

# 오토인코더를 이용한 파랑 비디오 영상에서의 수리동역학적 장면 분리 연구

김태경<sup>01</sup>      김재일<sup>2</sup>      김진아<sup>\*3</sup>

경북대학교<sup>01,2</sup>, 한국해양과학기술원<sup>\*3</sup>

paperrune@knu.ac.kr<sup>01</sup>, jaeilkim@knu.ac.kr<sup>2</sup>, jakim@kiost.ac.kr<sup>\*3</sup>

## Hydrodynamic scene separation from video imagery of ocean wave using autoencoder

Taekyung Kim<sup>01</sup>      Jaeil Kim<sup>2</sup>      Jinah Kim<sup>\*3</sup>

Kyungpook National University<sup>01,2</sup>, Korea Institute of Ocean Science and Technology<sup>\*3</sup>

### 요약

본 논문에서는 대용량 비디오 영상에서 오토인코더를 이용하여 파랑 전파시 수리동역학적 장면만을 분리하는 방법에 대해 소개한다. 연안에서 센서를 이용한 파랑의 직접적 관측의 어려움으로 인해 비디오 영상을 이용한 입자 추적, 옵티컬 플로우 등의 이미지 분석 방법이 주로 활용되고 있다. 하지만 이미지 분석 방법은 주변광 및 기상상태 등 외부 요인에 의한 영향으로 파랑에 대한 정확한 분석에 어려움이 있다. 제안하는 방법은 비디오 영상으로부터 주변광의 영향을 최소화하고, 순수 파랑 전파시 파랑의 움직임만을 분리하여 수리동역학적 장면을 추출한다. 실제 해역 및 수리 모형 실험에서 촬영된 비디오 영상에 제안하는 방법을 적용하여 원본 영상으로부터 주변광에 의한 영향과 배경을 잘 분리하여 파랑 전파에 따른 수리동역학적 파랑 이동 장면이 잘 추출되었음을 시각적으로 확인하였다. 또한 변분 오토인코더의 잠재표현 학습을 통해 얻은 원본 비디오 영상에 대한 잠재 표현은 주변광과 배경 요인에 의해 지배적으로 결정되는 반면, 파랑 이동 장면은 해당 요인에 관계없이 독립적으로 잘 표현되는 것을 알 수 있었다.

### Abstract

In this paper, we propose a hydrodynamic scene separation method for wave propagation from video imagery using autoencoder. In the coastal area, image analysis methods such as particle tracking and optical flow with video imagery are usually applied to measure ocean waves owing to some difficulties of direct wave observation using sensors. However, external factors such as ambient light and weather conditions considerably hamper accurate wave analysis in coastal video imagery. The proposed method extracts hydrodynamic scenes by separating only the wave motions through minimizing the effect of ambient light during wave propagation. We have visually confirmed that the separation of hydrodynamic scenes is reasonably well extracted from the ambient light and backgrounds in the two videos datasets acquired from real beach and wave flume experiments. In addition, the latent representation of the original video imagery obtained through the latent representation learning by the variational autoencoder was dominantly determined by ambient light and backgrounds, while the hydrodynamic scenes of wave propagation independently expressed well regardless of the external factors.

**키워드:** 비지도 학습, 수리동역학적 장면 분리, 오토인코더, 파랑 전파

**Keywords:** Unsupervised learning, Hydrodynamic scene separation, Autoencoder, Wave propagation

## 1. 서론

연안은 바다와 육지가 맞닿아 서로 밀접한 영향을 주고받는 공간으로 산업공간으로의 활용뿐 아니라 관광, 해양레저 등 문화공간

으로 다양한 형태의 이용 집약도가 높은 지역이다. 이와 동시에 연안은 해일 등 연안 재해에 매우 취약한 지역으로 급속한 연안 개발로 자연 복원력이 저하되면서 연안 침식, 방파제 월파, 태풍

\*corresponding author: Jinah Kim/Korea Institute of Ocean Science and Technology(jakim@kiost.ac.kr)

해일 침수범람 등 파랑 기인 연안 재해 발생으로 인한 피해가 급증하고 있다 [1]. 따라서 연안에서의 정확한 파랑 관측과 예측은 매우 중요하나, 연안에서의 파랑은 현상 자체의 높은 비선형성과 센서 기반의 직접 관측에 어려움이 있어 현상에 대한 이해가 충분하지 않다.

이에 연안에서의 연속적인 파랑 모니터링을 통한 충분한 시공간 데이터 확보를 위한 CCTV 비디오 및 레이더 영상 등의 원격탐사 기술이 활발하게 적용되고 있다 [2]. 이를 통해 방대한 시공간 데이터 확보와 유지보수가 용이해졌으며, 획득한 대용량 시계열 광학 영상에서 파랑 현상을 정확히 정량화하고 모델링할 수 있는 분석 기술에 대한 요구가 높아지고 있다. 기존의 연구에서는 Optical Current Meter(OCM) [3], Particle Image Velocimetry(PIV) [4], Optical Flow [5] 등 다양한 영상처리 방법이 적용되어 왔고, 최근에는 컴퓨터 비전 분야의 딥러닝 기술을 이용한 파랑 모델링 융합 연구가 시도되고 있다.

