



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년11월29일

(11) 등록번호 10-2607905

(24) 등록일자 2023년11월24일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2023.01) G06N 3/04 (2023.01)

(52) CPC특허분류  
G06N 3/08 (2023.01)  
G06N 3/04 (2023.01)

(21) 출원번호 10-2021-0034638

(22) 출원일자 2021년03월17일

심사청구일자 2021년03월17일

(65) 공개번호 10-2022-0129828

(43) 공개일자 2022년09월26일

(56) 선행기술조사문헌

S. Mo 등. "InstaGAN: Instance-aware Image-to-Image Translation".  
arXiv:1812.10889v2\*

H. Maeda 등. "Generative adversarial network for road damage detection". Computer-Aided Civil and Infrastructure Engineering. Vol. 36, Issue 1\*

X. Ouyang 등. "Pedestrian-Synthesis-GAN: Generating Pedestrian Data in Real Scene and Beyond". arXiv:1804.02047v2\*

X. Mao 등. "Least Squares Generative Adversarial Networks". arXiv:1611.04076v3\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

김은태

서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)

이상용

서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)

이수현

서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)

(74) 대리인

특허법인우인

전체 청구항 수 : 총 5 항

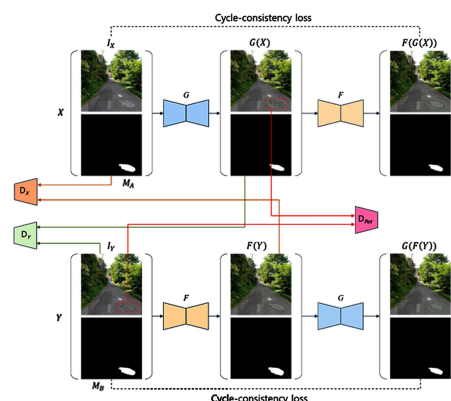
심사관 : 성민용

(54) 발명의 명칭 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치

## (57) 요약

본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치는, 적대적 생성 신경망(generative adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 기반으로 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 소스 단일 영상에 대한 포트홀 영상을 생성함으로써, 다양한 포트홀 영상을 손쉽게 적은 비용으로 생성할 수 있다.

## 대표도



이 발명을 지원한 국가연구개발사업

과제고유번호	1711110608
과제번호	2019R1A2C1007153
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	자율주행차량에서 사고 방지를 위한 강인한 객체 검출 및 추적 시스템 개발(2/3)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.03.01 ~ 2021.02.28
공지예외적용	: 있음

---

## 명세서

### 청구범위

#### 청구항 1

프로세서가, 포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하는 단계;

상기 프로세서가, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하는 단계; 및

상기 프로세서가, 상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경 부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하는 단계를 포함하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 생성 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 포트홀 영상을 생성하는 단계는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하는 단계;

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하는 단계; 및

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는 단계를 포함하는,

포트홀 영상 생성 방법.

## 청구항 2

삭제

## 청구항 3

삭제

## 청구항 4

삭제

## 청구항 5

삭제

## 청구항 6

삭제

## 청구항 7

제1항에서,

상기 제2 손실 함수는,

$$L_{advPot} = \mathbb{E}_{I_{YPot} \sim P_{data}(I_{YPot})} [D_{Pot}(I_{YPot}) - 1^2] + \mathbb{E}_{I_{XPot} \sim P_{data}(I_{XPot})} [D_{Pot}(G(I_{XPot})) - 1^2] \text{이며,}$$

상기  $D_{Pot}$ 은, 상기 제3 판별자를 나타내고,

상기  $I_{YPot}$ 은, 상기 제1 단일 영상에 포함된 포트홀 영역 영상을 나타내며,

상기  $I_{XPot}$ 은 상기 제2 단일 영상에 포함된 포트홀 영역 영상을 나타내고,

상기  $G(I_{XPot})$ 는, 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 나타내는,

포트홀 영상 생성 방법.

## 청구항 8

제1항 및 제7항 중 어느 한 항에 기재된 포트홀 영상 생성 방법을 컴퓨터에서 실행시키기 위하여 컴퓨터 판독 가능한 저장 매체에 저장된 컴퓨터 프로그램.

