때 중요하고 보통 연구 초반에는 해당 연구 도메인에서 자주 사용하는 배치 크기를 기준으로 다양한 실험을 통해 최적의 값을 탐색한다.

IoU(Intersection over union)는 경계 상자가 의미하는 객체를 정확히 판단하기 위한 지표로서 기존에 생성한 경계 상자로부터 예측된 상자가 동일하다면 IoU는 100%를 나타낸다. 일반적으로 IoU가 50% 이상이면 '동일 객체'라고 판단하고, 본 논문에서는 점 데이터가 비계 시스템으로 예측되는지 검토하여 제거하는 것이 목표이므로 경계 상자 생성 작업을 보류하였다. 이에 대한 정확도 지표인 IoU는 추후, 경계 상자가 뜻하는 객체를 정확히 판단하기 위한 연구 수행에 적극 활용할 예정이다(Fig. 15).

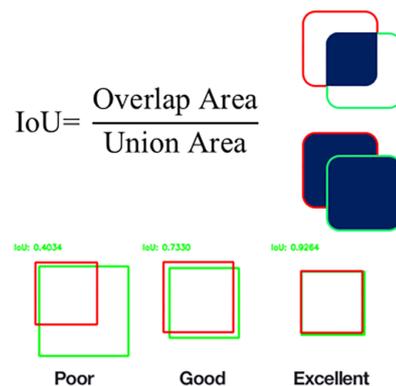


Fig. 15 Performance with intersection over union

4.4 3D 객체 탐지 및 결과 분석

본 연구는 해당 기법 도메인인 PointNet에서 활용된 배치 크기를 기반으로 모델을 생성하였다. 총 5가지 모델을 생성하였으

며 학습/검증 데이터 셋을 예측해야 할 원 데이터(LNGC 화물창 스캐닝 점군 데이터)와 동일하게 라벨링하고 일반적인 기계학습 데이터 분할 방법에 따라 80개의 데이터를 8:2 비율(학습:검증)인 64:16으로 분할하였다. 이를 입력 데이터로 적용하여 3D 분할 아키텍처에 따라 스캐닝 점군 데이터의 점 각각에 대한 라벨 예측을 수행하였다. 딥러닝은 배치 크기로 분할, 에포크 횟수만큼 반복 및 예측하도록 학습하는 과정으로 수행되고, 데이터를 학습하면서 사전에 설정한 에포크 횟수에 도달하면 학습이 마무리되고 모델이 생성된다. 기존의 라벨링이 완료된 데이터를 딥러닝 모델에 적용한 결과 C 모델의 예측 정확도인 IoU, Accuracy가 가장 높게 산출되었다. 이때, 정확도(accuracy)는 기존에 주어진 라벨(ground truth)에 대한 모델의 예측 정확도를 뜻한다. Ground truth는 Fig. 16과 같이 초기에 지정한 예측 시킬 데이터의 라벨 정보와 같으며 PointNet 기법을 통해 생성된 딥러닝 모델은 예측시킬 데이터의 라벨 정보가 없는 스캐닝 점군 데이터를 학습하여 Ground truth와 같은 라벨을 예측하도록 한다.