그러나 비디오 영상을 이용한 머신러닝 기반의 파랑 분석 연구는 비디오 영상의 레이블링이 까다롭고 자연상태에서 촬영되는 경우 포함되는 주변광 및 기상상태 등 외부 요인에 의해 영향을 받을 수 있기 때문에 기존 영상처리 방법만으로는 실험역 파랑 영상을 이용한 정확한 분석이 어렵다. 예를 들어, OCM, PIV 방법은 유체 역학 연구에서 미소시간 동안 움직인 입자를 추적하는데 사용되는 이미지 분석 방법이며, Optical Flow 방법은 분석 영상에 대해 고정된 위치의 일정한 광원에서 획득한 영상이어야 한다는 가정을 하고 있어 주변광에 의한 영향이 큰 실험역 영상에서 분석 정확도가 높지 않다. 또한 동적 물체 검출을 위해 사용되는 배경 감산(background subtraction) [6] 은 초기 프레임에서 배경에 대한 사전 정의가 요구되나 파랑 영상에서는 이를 정의하기 어렵다.

따라서 본 논문에서는 오토인코더를 이용한 비지도 학습을 통해 비디오 영상으로부터 일광 및 조명과 같은 주변광의 영향을 분리한 파랑 이동에 따른 수리동역학적 장면 추출 방법을 제안하고, 이를 실제 해역 및 수리모형 실험에서 촬영된 파랑 비디오 영상에 적용하여 파랑 이동 장면 추출 성능을 살펴보고자 한다.

## 2. 오토인코더를 이용한 수리동역학적 장면 분리

본 연구에서 제안하는 비디오 분리 모델은 오토인코더(autoencoder) 형태의 인공신경망을 이용해 비디오 영상을 압축 및 복원하는 과정에서 영상 전체에 걸쳐 변화하지 않는 영역에 해당하는 배경 프레임과 변화하는 영역에 해당하는 전경 영상을 분리하여 생성하고, 두 영상을 결합하기 위한 픽셀단위 가중치를 결정하는 마스크 영상을 생성한다. 이 과정에서 모델은 주변광에 의한 고정된 영향을 배경 프레임으로 분리하고 움직임을 표현하기 위해 전경을 활용하는 정도를 마스크 영상이 표현하도록 한다.

제안하는 비디오 분리 모델은 Figure 1과 같이 인코더(en-

coder), 전경(foreground) 스트림 및 배경(background) 스트림으로 구성된다. 모델은 전경과 배경을 분리하여 자연스러운 비디오 영상을 생성 [7]하는 적대적 생성 신경망(generative adversarial network) [8] 구조에서 판별기(discriminator)를 제거하고 생성기(generator) 앞에 인코더를 추가하여 입력 비디오 영상을 압축 후 복원하는 오토인코더의 형태로 확장하였다. 그림에서 각 실선 화살표는 입력 비디오 영상을 전파하는 컨볼루션 연산, 각 육각형과 사각형은 컨볼루션 연산을 거친 후의 영상의 크기가 각 축으로 절반씩 감소 혹은 2배로 증가하는 변화를 나타낸다. 인코더는 3차원 컨볼루션 연산으로 고차원의 입력 비디오 영상을 저차원의 벡터로 압축하고 두 개의 스트림으로 구성된 디코더(decoder)는 압축된 벡터를 원래의 영상으로 복원하는데, 그 과정에서 전경 영상과 배경 프레임으로의 분리를 수행한다. 전경 스트림은 입력 비디오 영상과 동일한 크기의 전경 영상과 1개의 채널로 구성된 마스크 영상을 생성하고, 배경 스트림은 단일 프레임 크기의 배경 프레임을 생성한다. 이 때, 전경 스트림과 인코더를 스킵커넥션(skip-connection)으로 연결하여 압축이 완전히 진행되지 않은 각 은닉 층의 특징을 활용하여 선명한 영상을 복원할 수 있도록 한다. 반대로, 배경 스트림은 입력 영상 전체 프레임에 대한 공통된 배경 정보를 생성할 수 있도록 압축이 완료된 벡터로부터 생성한다. 생성된 전경 및 마스크 영상과 배경 프레임을 사용하여 최종 영상을 합성하는 방법은 다음과 같이 표현할 수 있다:

