

해상교통 상황인지 향상을 위한 합성 데이터셋 구축방안 연구

이영채* · 박세길**

A Study on Synthetic Dataset Generation Method for Maritime Traffic Situation Awareness

Youngchae Lee* · Sekil Park**

Abstract

Ship collision accidents not only cause loss of life and property damage, but also cause marine pollution and can become national disasters, so prevention is very important. Most of these ship collision accidents are caused by human factors due to the navigation officer's lack of vigilance and carelessness, and in many cases, they can be prevented through the support of a system that helps with situation awareness. Recently, artificial intelligence has been used to develop systems that help navigators recognize the situation, but the sea is very wide and deep, so it is difficult to secure maritime traffic datasets, which also makes it difficult to develop artificial intelligence models. In this paper, to solve these difficulties, we propose a method to build a dataset with characteristics similar to actual maritime traffic datasets. The proposed method uses segmentation and inpainting technologies to build a foreground and background dataset, and then applies compositing technology to create a synthetic dataset. Through prototype implementation and result analysis of the proposed method, it was confirmed that the proposed method is effective in overcoming the difficulties of dataset construction and complementing various scenes similar to reality.

Keywords : Maritime Safety, Maritime Traffic, Situation Awareness, Synthetic Dataset, Deep Learning

Received : 2023. 11. 20. Revised : 2023. 12. 18. Final Acceptance : 2023. 12. 26.

※ This paper was carried out under the basic project "Development of open platform technology for smart marine safety and corporate support" of the Ship & Offshore Plant Research Institute with funding from the Ministry of Oceans and Fisheries (1525014880, PES4880).

* First Author, Graduate School of Environmental Engineering, Chungnam National University, e-mail : dudco5858@gmail.com

** Corresponding Author, Principal Researcher, KRISO (Korea Research Institute of Ships & Ocean Engineering), 32 Yuseong-daero 1312beon-gil, Yuseong-gu, Daejeon, 34103, Korea, Tel : +82-42-866-3647, e-mail : skpark@kriso.re.kr

1. 서 론

선박 충돌사고는 인명 손실과 재산상의 피해는 물론 해양오염을 일으키기도 하며 상황에 따라 국가적 재난으로 연결되기도 하여 예방이 매우 중요하다. 이러한 선박 충돌사고의 대부분은 해기사의 경계 소홀 및 부주의로 인한 인적요인에 의한 것으로 많은 경우 장비 및 시스템의 지원을 통한 예방이 가능하다. 그 동안에도 자율운항선 연구를 포함하여 항해 선박 주변의 해상교통 상황을 인지하고 이를 해기사에게 전달하여 충돌사고를 예방하려는 알고리즘과 시스템이 다수 제안되었다[Oh et al., 2016; Jeon et al., 2019; Ali et al., 2023]. 그러나 해상의 선박 및 장애물을 정확하고 신뢰성 있게 탐지 및 식별, 추적하는 부분은 여전히 개선이 필요한 상황으로 최근에는 다양한 분야에서 혁신적인 변화를 가져오고 있는 인공지능 기술을 활용하여 이를 해결하려는 연구들이 활발히 진행되고 있다[Er et al., 2023].

다만 우수한 성능의 인공지능 모델의 개발은 태스크에 적합한 모델 아키텍처의 선택, 하이퍼파라미터 튜닝, 정규화 등 여러 기술적 접근 방법만으로는 한계가 있으며 인공지능 모델 학습용 데이터셋이 무엇보다 중요하다. 데이터셋은 훈련과 시험, 검증의 역할을 수행하며 인공지능 모델의 정확도와 학습 결과를 좌우한다. 데이터셋은 양과 품질이 모두 중요한데 인공지능 모델이 풀고자 하는 태스크와 관련된 보다 많은 상황과 패턴을 정확하게 학습할 수 있도록 다양성을 확보하여 준비해야 우수한 성능의 인공지능 모델 개발이 가능하다. 그러나 데이터셋 구축은 구체적인 목표와 계획을 바탕으로 데이터의 수집과 정제, 레이블링 등 시간과 비용이 많이 소요되는 과정을 거쳐야하기 때문에 대부분의 경우 매우 어려운 작업이며, 때로는 원하는 데이터를 얻는 것이 현실적으로 불가능할 수 있다.

해상에서의 데이터셋 구축은 특히나 더 어려운데 육상에서 먼 바다의 선박을 촬영하는 경우 깨끗한 영상을 얻기 어렵고 선박의 특정 부분만 획득 가능한 문제점이 있다. 또한 바다가 매우 넓고 수심이 깊어 해상 곳곳에 고정형 CCTV와 같은 카메라를 설치하는 것이 거의 불가능하며 선박이나 드론을 이용하더라도 원하는 선박이 출항하지 않으면 데이터를 획득하기가 어렵고 육상에서의 데이터셋 구축과 비교하였을 때 시간과 비용

이 많이 소요되는 문제점이 있다. 이에 대한 해결책으로 실제 데이터는 아니지만 실제 데이터와 유사한 특성을 갖도록 제작된 합성 데이터셋의 활용을 고려해 볼 수 있다. 합성 데이터셋은 가상의 데이터이므로 데이터의 양과 다양성을 확보할 수 있으며 획득이 어려운 상황에 대한 데이터도 확보가 가능하여 데이터셋 구축의 어려움을 극복하는데 효과적이다.