## 청구항 9

포트홀 영상 생성 네트워크를 이용하여 포트홀 영상을 생성하는 포트홀 영상 생성 장치로서,

상기 포트홀 영상 생성 네트워크를 이용하여 상기 포트홀 영상을 생성하기 위한 하나 이상의 프로그램을 저장하는 메모리; 및

상기 메모리에 저장된 상기 하나 이상의 프로그램에 따라 상기 포트홀 영상을 생성하기 위한 동작을 수행하는 하나 이상의 프로세서;

를 포함하고,

상기 프로세서는,

포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으

로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하고,

상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하고,

상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경 부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 생성 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 프로세서는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하고,

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하고,

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는,

포트홀 영상 생성 장치.

## 청구항 10

삭제

## 청구항 11

삭제

## 청구항 12

삭제

## 청구항 13

제9항에서,

상기 제2 손실 함수는,

$$\text{식 } L_{adv_{Pot}} = \mathbb{E}_{I_{Y_{Pot}} \sim P_{data}(I_{Y_{Pot}})} [D_{Pot}(I_{Y_{Pot}}) - 1]^2 + \mathbb{E}_{I_{X_{Pot}} \sim P_{data}(I_{X_{Pot}})} [D_{Pot}(G(I_{X_{Pot}})) - 1]^2 \text{이며,}$$

상기  $D_{Pot}$ 은, 상기 제3 판별자를 나타내고,

상기  $I_{Y_{Pot}}$ 은, 상기 제1 단일 영상에 포함된 포트홀 영역 영상을 나타내며,

상기  $I_{X_{Pot}}$ 은 상기 제2 단일 영상에 포함된 포트홀 영역 영상을 나타내고,

상기  $G(I_{X_{Pot}})$ 는, 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 나타내는,

포트홀 영상 생성 장치.

## 발명의 설명

## 기술 분야

[0001] 본 발명은 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치에 관한 것으로, 더욱 상세하게는 소스 단일 영상에 대한 포트홀(pot hole) 영상을 생성하는, 방법 및 장치에 관한 것이다.

## 배경 기술

[0002] 머신러닝 학습에는 많은 양의 품질 좋은 데이터를 요구한다. 고양이, 강아지, 사람, 차량 등 생활 속에서 쉽게 접할 수 있는 물체의 경우 머신러닝 학습에 사용할 수 있는 공개된 데이터가 충분하지만, 포트홀, 수질 오염, 용암 등의 특수 환경에서의 데이터의 경우 공개된 양질의 데이터는 턱없이 부족하고 직접 취득하기도 어렵다.

[0003] 이렇게 흔하지 않은 객체에 대한 데이터를 직접 취득하지 않고, 인터넷 등을 이용하여 수집한 소량의 실제 데이터를 컴퓨터 영상 처리 방법으로 증량시키는 연구가 활발히 진행되고 있다. 예컨대, 기존에 존재하던 소량의 객체가 담긴 단일 영상을 회전, 밝기 조절, 크기 변환, 위치 이동 등을 이용하여 단순히 단일 영상을 증량할 수 있지만, 이 방법은 어디까지나 실제 단일 영상 데이터 기반의 증량 방법이기 때문에 새롭거나 다양한 데이터를 생성하기에는 한계가 있다.

[0004] 이러한 한계를 해결할 수 있도록 실제 단일 영상 데이터를 기반으로 새로운 형태의 합성 단일 영상을 생성하는 적대적 생성 신경망(Generative Adversarial Network, GAN)이 개발되었다. 적대적 생성 신경망은 생성자라고 불리는 장치가 불규칙 잡음을 입력 받아 새로운 데이터를 생성하면, 판별자라고 불리는 장치는 실제 훈련 데이터와 생성자가 만들어 낸 데이터를 각각 입력 받아 해당 데이터가 실제 데이터인지 혹은 생성자가 만들어 낸 가짜 데이터인지 판별하여 판별 정보를 생성자에게 전달하고, 생성자는 전달받은 정보를 바탕으로 더욱 진짜 같은 데이터를 만들어내려고 학습한다. 이 장치는 불규칙 잡음을 입력 받아 실제 데이터의 분포에 가까운 영상 데이터를 생성하기 위해 학습되지만, 생성되는 데이터에 위치 정보가 반영되지 않아 검출기를 학습시키기 위한 데이터를 생성하는 데는 무리가 있다.

[0005] 단일 영상 간 도메인 변환은 단일 영상의 특정한 도메인을 다른 도메인으로 매핑하여 변환하는 문제를 말한다. 도메인은 단일 영상이 변환할 때 적용되는 특정한 특징을 말하며, 이때 소스 원본 단일 영상으로부터 타겟 단일

영상으로 변환되는 과정은 항상 같은 값을 출력하는 한계가 있다.