$$H(x) = m \odot f + (1 - m) \odot b \quad (1)$$

여기서  $f$ 는 전경 영상,  $m$ 는 마스크 영상,  $b$ 는 배경 프레임, 1은 모든 원소가 1인  $m$ 와 모양이 같은 3차원 행렬이고,  $\odot$ 은 원소별 곱셈을 의미한다. 출력 영상을 합성할 때 각 항에서 상대적으로 크기가 작은 축 즉, 마스크 영상의 채널 축 또는 배경 프레임의 시간 축을 복제하여 큰 영상과 모양을 동일하게 한다. 마스크 영상에 대해  $[0, 1]$ 의 범위를 가지는 시그모이드 함수를 활성화 함수(activation function)로 사용하여 합성된 출력 영상의 픽셀 값 범위를 전경 및 배경 영상과 동일하게 한다. 또한 마스크 영상의 L1 희소성(sparsity)을 손실 함수에 추가함으로써 모델이 전경 영상을 활용하는 것에 패널티를 부여하여 배경 프레임이 비디오 영상의 최대한 많은 정보를 포함하도록 한다. 마스크 영상에 대한 L1 희소성은 다음과 같이 각 원소의 절대값 평균을 손실 함수에 추가하여 부여할 수 있다:

$$\|m\|_1 = \frac{1}{THW} \sum_k \sum_j \sum_i |m_{ijk}| \quad (2)$$

여기서  $T$ 는 프레임 개수,  $H$ 와  $W$ 는 각각 영상의 세로와 가로 해상도이다. 따라서 비디오 분리 모델을 훈련하기 위한 손실 함수(loss function)는 다음과 같다:

$$\mathcal{L}_H = \|x - H(x)\|_2 + \lambda \|m\|_1 \quad (3)$$

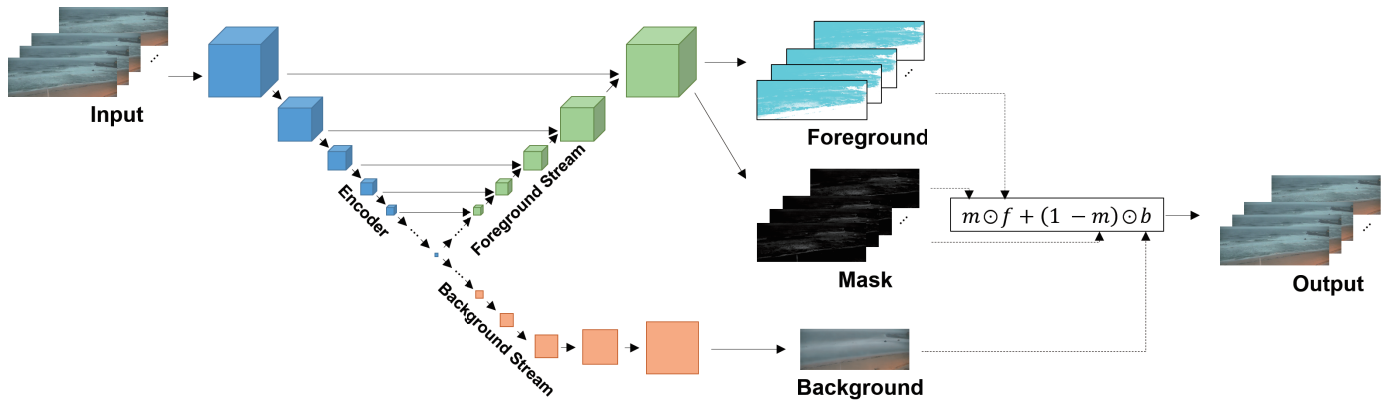


Figure 1: Model for hydrodynamic scene separation from video imagery of wave propagation

여기서 왼쪽 항은 입력 비디오 영상과 생성된 출력 비디오 영상 사이의 평균제곱오차(mean-squared-error) 이고  $\lambda$ 는 임의의 상수이다.  $\lambda$ 가 적절한 값으로 설정되었을 때 모델은 낮은 오차로 입력 비디오 영상을 복원하면서 동시에 전경과 배경의 분리를 달성할 수 있다. 전경 영상을 활용할 때 활성화되는 마스크에 대한 희소성을 달성하기 위해 모델은 배경 프레임만으로 입력 비디오 영상의 움직이지 않는 영역을 복원하고 배경 프레임으로 표현할 수 없는 움직이는 영역에만 전경 영상을 활용하고자 노력한다. 배경 프레임이 입력 영상 전체 프레임에 대해 변하지 않는 공통된 영상을 정확히 포착한 경우 마스크는 오직 움직이는 영역을 표현하는 영상이 된다. 따라서 제안하는 모델은 다음과 같은 특징을 갖는다:

- 모델이 생성하는 마스크 영상은 1개의 채널로 구성되어 흑백 영상으로 표현된다. 각 RGB 색상에 대응되는 흑백 픽셀 값을 계산하는 대신 마스크의 각 원소는 전경 영상을 활용하는 정도를 나타내므로 주변광에 의한 영향을 받지 않는다.
- 손실 함수에 마스크의 희소성을 추가하여 배경 프레임 영역을 활용하는 영역에 대한 마스크 영상의 픽셀 값은 0에 수렴하여 배경 감산의 효과를 얻을 수 있다.
- 마스크 영상은 입력 비디오 영상을 압축 후 그대로 복원하도록 훈련하는 과정에서 생성되므로 비디오 영상에 대한 레이 블링이 없이도 파랑 추출에 활용할 수 있다.