본 논문에서는 선박 충돌사고 예방을 목표로 해상교통 상황인지 분야와 같이 데이터셋 구축이 어려운 환경에서 실제와 유사한 특성의 데이터셋을 구축하는 방법을 제안한다. 제안하는 방법은 크게 두 단계로 구성된다. 첫 번째 단계는 전경 데이터셋과 배경 데이터셋을 구축하는 단계이며, 두 번째 단계는 전경 데이터셋과 배경 데이터셋을 합성하여 실제와 유사한 가상의 장면을 구성하는 단계이다. 본 연구에서는 각 단계 구현을 위해 적합한 인공지능 모델을 도입하고 실제 구현 및 분석을 통해 제안하는 방법이 데이터셋 구축의 어려움을 극복하는데 효과적임을 확인하였다. 특히 제안하는 방법은 실제 데이터를 기반으로 데이터셋을 구축하고 다양한 시나리오의 장면을 생성할 수 있는 장점이 있다.

2. 연구 배경

2.1 인공지능 학습용 데이터

인공지능 학습용 데이터는 기준에 따라 달라질 수는 있으나 크게 정형데이터와 비정형데이터, 반정형데이터로 분류 가능하다. 정형데이터는 미리 정해 놓은 형식과 구조에 따라 저장되도록 구성된 데이터로 관계형 데이터베이스의 테이블, 스프레드시트 데이터, CSV 등의 데이터를 말한다. 데이터에 대한 부분 검색, 선택, 갱신, 삭제 등의 연산을 손쉽게 수행할 수 있어 정형화된 업무 또는 서비스에 사용하며, 사용자가 쉽게 이해하고 쉽게 시스템에 적용 가능하도록 잘 알려진 포맷 또는 명확한 데이터 구조 표현 방법을 사용한다는 특징이 있다[TTA structured data, 2023].

비정형데이터는 정의된 구조 없이 정형화되지 않은 데이터로, 동영상 파일, 오디오 파일, 사진, 보고서, 메일 본문 등이 해당된다. 데이터의 특징을 추출하여 정형데이터로 변환하는 전처리 과정이 필요하고, 지속적으로 생성되는 데이터 중 가장 큰 비중을 차지하며 데

〈Table 1〉 Comparing the Strengths and Weaknesses of Synthetic Data and Real Data

Type	Strengths	Weaknesses
Synthetic Data	<ul style="list-style-type: none"> - Increase data diversity and minimize bias - Easy to build dataset - Fast deployment and creation speed - Automatic data generation is possible, allowing unlimited creation, allowing the creation of large-scale datasets - Capable of generating data that is actually difficult to obtain and under desired conditions 	<ul style="list-style-type: none"> - Proper model learning may not occur, resulting in subtle differences from the actual image, resulting in a domain gap - Original data cannot be replaced in cases of low utility - Generalization can be difficult when training a model using only synthetic datasets - Large differences between synthetic and real data result in poor model performance and increased likelihood of prediction errors
Real Data	<ul style="list-style-type: none"> - Ground Truth Data, realistic results - High predictability of how it will behave - Provides good insight - Data quality is likely to be good 	<ul style="list-style-type: none"> - The time and cost required to collect data is relatively high compared to synthetic data - It is difficult to secure various data - Data is often obtained by chance and contains inaccurate elements or bias - The pre-processing process requires a lot of cost due to personal information protection, etc.

이터가 생성되는 시간대의 다양한 사회적, 문화적, 공간적 현상들을 반영하고 있기에 마케팅이나 비즈니스 인텔리전스 등의 분야에서 매우 중요한 데이터이다 [TTA unstructured data, 2023].

반정형데이터는 정형데이터와 비정형데이터의 사이의 특성을 가진다. 반정형데이터는 데이터의 형식과 구조가 변경될 수 있는 데이터로, 구조 정보를 데이터와 함께 제공하는 파일형식의 데이터이다 [TTA semi-structured data, 2023]. 상기와 같은 특성으로 인해 일반적으로 합성 데이터셋 구축 시 비정형 데이터에 대한 합성 데이터 생성이 보다 난해하다.

2.2 합성 데이터

합성 데이터는 실제 데이터의 수나 종류가 부족하거나 획득이 어려운 경우 인공지능 학습용 데이터를 보완하는데 효과적이다. 〈Table 1〉은 일반적으로 알려져 있는 합성 데이터와 실제 데이터 각각의 장단점을 비교하여 정리한 것이다.