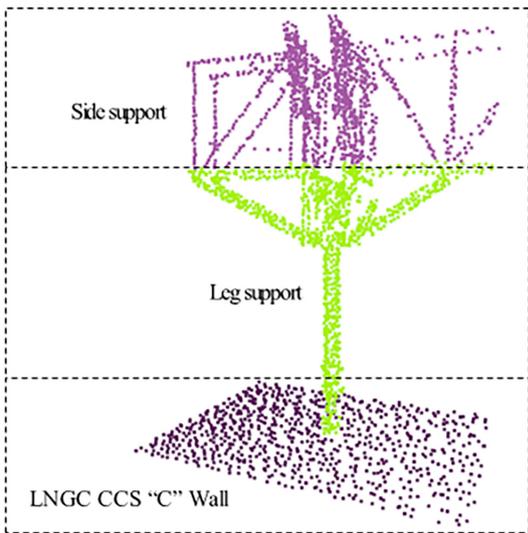


Fig. 16 Ground truth

일반적으로 에포크 횟수가 증가할수록 예측 정확도 또한 비례적으로 증가하는 경향이 있으나 본 연구에서 수행된 일부 특정 모델은 그러한 경향성을 보이지 않는다(Table 1; Fig. 17). 예측된 스캐닝 점군 데이터는 모델 A, B, E, D, C 순으로 예측 정확도가 점차 높아진다. C 모델의 예측 정확도는 약 70% 수준으로 타 모델보다 뛰어난 성능을 보인다. 학습/검증 데이터의 크기와 에포크 값(횟수)은 모든 모델이 동일하기 때문에 본 모델들의 정확도 차이는 배치 크기에 따른 결과라고 할 수 있다.

Fig. 18은 딥러닝 모델별로 예측된 스캐닝 점군 데이터를 정성적으로 검토하기 위해 시각화한 결과이다. A 모델은 C-wall, Leg support, Side support가 세로열 순으로 예측되었다. 본 연구에서 수행된 데이터는 3개의 파트가 가로열 순으로 예측되어야 한다. 하지만 이러한 오류가 발생하였는데, 이는 파트 간 경계 부분의 학습 데이터 부족, 복잡한 형상, 매우 낮은 배치 크기가 오직 원인으로 판단된다.

B 모델은 대부분의 영역이 C-wall로 예측되었다. Side support의 거의 모든 데이터가 C-wall로 예측되었으며 Leg support 하단부 기둥 데이터의 절반 이상이 C-wall로 예측되었다. 이는 모델 생성을 위해 학습하는 과정에서 모델이 C-wall 데이터에 과대적합 되어 다른 대부분의 데이터들을 C-wall로 예측한 것으로 판단된다.

C 모델은 가장 좋은 예측 정확도를 가지지만 Leg support 하단부와 Side support 일부를 C-wall로 예측한다. 이는 Leg support 하단부와 C-wall 사이의 경계 부근 학습 데이터의 부족, Side support 날개 부분 학습 데이터의 부족이 오직 원인으로 판단된다.

D 모델은 C-wall의 약 40%를 Side support로 예측한다. 이 영역만 올바르게 예측했다면 C 모델을 넘어서는 정확도를 가졌을 것이다. 이렇게 예측된 이유는 적정 수준을 넘어서는 배치 크기가 원인으로 판단된다.

E 모델은 대부분의 영역을 Leg support로 예측한다. 이는 타 모델 대비 가장 높은 배치 크기로 인해서 데이터를 학습할 때 점 데이터 간의 상호작용을 심히 건너뛰며 학습했기 때문이다.

Table 1 Model performance review

| Batch size | 1 (Model A) | | 2 (Model B) | | 4 (Model C) | | 8 (Model D) | | 16 (Model E) | |
|------------|-------------|----------|-------------|----------|-------------|----------|-------------|----------|--------------|----------|
| | IoU | Accuracy | IoU | Accuracy | IoU | Accuracy | IoU | Accuracy | IoU | Accuracy |
| 10 | 14.5% | 27.8% | 23.2% | 46.9% | 8.8% | 26.3% | 11.3% | 27.4% | 13.9% | 37.2% |
| 20 | 16.4% | 29.6% | 8.8% | 26.3% | 8.8% | 26.3% | 23.4% | 40.5% | 15.8% | 38.5% |
| 30 | 17.6% | 31.1% | 8.8% | 26.3% | 8.8% | 26.2% | 8.8% | 26.3% | 14.3% | 37.4% |
| 40 | 19.1% | 33.1% | 8.8% | 26.3% | 11.4% | 24.9% | 8.8% | 26.3% | 16.5% | 39.0% |
| 50 | 19.4% | 33.4% | 28.4% | 45.2% | 8.8% | 26.3% | 8.8% | 26.3% | 29.0% | 47.9% |
| 60 | 21.1% | 37.3% | 8.8% | 26.3% | 8.8% | 26.3% | 46.3% | 63.4% | 8.4% | 22.9% |
| 70 | 21.1% | 37.3% | 8.8% | 26.3% | 8.8% | 26.3% | 32.3% | 52.7% | 16.9% | 30.9% |
| 80 | 22.2% | 39.4% | 10.6% | 27.1% | 28.0% | 46.4% | 27.0% | 44.4% | 9.0% | 24.0% |
| 90 | 22.4% | 39.8% | 8.8% | 26.3% | 37.4% | 54.8% | 21.2% | 33.2% | 17.3% | 30.4% |
| 100 | 22.2% | 39.4% | 8.8% | 26.3% | 54.1% | 70.2% | 7.6% | 20.8% | 14.6% | 30.0% |