해수면 운동의 물리량을 관측하기 위해 파랑에 영상처리 방법을 적용할 때, 주변광과 고정된 배경 영상은 불필요한 정보일 뿐만 아니라 동일한 형태의 파랑을 카메라에서 다양한 모습으로 비치게 만드는 요인이므로 이러한 정보를 분리하는 제안하는 방법은 매우 효과적인 수리동역학적 장면 추출 기법이 될 수 있다.

### 3. 모델 구현 및 실험

#### 3.1 데이터 집합

**수리모형실험 비디오 영상** 2019년 1월 전남대학교 해양항만실험센터에서 수행된 파랑 수리모형실험을 CCTV로 촬영한 비디오 영상을 사용하여 제안하는 방법을 적용·평가했다. 실험은 5가지 폭풍 시나리오에 따라 파랑을 조파하는 과정을 반복 수행하였으며, 동일한 시나리오를 재현한 2개의 독립적인 실험의 비디오 영상을 가져와 분석하였다. 두 실험은 동일한 파랑을 생성하였으나 2시간의 시간 차이를 두고 수행되어 비디오 영상에서는 태양의 위치 변화에 따른 전반적인 픽셀 값의 차이가 관측된다. 제안하는 방법을 적용하기에 앞서 비디오 영상이 포함하는 수로외의 불필요한 영역을 제거하고자 원근투영변환(perspective projective transformation) 기법을 적용하여 1080 x 1920 x 3 크기의 영상으로부터 1688 x 200 x 3 크기의 수로 영상을 얻었다 (Figure 2). 그리고 비디오 영상의 모든 프레임을 추출하여 각 실험으로부터 26,992개의 프레임 영상을 얻었으며 추출한 프레임 영상을 16개씩 묶어 얻은 총 53,954개의 16 프레임 길이의 짧은 비디오 영상을 훈련 데이터로 사용했다. 손실 함수에서 마스크에 희소성을 부여하는 상수  $\lambda$ 는 0.01로 설정했다. 상수가 너무 높게 설정된 경우에는 전경을 적절히 활용하지 못해 복원된 영상에서 파랑이 정상적으로 생성되지 않았고, 너무 낮게 설정된 경우에는 고정된 영역에도 전경을 활용하여 배경의 분리를 달성하지 못했다.

**안목 해변 비디오 영상** 제안하는 방법이 실제 해변에서도 효과적으로 수리동역학적 장면을 분리할 수 있는지 검증하기 위해 실제 안목의 바다를 촬영한 비디오 영상을 이용하였다 (Figure 5 좌측열 입력 영상). 2015년 11월 19일 바람이 강하게 불어 파도가 뚜렷하게 드러나는 날 하루 (5시 ~ 18시) 를 촬영한 영상을 사용하였다. 해당 비디오 영상은 일출부터 일몰까지 태양의 위치가 픽셀 값에 미치는 영향을 포함한다. 불필요한 연안유역 부분을 잘라내어 720 x 1920 x 3 크기의 영상을 얻었으며 GPU의 메모리



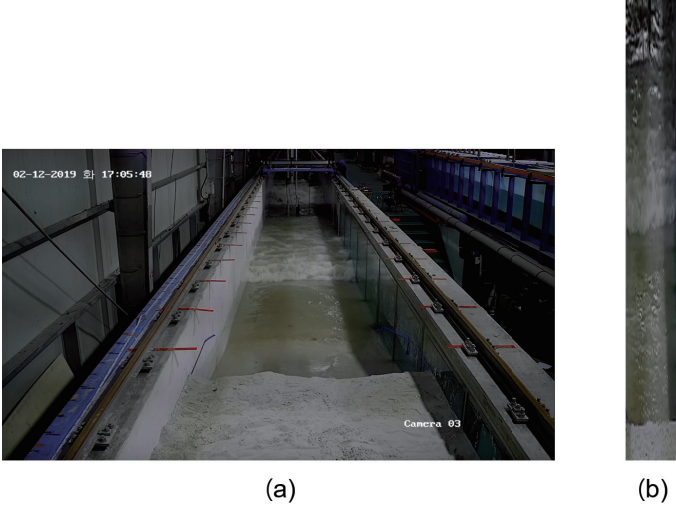


Figure 2: (a) Sample image from the wave flume experiment and (b) Rectified image by applying perspective projective transformation to (a)

크기를 고려해 영상을 세로 및 가로 해상도를 1/3씩 줄여 최종적으로  $240 \times 640 \times 3$  크기의 비디오 영상을 얻었다. 비디오 영상으로부터 30 프레임마다 1장씩 프레임을 추출하여 총 43,883개의 프레임 영상을 얻었으며 추출한 프레임 영상을 16개씩 묶어 얻은 총 43,868개의 16 프레임 길이의 짧은 비디오 영상을 훈련 데이터로 사용했다. 손실 함수에서 마스크에 희소성을 부여하는  $\lambda$ 는 0.05로 설정했다.