2.3 합성 데이터 생성 기술

지금까지 다수의 합성 데이터 생성 기술들이 연구되었는데 이들을 접근 방법에 따라 구분하면 크게 데이터 증강(data augmentation) 기반 접근법, 시뮬레이션

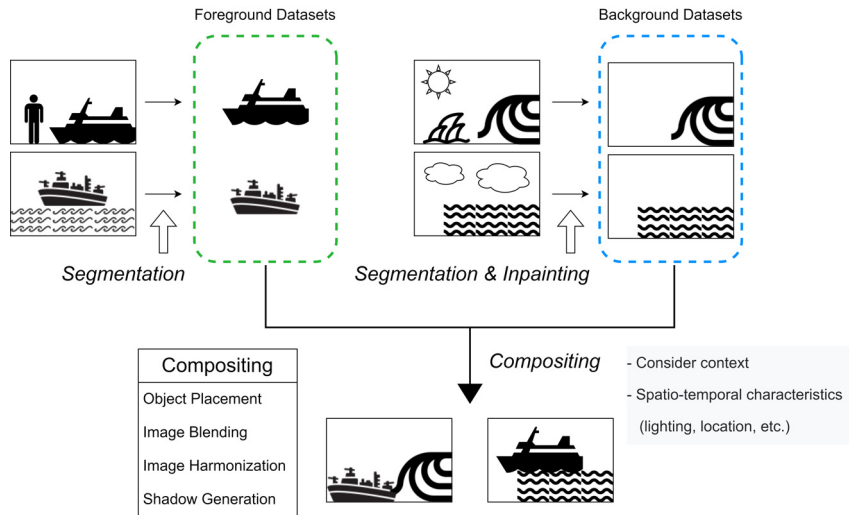
(simulation) 기반 접근법, 생성 모델(generative model) 기반 접근법으로 나눌 수 있다.

데이터 증강 기반 접근법은 원본 데이터에 노이즈 삽입, 색상 및 질감 변경, 자르기, 가르기 등의 변형을 가한 데이터를 학습데이터로 추가 활용하여 다양성을 보강하는 방법이다. 시뮬레이션 기반 접근법은 게임엔진을 활용하여 실제와 비슷한 형상의 데이터를 확보하거나 CAD 등을 이용해 실제와 유사한 데이터를 생성하는 기술이다. 그리고 생성 모델 기반 접근법은 기존에 주어진 소량의 학습데이터의 통계적 특성과 확률 분포를 학습하여 AI 알고리즘으로 유사한 특성을 갖는 가상의 합성데이터를 생성하는 기술로, GAN(Generative Adversarial Network)과 VAE(Variational Auto Encoder)가 대표적인 모델이다 [Lee and Ma, 2023].

본 논문에서는 비정형 데이터인 해상교통 이미지에 대한 증강을 위해 상기한 기존의 합성 데이터셋 생성 기술들을 조합하여 적용이 가능한 유연한 구조의 새로운 합성 데이터 생성 기술을 제안한다.

3. 합성 데이터셋 생성 방안

제안하는 합성 데이터셋 생성 방안은 크게 두 단계로 구성된다. 첫 번째 단계는 전경 데이터셋과 배경 데이터셋을 구축하는 단계이며, 두 번째 단계는 전경 데이터셋과 배경 데이터셋을 합성하여 실제와 유사한 가



〈Figure 1〉 Flowchart of Creating Foreground Datasets and Background Datasets

상의 합성 장면을 생성하는 단계이다.

〈Figure 1〉은 이러한 전경 데이터셋과 배경 데이터셋을 구축하고 합성 이미지 데이터를 구성하는 과정을 나타낸 흐름도이다. 전경 데이터셋은 해상에 존재하는 선박이나 항로표지 등 선박 항해 시 탐지, 추적 및 회피가 필요한 객체들의 데이터셋으로 배경색이 투명하게 처리된 이미지 형태로 구축되며, 원하는 객체를 추출해내기 위해 Segmentation 단계를 거치게 된다. 배경 데이터셋은 해상 또는 항만 배경의 이미지들로 필요에 따라 이미지에 포함된 불필요한 객체를 사전에 제거한 이미지 모음이다. 불필요한 객체를 추출 및 제거하기 위해 Segmentation 단계를 거친 후 객체가 제거된 부분의 배경을 복원하기 위해 Inpainting 단계를 수행한다. 합성 장면 생성 단계는 구축된 전경 데이터셋과 배경 데이터셋을 조합하여 자동으로 자연스럽게 합성되도록 처리되는데, 이 때 배경 데이터셋의 조명이나 위치와 같은 시공간적 특성과 문맥을 고려하여 전경 데이터셋을 합성하도록 Compositing 기술이 활용된다. Compositing 과정에서 전경 객체를 위치시키는 Object Placement와 경계를 다듬어주는 Image Blending, 조명 특성을 호환시켜주는 Image Harmonization, 그림자를 생성하는 Shadow Generation을 통해 더욱 사실적인 합성 이미지를 얻을 수 있다. 전경 데이터셋 및 배경 데이터셋 구축은 경우에 따라 실제 사진 외에 가상 데이터나 생성형 인공지능을

활용한 이미지를 사용할 수도 있다.

3.1 전경 데이터셋 생성

본 연구에서는 전경(해상 장애물 객체) 데이터셋 생성을 위해 Segmentation 기술을 활용하였다. Segmentation은 이미지 내 픽셀의 특징을 기반으로 하나의 이미지를 다수의 부분이나 영역으로 분할하는 기법으로, 디지털 영상 처리와 분석에서 널리 사용된다[MathWorks, 2023]. Segmentation은 다시 semantic segmentation과 instance segmentation로 분류할 수 있다. Semantic segmentation은 이미지 상의 픽셀별로 각각 어떤 클래스에 속하는지 레이블을 구해주는 방법으로 이미지 사이즈로 클래스 개수만큼 출력 채널이 존재하여 같은 클래스의 객체들에 대해 서로 구분이 불가능하다는 특징을 가진다. Instance segmentation은 이미지 상의 각 픽셀별로 객체 존재 여부를 계산하며 관심영역(ROI)별로 클래스 개수만큼 출력 채널이 존재하여 같은 클래스의 객체여도 구분이 가능하다는 특징을 가진다[Kirillov et al., 2019].