## 발명의 내용

### 해결하려는 과제

- [0006] 본 발명이 이루고자 하는 목적은, 적대적 생성 신경망(generative adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 기반으로 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 소스 단일 영상에 대한 포트홀 영상을 생성하는, 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치를 제공하는 데 있다.
- [0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

### 과제의 해결 수단

- [0008] 상기의 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법은, 적대적 생성 신경망(generative adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 학습하는 단계; 및 상기 포트홀 영상 생성 네트워크를 이용하여 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 상기 소스 단일 영상에 대한 포트홀 영상을 생성하는 단계;를 포함한다.
- [0009] 여기서, 상기 포트홀 영상 생성 네트워크 학습 단계는, 포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하는 단계; 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 기반으로, 상기 포트홀 영상 생성 네트워크를 학습하는 단계;를 포함할 수 있다.
- [0010] 여기서, 상기 포트홀 영상 생성 네트워크는, 제1 생성자, 제2 생성자, 제1 판별자, 제2 판별자 및 제3 판별자를 포함하며, 상기 제1 생성자는, 입력된 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하고, 상기 제2 생성자는, 입력된 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하며, 상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득하고, 상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제1 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득하며, 상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 기반으로 판별 정보를 획득할 수 있다.
- [0011] 여기서, 상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상일 수 있다.
- [0012] 여기서, 상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역일 수 있다.
- [0013] 여기서, 상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 영상 및 마스크와 상기 제1 생성자 및 상기 제2 생성자를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 제3 손실 함수, 영상 및 마스크와 상기 제1 생성자 및 상기 제2 생성자 중 하나를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 제4 손실 함수, 및 영상의 배경은 유지하고 포트홀 객체 부분만 변환되도록 하는 제5 손실 함수를 포함할 수 있다.
- [0014] 여기서, 상기 제2 손실 함수는, 식 
$$L_{adv} = E_{I_{Y_{pot}} \sim P_{data}(I_{Y_{pot}})} [D_{Pot}(I_{Y_{pot}}) - 1]^2 + E_{I_{X_{pot}} \sim P_{data}(I_{X_{pot}})} [(D(G(I_{X_{pot}}))) - 1]^2$$
이며, 상기  $D_{Pot}$ 은, 상기 제3 판별자를 나타내고, 상기  $I_{Y_{pot}}$ 은, 상기 제1 단일 영상에 포함된 포트홀 영역 영상을 나타내며, 상기  $G(I_{X_{pot}})$ 는, 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 나타낼 수 있다.
- [0016] 상기의 기술적 과제를 달성하기 위한 본 발명의 바람직한 실시예에 따른 컴퓨터 프로그램은 컴퓨터 판독 가능한

저장 매체에 저장되어 상기한 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 중 어느 하나를 컴퓨터에서 실행시킨다.

[0018] 상기의 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 장치는, 적대적 생성 신경망(generative adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 이용하여 배경에 문맥적으로 적합한 포트홀(pot hole) 영상을 생성하는 포트홀 영상 생성 장치로서, 상기 포트홀 영상 생성 네트워크를 이용하여 배경에 문맥적으로 적합한 상기 포트홀 영상을 생성하기 위한 하나 이상의 프로그램을 저장하는 메모리; 및 상기 메모리에 저장된 상기 하나 이상의 프로그램에 따라 배경에 문맥적으로 적합한 상기 포트홀 영상을 생성하기 위한 동작을 수행하는 하나 이상의 프로세서;를 포함하고, 상기 프로세서는, 상기 포트홀 영상 생성 네트워크를 학습하고, 상기 포트홀 영상 생성 네트워크를 이용하여 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 상기 소스 단일 영상에 대한 포트홀 영상을 생성한다.

[0019] 여기서, 상기 프로세서는, 포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하고, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 기반으로, 상기 포트홀 영상 생성 네트워크를 학습할 수 있다.

[0020] 여기서, 상기 포트홀 영상 생성 네트워크는, 제1 생성자, 제2 생성자, 제1 판별자, 제2 판별자 및 제3 판별자를 포함하며, 상기 제1 생성자는, 입력된 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하고, 상기 제2 생성자는, 입력된 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하며, 상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득하고, 상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제1 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득하며, 상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 기반으로 판별 정보를 획득할 수 있다.