### 3.2 구현 세부사항

제안하는 비디오 분리 모델은 비디오 영상으로부터 특징을 추출 및 전파하기 위해 인코더 및 인코더와 전경 스트림을 연결하는 스킵커넥션에 3차원 컨볼루션을 사용한다. 디코더에는 전치 컨볼루션(transposed convolution)을 사용하여 영상의 크기를 늘려진 전경 스트림에는 3차원, 배경 스트림에는 2차원 전치 컨볼루션을 사용한다. 채널의 개수는 영상의 크기에 반비례하게 결정하여 인코더 부분에서는 최초 16개부터 영상을 압축할 때마다 2배씩 늘려 최대 512개까지 늘렸다가 디코더 부분에서는 반대로 영상을 복원할 때마다 2배씩 줄여 16개로 줄였다. 컨볼루션 층의 개수와 각 층의 크기 및 채널 개수는 GPU 메모리 용량 또는 입력 비디오 영상의 크기에 따라 유연하게 결정할 수 있으며 생성 영상의 품질에 큰 영향을 미치지 않는다. 모델 훈련을 위해 가중치를 Xavier initialization [9] 방법을 사용하여 초기화했고 아담 옵티마이저(ADAM optimizer) [10]를 사용하여 학습률(learning rate)은 0.001로 설정했다. 모든 실험에서 배치의 크기는 2, 프레임 개수는 16개로 진행하였으며 프레임 개수는 본 실험에서 사용한 GPU (NVIDIA TITAN XP 12GB)의 메모리 크기를 고려하여 결정했다.



Figure 3: Hydrodynamic scene separation of foreground, mask, and background in wave flume experiment video according to the proposed model in Figure 1

### 3.3 변분 오토인코더를 이용한 잠재 표현 학습

제안하는 방법의 배경 분리 및 수리동역학적 장면 추출 성능을 검증하기 위해 변분 오토인코더(Variational AutoEncoder, VAE)를 이용한 원본 비디오 영상과 추출된 수리동역학적 파랑 이동 장면에 대한 잠재 표현 실험을 진행하였다. VAE [11]는 입력 영상을 압축 후 복원하는 오토인코더와 동일한 목적을 달성하는 동시에 압축된 잠재 분포가 표준 정규 분포를 따르도록 정규화한다. VAE의 인코더는 입력 영상을 잠재 공간의 특정 분포에 매핑하고, 디코더는 잠재 공간의 벡터를 영상으로 복원한다. 다시말해 VAE는 정규 분포를 따르는 잠재 표현을 학습하는 차원 축소(dimension reduction) 기법으로 볼 수 있으며, 여러 데이터의 잠재 표현을 시각화 하였을 때 데이터의 분포와 데이터 간 유사성을 판단할 수 있다.

본 실험에서 제안하는 방법이 연안 비디오로부터 배경 및 환경 요인을 정확히 분리한 파랑 영상을 생성하는 경우, 파랑 영상의 분포는 더 이상 주변광 등 외부 요인에 영향을 받지 않을 것이라고 가정하였다. 흐린 날과 맑은 날처럼 서로 다른 환경에서 촬영되더라도, 영상에서 파랑의 움직임과 크기가 유사하다면, 두 비디오 영상에서 유사한 파랑 마스크가 추출되고, 서로 매우 유사한 파랑 특징에 따라서 VAE의 인코더가 두 마스크 영상을 잠재 공간에서 거의 동일한 위치에 매핑할 것이다.

잠재 공간의 시각화를 위하여, VAE의 잠재 표현을 2차원으로

구성하였다. 수리모형실험 비디오 영상은 동일 파랑 발생 시나리오로 서로 다른 시간에 촬영된 비디오 영상을 잠재 표현 학습 실험에 사용하였고, 촬영 시간을 기준으로 그룹화하여 시각화하였다. 안목 해변 비디오 영상은 1시간 단위로 구분하여 시각화하였는데, 주변광은 날씨 뿐 아니라 시간에 따라서도 연속적으로 변하기 때문에 본 실험에 적합한 변수이다. 시간에 따른 연안 비디오 영상의 주변광 변화는 Figure 5에서 확인할 수 있다.