본 연구에서는 instance segmentation을 진행하기 위해 픽셀 단위로 정확한 객체의 모양과 윤곽을 추출하며 학습하지 않은 물체도 분할 가능한 제로샷(zero-shot) 모델인 SAM(Segment Anything Model)을 활용하

였다. SAM은 Image encoder와 flexible Prompt encoder, fast Mask decoder로 구성되어 있다. Image encoder에서 MAE pretrained ViT 모델을 이용하여 이미지를 삽입하여 임베딩(embedding)을 얻고, Prompt encoder에서 Sparse prompt (points, boxes, text)와 Dense prompt(mask)라는 프롬프트의 종류를 고려하여 다른 방법으로 임베딩을 얻는다. 이후 Mask decoder에서 Image embedding과 Prompt embedding을 입력으로 받아 마스크를 예측 및 출력한다[Kirillov et al., 2023].

〈Figure 2〉는 해상교통 이미지로부터 SAM을 활용하여 선박 객체를 추출한 결과 이미지이다.



〈Figure 2〉 Segmentation Results

실험 결과 원본 이미지가 선박을 원거리에서 촬영하여 작은 크기를 가지거나, 선박의 형태가 비교적 간단한 경우에는 SAM을 활용한 객체 추출이 효과적으로 이루어짐을 확인하였다. 이러한 경우, 객체 추출 결과물은 높은 정확성을 보이며 깔끔하게 배경을 분리해 내었다. 반면에 원본 이미지 상 선박이 크거나 주변에 컨테이너가 많은 경우, 복잡한 구조물로 인해 세부적인 세그먼트가 필요한 경우에는 배경 제거가 불완전하거나 객체의 일부가 잘리는 등의 문제가 발생하였고 관련 데이터셋을 이용하여 파인튜닝을 통한 모델 업데이트가 필요하였다.

3.2 배경 데이터셋 생성

배경 데이터셋 생성은 SAM을 이용하여 객체를 선택한 이후 Inpainting을 통해 배경을 복원하는 방식으로 진행하였다. Inpainting은 이미지나 비디오에서 훼손된 영역을 훼손되지 않은 영역의 정보를 이용하여 특징을 추출하고 최대한 의미적, 인지적으로 자연스러운 이미지를 생성하는 이미지 복원의 한 종류이다 [Kim et al., 2021].

본 연구에서는 Inpaint Anything Model의 연구 결과를 활용하였고 이 중 특정한 객체를 제거할 수 있는 Remove Anything 기능을 활용하였다. 사용자가 이미지에서 제거하려는 객체를 클릭하여 선택하면 SAM을 통해 클릭 위치를 기준으로 객체 자동 분할을 수행한 후 생성된 마스크 예측 결과를 Inpainting Model의 입력으로 사용하여 제거 후 채울 객체 영역에 대한 표시기를 제공해준다. 이후 LaMa와 같은 SOTA Inpainter로 제거된 영역을 시각적으로 자연스러운 구조와 질감으로 대체하는 과정을 거친다. 해당 전체 과정동안 사용자는 이미지에서 제거하려는 객체를 클릭하기만 하면 된다는 편리함을 가진다[Yu et al., 2023].

〈Table 2〉 Image Inpainting Results

Before	After

〈Table 2〉는 해상교통 이미지로부터 Inpaint Anything 모델로 선박 객체를 제거하고 배경을 복원한 결과인 배경 이미지 데이터이다.

실험 결과 선박이 비교적 간단한 형태를 가지며 배경이 단조로운 경우와 세밀한 세그먼트 및 Inpaint 작업이 필요하지 않은 경우, 객체 추출은 효과적으로 수행되었으며, 객체가 자연스럽게 배경과 통합되어 원하는 결과물을 얻을 수 있었다. 그러나 원본 이미지에서 선박의 크기가 크거나 세부적인 세그먼트 및 객체 선택이 필요한 경우에는 일부 부분만 복원되거나 Inpaint

작업이 어색하게 나타나기도 하였다. 이러한 현상은 Inpainting 모델이 해양 이미지와 관련된 학습 데이터가 충분하지 않아 발생한 것으로 이를 해결하기 위해서는 공개 Inpainting 모델에 대한 수정이 필요하였다.

3.3 합성 장면 생성

합성 장면 생성은 Image Compositing 기술을 활용하여 구현하였다. Image Compositing은 한 이미지에서 전경을 잘라 다른 이미지에 붙여 넣어 합성 이미지를 생성하는 기술이다[Niu et al., 2021]. 이때 자연스러운 합성 이미지를 생성하기 위해 본 연구에서는 Object Placement, Image Blending, Image Harmonization, Shadow Generation 순서대로 Compositing을 진행하였고 이미지 붙여넣기나 블렌딩(blending)에 비해 보다 자연스러운 이미지를 얻을 수 있도록 하였다.