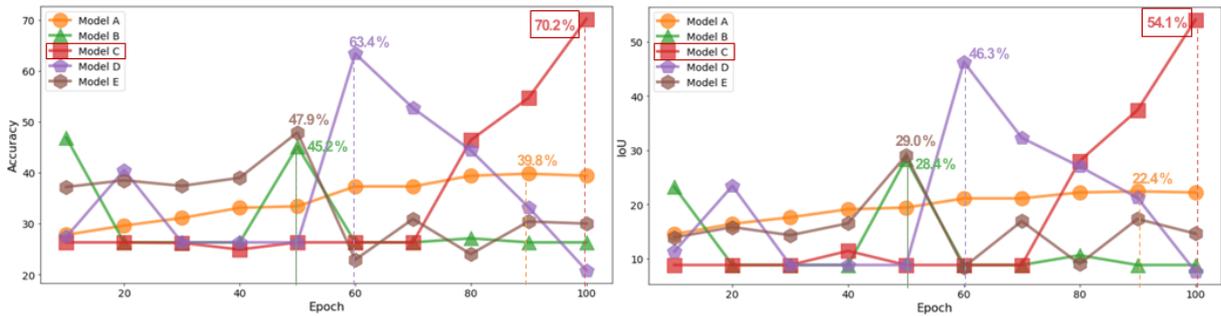


Fig. 17 Accuracy and IoU of deep learning models

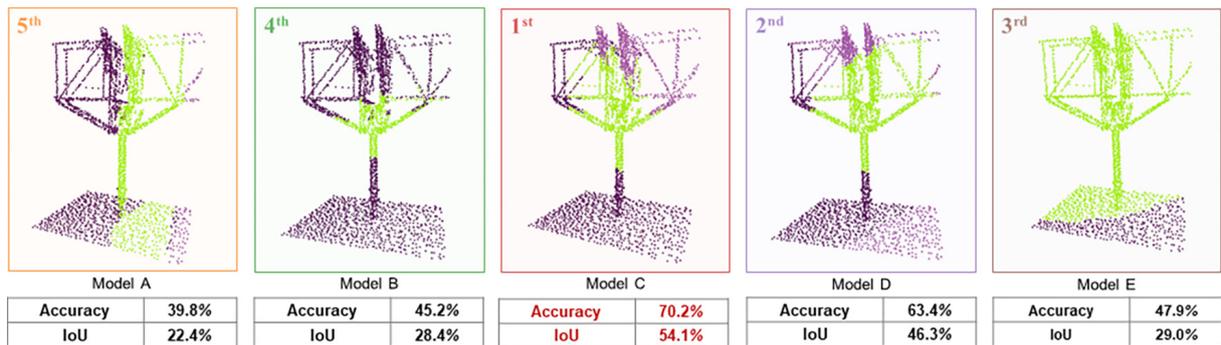


Fig. 18 Prediction results about deep learning models

이처럼 딥러닝을 통해 학습된 모델들은 서로 간에 성능 차이가 발생한다. 모델의 성능은 배치 크기, 에포크 횟수에 따라 달라지며 본 모델들의 성능 미비는 학습 데이터의 부족이 주 원인으로 파악되는데 이 내용은 4.5절에서 검토하였다.