## 4. 결과 및 성능 평가

Figure 3은 수리모형실험 영상에 대해 훈련된 비디오 분리 모델이 생성하는 영상으로, 생성된 16프레임 중 첫 프레임을 보여준다. 모델은 낮은 복원 오차 ( $< 0.001$ )로 입력 비디오 영상을 복원하였으며 그 과정에서 전경과 배경 영상 및 마스크로의 분리를 달성했다. 생성된 배경 영상은 파랑은 생성하지 않지만 입력 비디오 영상의 전반적인 색상을 충실히 복원했다. 변화하는 연속 비디오 프레임에 대해 동일한 배경을 사용하면서, 높은 복원도를 갖는 것은 분리된 마스크와 전경이 비디오 프레임의 변화를 충분히 표현하고 있음을 의미한다. 다만, 전경 영상은 전반적으로 매우 높은 픽셀 값을 가지는 경향이 있는데 곱해지는 마스크에 희소성을 부여하였기 때문에 영상 복원 시 전경의 가중치를 의미하는 마스크의 화소값이 작아져서 영상 복원 시 이를 극복하기 위함으로 판단된다. 마스크에 대한 희소성은 손실 함수 중  $\lambda$  파라미터를 사용해 조절할 수 있는데,  $\lambda$ 를 작게 설정할 경우, 파랑 외에 조명의 변화 등 다른 변화도 포함하게 되며, 전경의 화소값은 작아진다.

Figure 4는 원근투영변환 기법만을 적용한 원본 비디오 영상과 그로부터 생성한 마스크 영상을 사용하여 훈련된 VAE에서 얻은 잠재 표현의 시각화 결과이다. Figure 4 도표의 축은 VAE를 통해 학습한 2차원 잠재 벡터의 각 요소에 해당한다. 원본 비디오 영상의 잠재 표현을 촬영 시간에 따라 그룹화하여 시각화하였을 때 (Figure 4의 (a)), 영상 분포가 촬영 시간에 의해 명확히 구분되는 분포를 보였으며, 이는 시간에 따른 일광의 변화가 비디오 영상의 잠재 표현에 지배적인 특징으로 인코딩되었음을 의미한다. 그러나 제안한 기법을 이용해 생성한 마스크의 잠재 표현 시각화에서는 인코더가 두 영상 간의 차이를 포착하지 못해 잠재 표현의 분포가 서로 겹쳐진 것을 확인하였다(Figure 4의 (b)). 이는 제안하는 방법이 동일 파랑 시나리오에 따라 촬영된 비디오 영상으로부터 일광에 의한 영향을 효과적으로 분리하여, 동일한 파랑 움직임을 제대로 표현하였기 때문이다.

Figure 5는 안목 비디오 영상에 대한 비디오 분리 모델의 결과를 보여준다. 배경 영상은 모든 프레임에 동일하게 적용됨으로 시각화를 위하여 동일 배경 영상을 프레임마다 배치하였다. 모델은 낮은 복원 오차 ( $< 0.001$ )로 입력 비디오 영상을 복원하였으며 그 과정에서 전경과 배경 영상 및 마스크의 분리를 달성했다. 제안하는 방법은 태양의 위치가 전반적인 픽셀 값에 미치는 영향을 배경 스트림으로 전달함으로써 시간에 따라 명확히 달라지는

일광의 영향을 분리하고 프레임 간에 변하지 않는 영역을 잘 포착한 배경 영상을 생성하여 시간에 관계없이 움직이는 파랑을 일관되게 검출하는 마스크를 생성했다.

Figure 6 (a)는 안목 해안 비디오 영상으로부터 얻은 16 프레임 길이의 비디오 영상 43,868개를 VAE가 잠재 공간에 매핑한 모습으로 각 점은 비디오 영상 하나를 매핑한 분포의 평균에 대응된다. Figure 6 (b)는 분리된 수리동역학적 파랑 이동 장면에 대한 마스크 영상의 잠재표현 분포이며, Figure 6 (c)의 컬러바는 각 색이 어떤 시간대 (5시 ~ 18시)를 의미하는지 나타내는 색상표이다. 생성된 마스크 영상의 잠재 표현은 시간이 아닌 파도 특성에 따라 결정됨으로써 낮 시간에 해당하는 마스크 영상은 원영상과 달리 시간에 따라 분류되지 않았다. 다만, 낮과 밤으로 분류되는 것은 밤에 주변광의 하락으로 파랑이 충분히 관찰되지 않았기 때문으로 판단된다.

## 5. 결론 및 향후 연구

본 논문에서는 연안에서의 촬영된 대용량 비디오 영상에서 파랑 이동에 대한 수리동역학적 장면 분리 방법을 제시하였다. 이는 비지도 학습을 통해 비디오 영상으로부터 주변광의 영향을 최소화하고, 순수 파랑의 이동에 따른 움직임만을 분리하여 수리동역학적 장면을 효과적으로 추출하였다. 또한 잠재 표현 학습을 이용한 실험에서는 원본 비디오 영상과 달리 파랑 이동 장면이 주변광과 배경 정보에 관계없이 독립적으로 표현될 수 있음을 확인하였다. 향후 이를 이용하여 분리된 수리동역학적 파랑 이동 장면을 이용하여 파랑 추적을 통한 파랑의 이동 속도 측정 기법을 개발할 예정이다.

## 감사의 글

본 연구는 한국해양과학기술원 (PE99742, ‘AI기반 파랑기인 연안재해 모델링 플랫폼 및 해무 예측기술 개발’) 연구비 지원으로 수행하였습니다.