3.3.1 Object Placement

Object Placement는 적절한 위치, 크기 및 모양으로 전경 객체를 배경에 배치하는 과정이다. 객체 배치는 객체 크기가 부적절하거나 전경 객체에 의해 불합리하게 배경 물체가 가려지는 경우, 객체가 합리적인 지지력을 갖지 않는 경우, 객체가 의미상 불합리한 장소에 위치 또는 객체와 배경 사이의 일관되지 않은 관점 등과 같은 요소들을 고려하여야 한다. 이처럼 Object Placement는 이미지 분석을 통해 장면을 이해하고 배치가 필요하여 난이도 높은 과정이다. 본 연구에서는 해상 영역을 고려한 스케일링과 배치를 진행하기는 하였으나 기술적으로 완벽한 배치에 어려움이 있어 합성 장면 생성 후 수동으로 적격 여부를 판단토록 진행하였다.

3.3.2 Image Blending

Image Blending은 이미지 전체의 투명도 값을 조정하여 합성하는 일반적인 블렌딩과 다르게 전경 객체와 배경 사이의 급격한 강도 변화와 흐릿함을 줄이고 경계를 다듬어줌으로써 부자연스러운 경계를 해결해주는 과정이다. 이는 다음 과정인 Image Harmonization과 함께 이미지의 이질감을 줄여주는데 효과적이며 매우

중요한 과정이다.

3.3.3 Image Harmonization

〈Table 3〉 Image Harmonization Results

	Before	After
Good cases		
Bad cases		

Image Harmonization은 배경과 전경의 조명 특성이 호환되도록 배경에 따라 전경 이미지의 외관을 조정하는 작업으로 전경 이미지를 배경에 조화롭게 통합하는 과정이다. Image Harmonization을 수행하기 위해 지금까지 딥러닝을 활용한 다양한 방법들이 개발되었는데 CNN 분류기, CDTNet, RainNet, IiH, iSSAM, DoveNet 등이 이에 해당한다[Niu et al., 2021].

본 연구에서는 INR-Harmonization 모델을 사용하였다. 이 모델은 추가 전역 기능을 중간 계층에

통합시켜주는 인코더와 암시적 신경 표현(INR) 패러다임을 사용하며, MLP 스택을 활용하는 디코더를 결합한 방법으로 조밀한 픽셀 간의 변환을 통해 고해상도(HR) 이미지를 조화하고 암시적 신경 표현을 활용한다[Chen et al., 2023]. 이를 통해 INR-Harmonization 모델은 배경과 전경의 시각적 호환성을 더욱 향상시키고, Image Harmonization 작업에서 암시적 신경 표현을 효과적으로 활용하여 높은 품질의 결과물을 얻을 수 있도록 제한한다. <Table 3>은 구축한 전경 데이터셋과 배경 데이터셋을 활용하여 INR-Harmonization 모델로 이미지 조화를 진행하기 전과 후의 이미지를 조화가 잘 이루어진 경우와 그렇지 못한 경우로 분류하여 나타낸 표이다.

실험결과 Image Harmonization이 잘 된 경우에는 배경 이미지의 객체 색상, 조도 등에 맞춰서 Harmonization이 진행되어 배경과 전경이 자연스럽게 어우러지는 결과를 확인할 수 있었다. 또한 동일한 객체를 각각 다른 배경 이미지에 합성하는 경우에도 각 배경의 조명 및 색감에 맞춘 다양한 결과를 얻었다. 반면에 배경 조명이 매우 복잡하거나 배경과 전경 사이의 조명 통계가 크게 다른 경우 Harmonization 작업에 어려움을 겪어 색상과 조도 등에서 어색한 결과를 얻었다. 좋지 않은 결과(Bad cases) 중 첫 번째 케이스는 이미지 조화 후 전경 이미지가 너무 어두워져서 전경 이미지와 배경 이미지의 색감이 유사하지 않았고, 두 번째 케이스의 경우에는 전경이 배경 이미지 내에 있는 다른 선박과 비교하였을 때 너무 밝아 보여 전경과 배경 이미지의 색감이 유사하지 않은 결과가 나온

것을 확인할 수 있었다. 다만 이러한 결과는 판정하는 사람의 주관에 따라 차이가 발생할 수 있는 부분으로 판단된다.

3.3.4 Shadow Generation

Shadow Generation은 그림자가 없어 합성 이미지가 비현실적으로 보이는 문제를 해결하고 보다 사실적으로 만들기 위해 배경 조명 정보를 활용하여 전경 객체에 자연스러운 그림자를 생성하는 작업이다. 이를 위한 딥러닝 기반 방법으로는 SGRNet, ARShadowGAN, ShadowGAN, SSN 등이 있다[Niu et al., 2021].

실험결과 SGRNet이 상대적으로 설득력 있는 그림자를 생성하는 결과를 보였으나 그림자의 모양이 비현실적이고 개선이 필요한 부분이 많았으며 복잡한 장면이나 복잡한 조명 조건을 가진 합성 이미지에서 합리적인 그림자를 생성하기 어려웠다.