4.5 딥러닝 모델의 오측 원인 검토

성능이 가장 좋은 C 모델은 약 29.8%의 예측 실패를 보인다. 본 연구의 목표는 비계 시스템 객체를 구성하는 스캐닝 점군 데이터의 제거 자동화를 통해 LNGC 화물창 스캐닝 점군 데이터만 남겨두어 복수 정합의 계산 비용을 절감하는 것이므로 C-wall로 예측된 스캐닝 점군 데이터를 제외한 모든 데이터를 제거하였다(Fig. 19). 이에, 2,960 스캐닝 점군 데이터 중 41.6%인 1,232 스캐닝 점군 데이터가 C-wall로 예측되었으나 그 중 비계 시스템을 구성하는 데이터 일부가 C-wall로 예측되는 오측이 발생하였다.

C 모델이 나타내는 70%의 정확도 수준은 Leg support를 Side support로, Side support를 Leg support로 오측한 결과도 반영되어 있다. 그러나 본 연구는 비계 시스템으로 예측된 데이터는 모두 제거할 것이므로 Leg support와 Side support가 서로 반대된 예측 결과를 나타내는 것은 목표 달성에 지장이 없다. 하지만 비계 시스템을 구성하는 점 데이터를 C-wall로 예측하는 것과 C-wall을 구성하는 점 데이터를 비계 시스템으로 예측하는 것은 비계 시스템을 구성하는 데이터를 제거할 때 오류를 초래한다. 이에, Fig. 20에는 C 모델이 나타낸 오측 데이터(빨간색 점 데이터)로써 비계 시스템을 구성하는 데이터이지만

C-wall로 오측되고, 비계 시스템 객체로 탐지된 데이터를 제거할 때 발생하는 악영향 데이터를 명시(Fatal error)하였다. C-wall로 예측된 1,232 포인트 중 진실된(True positive) C-wall 포인트는 780개로, 나머지 452 포인트는 비계 시스템이지만 C-wall로 오측(False positive)되었다. 딥러닝 모델을 통해 예측시킨 데이터는 총 2,960 포인트로, 그 중 15.3%를 오측하는 결과가 발생하였다. 이에, C-wall 학습 데이터는 예측시킬

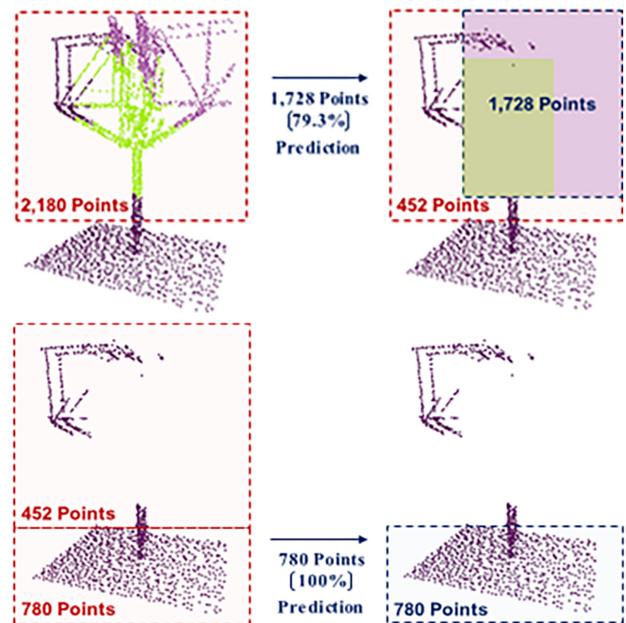


Fig. 19 Post-processing of prediction data(2,960 to 1,232 points)