## References

- [1] S. Caires, J. Kim, and J. Groeneweg, “Korean east coast wave predictions by means of ensemble kalman filter data assimilation,” *Ocean Dynamics*, vol. 68, no. 11, pp. 1571–1592, 2018.
- [2] R. Holman and M. C. Haller, “Remote sensing of the nearshore,” *Annual Review of Marine Science*, vol. 5, pp. 95–113, 2013.
- [3] C. Chickadel, R. A. Holman, and M. H. Freilich, “An optical technique for the measurement of longshore currents,”



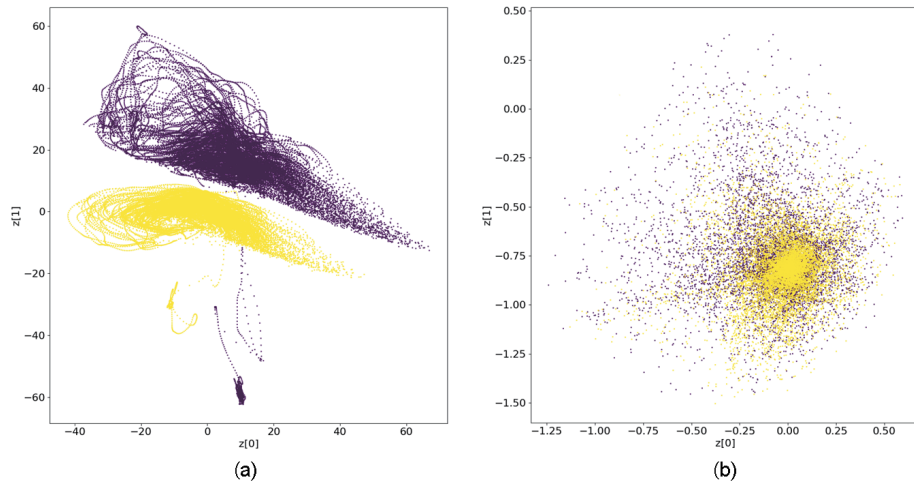


Figure 4: Visualization of latent representation of (a)original video imagery and (b)separated hydrodynamics scenes in wave flume experiment video