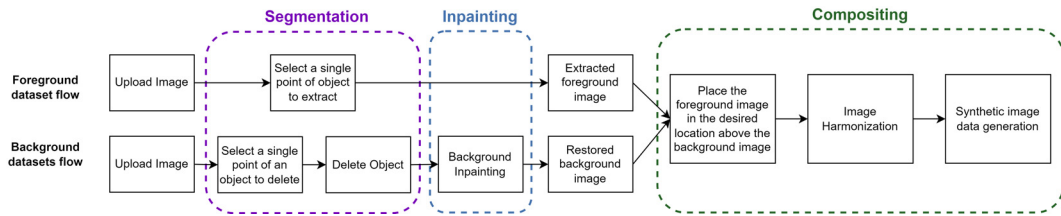
3.4 프로토타입 구축

제안하는 방법의 효율성 및 효과성을 확인하기 위해 프로토타입 구축을 진행하였다. 이를 위해 먼저 합성 이미지 데이터를 생성하기 위한 Segmentation 및 Inpainting, Compositing 과정을 구현해 놓은 기존 데모를 직접 사용해보고 분석하였다. <Table 4>는 해당 데모를 사용하면서 그 기능과 성능 상 장단점을 비교하여 정리한 표이다.

사용자 편의성 관점에서 Segmentation과 Inpainting의 경우 사용자가 원하는 이미지를 바로 업로드하여

<Table 4> Comparing the Strengths and Weaknesses of Existing Demos

Type	Strengths	Weaknesses
Segmentation (Segment Anything Model)	<ul style="list-style-type: none"> - Users can upload desired images - Objects can be extracted and images saved (PNG) at high speed - Various object extraction methods (point, box, automatic, etc.) 	<ul style="list-style-type: none"> - Multiple objects cannot be extracted at once as multiple boxes cannot be extracted.
Inpainting (Inpaint Anything)	<ul style="list-style-type: none"> - Users can upload desired images 	<ul style="list-style-type: none"> - It takes a long time to run (the Hugging Face webpage also takes a long time to load)
Compositing (Image Composition)	<ul style="list-style-type: none"> - Composition possible at high speed - Users can conveniently adjust the size and location of objects 	<ul style="list-style-type: none"> - User cannot upload desired images - Object image rotation (angle adjustment) is not possible - Image saving is not possible



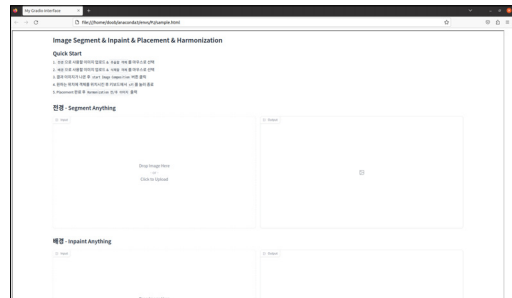
〈Figure 3〉 Web based Prototype Workflow

결과를 확인할 수 있다는 것이 큰 장점으로 판단되었다. 특히 Segmentation은 다양한 객체 추출 방법을 지원하고 빠른 추출 속도로 PNG 형식의 이미지를 저장할 수 있어 사용자에게 편의성을 제공하였다. 반면 Inpainting은 웹페이지 로딩 및 실행 시간이 오래 소요되어 알고리즘 및 서버 성능 개선이 필요함을 확인하였다. Compositing의 경우 처리 속도는 빠른 편이었으나 사용자가 원하는 이미지를 업로드 할 수 없었고 결과 이미지 저장 기능이 구현되지 않아 활용에 한계가 있었다.

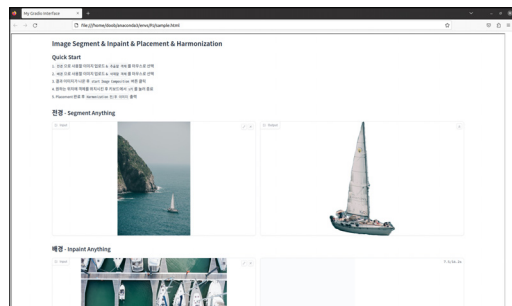
이러한 장단점을 고려하여 필수적인 기능과 편의 기능을 결정하고 본 연구를 통해 개선한 모델의 성능과 더불어 사용자가 세 가지 기능을 하나의 웹페이지에서 편리하게 활용할 수 있도록 프로토타입을 구축하였다.

먼저 사용자가 자신이 원하는 이미지를 업로드 하여 합성 이미지를 생성하고 저장할 수 있도록 하였으며, 객체 선택 시 포인트의 x, y좌표를 코드 내에 직접 작성하는 번거로움을 해소하기 위해 사용자가 마우스를 사용하여 프롬프트를 포인트로 전달하여 객체를 간편하게 지정할 수 있도록 하였다. 이로써 사용자가 편리하게 사용할 수 있도록 합과 동시에 기능 사용의 효율성을 향상시켰다.