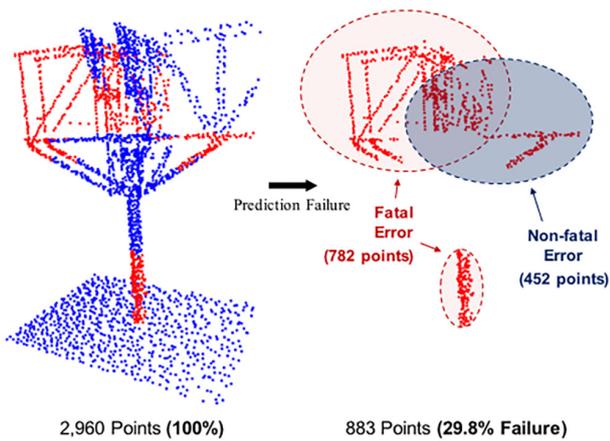


Fig. 20 Review of post-processing results

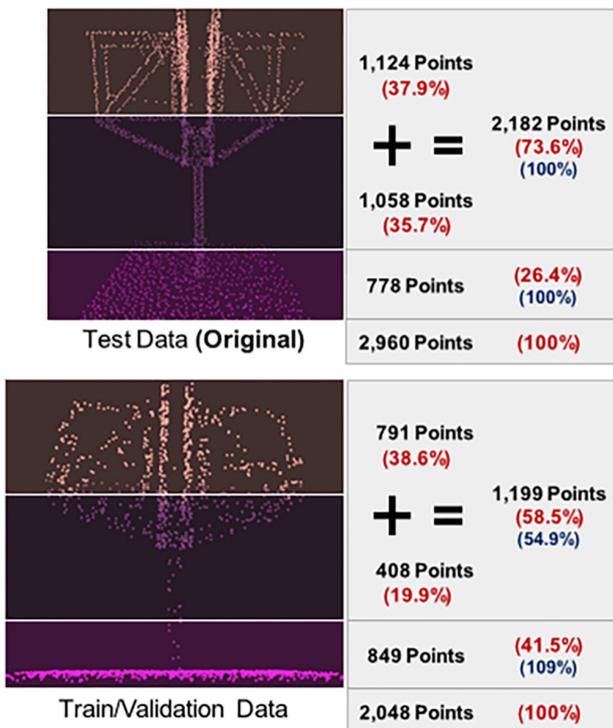


Fig. 21 Consideration of prediction accuracy

데이터에 준하는 양(109%)을 보유하고 있었으나 비계 시스템을 구성하는 Leg support, Side support의 학습 데이터는 예측시켜야 할 데이터에 비해 45.1% 부족하였으며, 그렇기 때문에 학습 데이터가 부족한 영역의 예측에 대한 오차 비율이 높았던 것으로 사료된다(Fig. 21).

5. 결론 및 향후 연구방향

본 연구를 통해 스캐닝 점군 데이터를 활용한 딥러닝 기반 3D 객체 탐지 기법의 성능을 대략적으로 검토하였다. 비계 시스템을 구성하는 데이터를 제거하고자 함에 있어 70% 수준의 예측 정확도를 확보하였고 학습 데이터의 크기, 배치 크기, 에포크 값 등 특정 조건을 전제하였기 때문에 파라미터 조절을 통