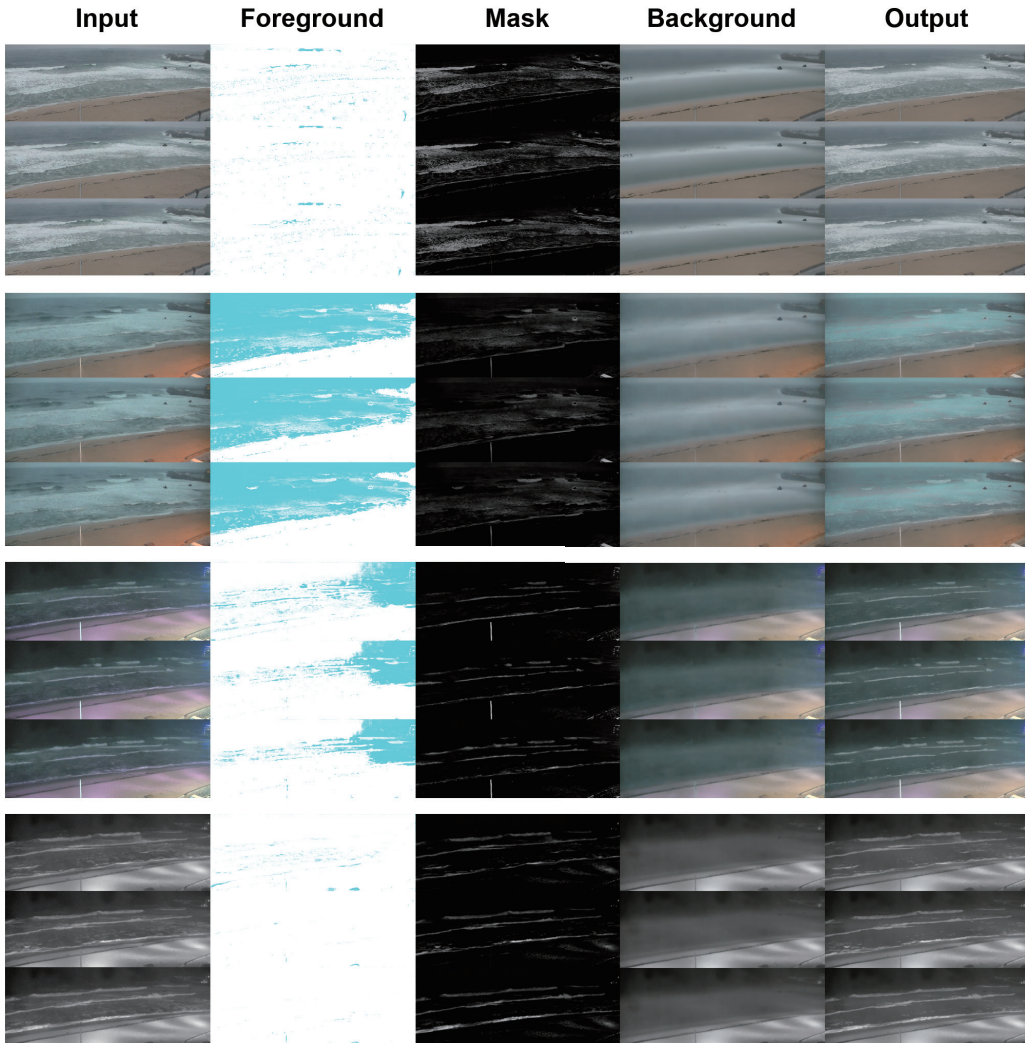


Figure 5: Hydrodynamic scene separation of foreground, mask, and background in Anmok beach video according to the proposed model in Figure 1

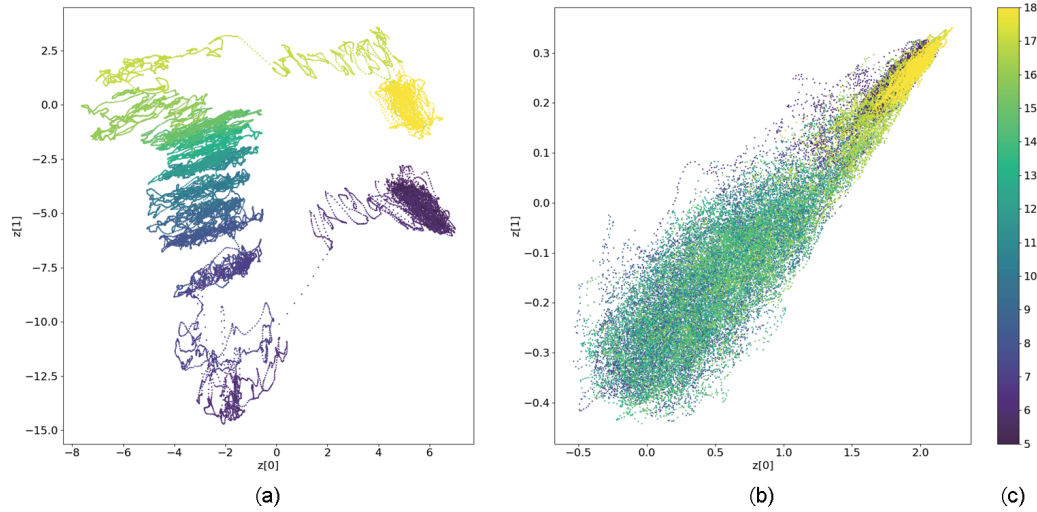


Figure 6: Visualization of latent representation of (a)original video imagery and (b)separated hydrodynamics scenes in Anmok beach video

- Journal of Geophysical Research: Oceans*, vol. 108, no. C11, 2003.
- [4] K. Holland, J. Puleo, and T. Kooney, “Quantification of swash flows using video-based particle image velocimetry,” *Coastal Engineering*, vol. 44, no. 2, pp. 65–77, 2001.
- [5] B. K. Horn and B. G. Schunck, “Determining optical flow,” *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.
- [6] A. Thangarajah, Q. M. J. Wu, W. Jiang, A. Safaei, and J. Huo, “New trend in video foreground detection using deep learning,” *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 889–892, 2018.
- [7] C. Vondrick, H. Pirsiavash, and A. Torralba, “Generating videos with scene dynamics,” in *Advances In Neural Information Processing Systems*, 2016, pp. 613–621.
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [9] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [10] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *International Conference on Learning Representations*, 12 2014.
- [11] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *International Conference on Learning Representations*, 12 2013.

## 〈 저 자 소 개 〉



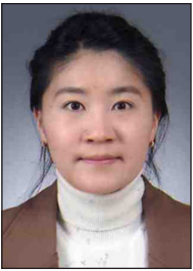
김 태 경

- 2019년 ~ 현재 경북대학교 인공지능학과 석사과정
- 관심분야: 딥러닝, 인공지능, 자율 자동차, 시계열 데이터 예측 등
- <https://orcid.org/0000-0002-9636-9560>



김 재 일

- 2014년 ~ 2016년 삼성전자 책임연구원
- 2016년 ~ 2018년 University of North Carolina at Chapel Hill 박사후연구원
- 2018년 ~ 현재 경북대학교 IT대학 컴퓨터학부 조교수
- 관심분야: 영상 처리, 시계열 데이터 모델링, 예외 검출, 기계학습 등
- <https://orcid.org/0000-0002-9799-1773>



김 진 아

- 2015년 한국과학기술원 전산학과 박사
- 2005년 ~ 현재 한국해양과학기술원 연안재해재난연구센터 책임연구원
- 관심분야: 해양 및 기상자료의 데이터 사이언스(Data Science)
- <https://orcid.org/0000-0002-8110-6047>