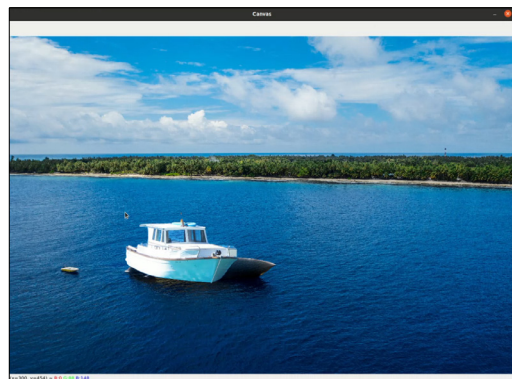
〈Figure 3〉은 프로토타입 웹페이지에서의 합성 이미지 데이터 생성 작업 흐름도이다. 전경 데이터셋의 경우 사용자가 이미지 업로드 후 추출할 객체를 단일 포인트로 선택하면 객체를 제외한 배경은 알파채널의 값을 조정하여 투명 처리된 이미지를 생성한다. 배경 데이터셋은 사용자가 이미지 업로드 후 삭제할 객체를 단일 포인트로 선택한 후 객체를 제거하고, 해당 위치의 배경을 복원한 이미지를 얻을 수 있다. 이후 이러한 과정을 통해 구축한 전경 이미지와 배경 이미지의 합성을 진행한다. 전경 이미지를 배경 이미지 위에 원하는 위치에 위치시킨 후 Image Harmonization을 진행하여 Compositing



(a) Prototype User Interface



(b) Image Segmentation and Inpainting



(c) Image Compositing
〈Figure 4〉 Prototype Webpage

까지 완료된 합성 이미지 데이터를 생성한다. Segmentation, Inpainting, Compositing 각 단계에서 완료된 이미지는 모두 저장할 수 있도록 하였으며, Compositing 과정에서 전경 객체를 위치시킨 후 Harmonization을 진행하기 전과 진행한 후의 결과 이미지를 확인하고 저장할 수 있도록 구현하였다.

(Figure 4)는 구축 완료한 프로토타입 웹사이트의 모습이다. 이 프로토타입은 구현이 용이하고 사용자에게 최적화된 사용자 인터페이스(UI)를 제공할 수 있는 Gradio를 활용하여 구축하였다. 사용자가 별도의 설명 없이도 편리하고 간편하게 웹페이지를 이용할 수 있도록 간단하고 직관적으로 제작하였다. 상단에는 이용 방법 설명을 배치하여 웹페이지를 방문한 사용자들이 작업을 진행하기 전에 사용법을 익힐 수 있도록 하였으며, 이미지를 업로드하고 포인트를 전달할 수 있는 창의 사이즈를 크게 만들어 편리하게 작업할 수 있도록 하였다. 하단에는 각 단계에서 사용한 모델들의 Github 링크를 연결해두어 더 많은 정보가 필요하거나 직접 모델을 사용하고 싶은 경우 바로 이동할 수 있도록 하였다.

3.5 한계점

프로토타입 구축을 통해 사용자가 실제 해상교통 데이터를 획득하는 것에 비해 보다 적은 시간과 비용으로 합성 데이터셋을 생성할 수 있음을 확인 하였다. 그러나 합성데이터셋이 실제 데이터셋에 비해 얼마나 효율적인지에 대한 효용성 측정과 관련하여 시간적으로 얼마나 빠르게 데이터셋을 구축할 수 있는지와 시각적으로 얼마나 사실적이고 구축된 데이터셋을 실제 태스크에 적용했을 때 실제 데이터셋과 유사한 수준의 기여를 할 수 있는지에 대한 추가적인 정량화 연구가 필요함도 확인 하였다. 그리고 유사한 관점에서 합성 데이터셋 생성 절차상 합성 장면 생성 과정이후 객체 배치가 실제로 발생 가능한 수준으로 적절함에 대한 부분을 수동으로 진행해야 하는 문제점과 합성된 전경 이미지의 색상이 얼마나 자연스러운지에 대한 정량적 평가 부분은 추가적인 연구를 통해 보완이 필요함을 확인하였다.

4. 결 론

본 논문에서는 해상교통 상황인지 향상을 위한 합성

데이터셋 구축방안을 제안하였다. 상당수가 인적오류로 인해 발생하는 선박 충돌사고는 예방을 위해 보다 효과적인 상황인지 지원 시스템 개발이 필요한 상황이다. 그리고 인공지능 모델 기반의 상황인지 지원 시스템이 그 해결책이 될 수 있으나 넓고 깊은 바다의 특성상 상황인지 향상을 위한 해상교통 데이터셋의 구축이 매우 어려운 것이 현실이다. 이에 대한 해결책으로 본 논문에서 제안하는 합성 데이터셋 생성 방법은 전경 데이터셋과 배경 데이터셋 개념을 도입하고 배경 데이터셋의 시공간 파라미터를 분석하고 이를 바탕으로 사실적인 합성 장면 생성이 가능하다. 제안하는 방법의 효과를 확인하기 위해 기 연구되어 공개된 Segment Anything Model과 Inpaint Anything, INR-Harmonization 모델을 해상 환경에 맞게 수정하여 활용하였고 사용자가 쉽고 간단하게 사용할 수 있도록 프로토타입을 구축하였다. 각 과정별 결과 데이터 및 수행 절차에 대한 시험을 통해 제안하는 방법을 이용하여 생성한 합성 이미지 데이터가 부족한 해상교통 분야 데이터셋 보완에 활용 가능한 수준으로 생성됨을 확인하였다. 또한 전경 및 배경 데이터셋 구축 이후의 과정은 자동화를 통해 다양한 합성 이미지 데이터를 쉽고 빠르게 생성할 수 있음을 확인하였다. 제안하는 방안을 통해 구축한 합성 데이터셋은 실질적으로 해상교통 상황인지 및 사고 예방 분야에 활용 가능할 것으로 기대되며 개인정보 보호가 용이하고 획득이 어려운 상황에 대한 데이터셋 구축이 가능하여 보다 안정적이고 신뢰할 수 있는 해상 객체 탐지 및 추적 모델 구축이 가능하고 나아가 고품질의 운항 지원 시스템 개발에 기여가 가능할 것으로 기대된다.