한 예측 정확도 향상 또한 가능성이 있다. PointNet 기법에서 활용된 로켓형상 데이터의 경우 딥러닝 모델의 성능으로 약 58%의 IoU를 나타내지만 LNGC 화물창 데이터의 IoU는 54% 수준으로 최소 목표수준에 도달하지 못하였다. 하지만 일반적인 동일 객체 판단 기준은 50% 수준이고, 비계 시스템 학습 데이터가 상당히 부족하다는 것을 검토하였기 때문에 향후 연구에서는 이를 보완 및 개선하여 안정적인 예측 수준의 결과를 산출할 수 있을 것으로 예상된다. 선박 건조 공정에서 수작업으로 진행되어온 기존 작업들에 객체 탐지 기법을 적용하면 반자동 작업수행이 가능하고 모델 생성이 완료되면 예측하는 것은 곧바로 처리된다는 장점이 있다. 또한, 객체 탐지 기법은 딥러닝을 기반하기 때문에 찾고자 하는 객체의 특징(모서리, 꼭짓점 등)을 찾아 탐지하며, 따라서 LNGC 화물창의 비계 시스템뿐만 아니라 곡면과 같은 기존 알고리즘으로 처리하기 힘들거나 불가능한 형상 등 복잡한 요소에도 제약 없는 적용이 가능하다.

타 산업에서는 2D 객체 탐지 기법을 제품의 정도 관리에 적용하여 상용화까지 진행된 사례가 있으며 심층 신경망 학습을 위한 데이터를 보완 및 개선하고 파라미터 조절을 통해 모델의 성능을 최적화한다면 객체 탐지 기법을 조선 산업에 적용하는 것이 가능하다고 판단된다. 향후 연구에서는 선박의 선수 및 선미부 곡률의 정도 관리 또는 내업공장의 조립 공정 진도를 모니터링 자동화 등 선박 건조의 생산성 고도화를 위해 인공지능 기술을 융합하고자 지속적인 연구를 수행할 예정이다.

References

Charles, R.Q. et al., 2018. Frustum PointNets for 3D object detection from rgb-d data. *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, arXiv: 1711.08488.

Charles, R.Q. Hao, S. Kaichun, M. & Guibas, L.J., 2017. PointNet: Deep learning on point sets for 3D classification and segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, arXiv: 1612.00593.

Ham, D. Lee, P. & Woo, J.H., 2016. A study of method for machine learning application for using shipyard big data. *Journal of Computational Design and Engineering*, pp.186-190.

Jaderberg, M. Simonvan, K. Zisserman, A. & Kavukcuoglu, K., 2015. Spatial transformer networks. *Part of Advances in Neural Information Processing Systems 28*, paper ID:1213. Montreal, Canada, 7 - 12th December 2015.

Jeong, Y.C. et al., 2021. Mechanical properties analysis of epoxy and polyurethane adhesive for accurate structural analysis of lng cargo hold. *Journal of the Society of Naval Architects of Korea*, 58(2), pp.66-72.

Kim, Y.J., Kim, T.W., Yoon, J.S. & Kim, M.K., 2019. Study

- of the construction of a coastal disaster prevention system using deep learning. *Journal of Ocean Engineering and Technology*, 33(6), pp.590–596.
- Kirillov, A. et al., 2019. Panoptic segmentation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, arXiv: 1801.00868.
- Kwon, H. & Ruy, W., 2020. A study on the work-time estimation for block erections using stacking ensemble learning. *Journal of the Society of Naval Architects of Korea*, 56(6), pp.488–496.
- Lee, D.K., Shin, J.G., Kim, Y. & Jeong, Y.K., 2014. Simulation-based work plan verification in shipyards. *Journal of Ship Production and Design*, 30(2), pp.49–57.
- Maturana, D. & Scherer, S., 2015. VoxNet: A 3D convolution neural network for real-time object recognition. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 922–928. Hamburg, Germany, 17 December 2015.
- Russakovsky, O. et al., 2015. ImageNet large scale visual recognition challenge. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, arXiv: 1409.0575.
- Woo, J.H., Hwang, Y.S., & Nam, J.H., 2016. An approach for construction of shipyard simulation environment based on neutral file format. *Journal of the Society of Naval Architects of Korea*, 53(1), pp.18–28.
- Yin, Z. & Tuzel, O., 2017. VoxelNet: End-to-end learning for point cloud based 3D object detection. *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, arXiv: 1711.06396.