향후 연구로 프로토타입 구축 및 실험을 통해 한계점으로 확인된 부분에 대한 보완을 진행할 예정이다. 그리고 절차 개선 측면에서 수차례의 실험을 통해 확인한 배경 데이터셋 이미지를 Inpainting 할 경우 Segmentation 범위를 세분화하여 여러 번 Inpainting하는 것이 보다 효과적이었던 부분 또한 보완할 예정이며, 이미지가 촬영된 위치와 시간, 광원 정보 등과 같은 메타 정보를 모델이 함께 학습하는 방법도 진행할 예정이다. 또한 전경 데이터셋과 배경 데이터셋에 NeRF를 적용하거나 배경 데이터셋의 시간과 기상을 변화시켜 보다 다채로운 데이터셋을 생성하는 고도화 연구도 진행하고자 한다.

References

- [1] Ali, H., Xiong, G., Tianci, Q., Kumar, R., Dong, X., and Shen, Z., "Autonomous ship navigation with an enhanced safety collision avoidance technique", *ISA transactions*, October, 2023.
- [2] Chen, J., Zhang, Y., Zou, Z., Chen, K., and Shi, Z., "Dense Pixel-to-Pixel Harmonization via Continuous Image Representation", *arXiv preprint arXiv:2303.01681*, 2023.
- [3] Er, M. J., Zhang, Y., Chen, J., and Gao, W., "Ship detection with deep learning: a survey", *Artificial Intelligence Review*, Vol. 56, No. 6, March, 2023, pp. 1-41.
- [4] Jeon, M., Park, J., and Woo, J., "Development of HHI's advanced navigation assistance system for safe voyage", *IFAC-PapersOnLine*, Vol. 52, No. 21, September, 2019, pp. 111-113.
- [5] Kim, D., "A Study on the Causes of Human Error in the OOW Considering the TIPOSC", *Korean Association of Maritime Police Science*, Vol. 8, No. 1, February, 2018, pp. 1-18.
- [6] Kim, K., Kong, K., and Kang, S., "A Review on Deep Learning-based Image Outpainting", *Journal of Broadcast Engineering*, Vol. 26, No. 1, January, 2021, pp. 61-69.
- [7] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A., Lo, W., Dollár, P., and Girshick, R., "Segment Anything", *arXiv preprint arXiv:2304.02643*, 2023.
- [8] Kirillov, A., He, K., Girshick, R., Rother, C., and Dollár, P., "Panoptic Segmentation", *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 9396-9405.
- [9] Lee, G. and Ma, J., "A study on improving weapon system detection performance using sim2real learning", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol. 24, No. 7, July, 2023, pp. 189-196.
- [10] MathWorks, "What Is Image Segmentation? 3 things you need to know", [cited 2023 October 10], Available from: <https://www.mathworks.com/discovery/image-segmentation.html>.
- [11] MIT Technology Review, "The era of synthetic data is coming", [cited 2023 October 10], Available from: <https://www.technologyreview.kr/2022-10-24-synthetic-data/>.
- [12] Niu, L., Cong, W., Liu, L., Hong, Y., Zhang, B., Liang, J., and Zhang, L., "Making Images Real Again: A Comprehensive Survey on Deep Image Composition", *arXiv preprint arXiv:2106.14490*, 2021.
- [13] Oh, J., Park, S., and Kwon, O. S., "Advanced navigation aids system based on augmented reality", *International Journal of e-Navigation and Maritime Economy*, Vol. 5, No. 1, December, 2016, pp. 21-31.
- [14] TTA, "semi-structured data", *Telecommunications Technology Association Information and Communication Terminology Dictionary*, [cited 2023 October 6], Available from: http://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=175129-2.
- [15] TTA, "structured data", *Telecommunications Technology Association Information and Communication Terminology Dictionary*, [cited 2023 October 6], Available from: http://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=175128-2.

- [16] TTA, “unstructured data”, Telecommunications Technology Association Information and Communication Terminology Dictionary, [cited 2023 October 6]. Available from: http://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=045801-7.
- [17] Yu, T., Feng, R., Feng, R., Liu, J., Jin, X., Zeng, W., and Chen, Z., “Inpaint Anything: Segment Anything Meets Image Inpainting”, arXiv preprint arXiv:2304.06790, 2023.

■ 저자소개



이 영 채

충남대학교 환경공학과에 학사 과정 재학 중이며 복수전공으로 환경 ICT를 이수 중이다. 주요 관심분야는 객체 인식 및 탐지, 알고리즘이다.



박 세 길

한국해양과학기술원 부설 선박해양플랜트연구소의 책임연구원으로 재직 중이다. 주요 관심분야는 확장현실, 시뮬레이션, 인공지능이다.