[0021] 여기서, 상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 영상 및 마스크와 상기 제1 생성자 및 상기 제2 생성자를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 제3 손실 함수, 영상 및 마스크와 상기 제1 생성자 및 상기 제2 생성자 중 하나를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 제4 손실 함수, 및 영상의 배경은 유지하고 포트홀 객체 부분만 변환되도록 하는 제5 손실 함수를 포함할 수 있다.

[0022] 여기서, 상기 제2 손실 함수는, 식 
$$L_{adv_{Pot}} = E_{I_{Y_{Pot}} \sim P_{data}(I_{Y_{Pot}})} [D_{Pot}(I_{Y_{Pot}}) - 1]^2 + E_{I_{X_{Pot}} \sim P_{data}(I_{X_{Pot}})} [(D(G(I_{X_{Pot}}))) - 1]^2$$
이며, 상기  $D_{Pot}$ 은, 상기 제3 판별자를 나타내고, 상기  $I_{Y_{Pot}}$ 은, 상기 제1 단일 영상에 포함된 포트홀 영역 영상을 나타내며, 상기  $G(I_{X_{Pot}})$ 는, 상기 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 나타낼 수 있다.

## 발명의 효과

[0023] 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치에 의하면, 적대적 생성 신경망(generative adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 기반으로 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 소스 단일 영상에 대한 포트홀 영상을 생성함으로써, 다양한 포트홀 영상을 손쉽게 적은 비용으로 생성할 수 있다.

[0024] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

## 도면의 간단한 설명

[0025] 도 1은 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상

생성 장치를 설명하기 위한 블록도이다.

도 2는 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법을 설명하기 흐름도이다.

도 3은 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크의 구조를 설명하기 위한 도면이다.

도 4는 본 발명의 바람직한 실시예에 따른 불규칙 잡음 추가 과정을 설명하기 위한 도면으로, 도 4의 (a)는 포트홀이 포함된 영상을 나타내고, 도 4의 (b)는 포트홀 객체 영상을 나타내며, 도 4의 (c)는 포트홀 마스크 추출을 위한 의미론적 분할 네트워크를 나타내고, 도 4의 (d)는 포트홀 마스크를 나타낸다.

도 5는 도 3에 도시한 판별자를 설명하기 위한 도면이다.

도 6은 도 3에 도시한 생성자를 설명하기 위한 도면이다.

도 7은 본 발명의 바람직한 실시예에 따른 테스트 영상 생성 과정을 설명하기 위한 도면이다.

도 8은 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크의 성능을 설명하기 위한 도면으로, 도 6의 (a)는 소스 단일 영상을 나타내고, 도 6의 (b)는 포트홀 객체 부분에 대한 판별자를 포함하지 않는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성된 포트홀 영상을 나타내며, 도 6의 (c)는 포트홀 객체 부분에 대한 판별자를 포함하는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성된 포트홀 영상을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0026] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0027] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또한, 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0028] 본 명세서에서 "제1", "제2" 등의 용어는 하나의 구성 요소를 다른 구성 요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예컨대, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고, 유사하게 제2 구성 요소도 제1 구성 요소로 명명될 수 있다.
- [0029] 본 명세서에서 각 단계들에 있어 식별부호(예컨대, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별 부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0030] 본 명세서에서, "가진다", "가질 수 있다", "포함한다" 또는 "포함할 수 있다" 등의 표현은 해당 특징(예컨대, 수치, 기능, 동작, 또는 부품 등의 구성 요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.
- [0033] 이하에서 첨부한 도면을 참조하여 본 발명에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법 및 장치의 바람직한 실시예에 대해 상세하게 설명한다.
- [0035] 먼저, 도 1을 참조하여 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 장치에 대하여 설명한다.
- [0036] 도 1은 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 장치를 설명하기 위한 블록도이다.
- [0037] 도 1을 참조하면, 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 장치(이하 '포트홀 영상 생성 장치'라 한다)(100)는 적대적 생성 신경망(generative

adversarial network, GAN) 및 도메인(domain) 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 기반으로 소스 단일 영상의 배경 부분은 유지하면서 포트홀(pot hole) 객체 부분만 변환하여 소스 단일 영상에 대한 포트홀 영상을 생성한다.

- [0038] 이에 따라, 본 발명은 다양한 포트홀 영상을 손쉽게 적은 비용으로 생성할 수 있다.
- [0040] 이를 위해, 포트홀 영상 생성 장치(100)는 하나 이상의 프로세서(110), 컴퓨터 판독 가능한 저장 매체(130) 및 통신 버스(150)를 포함할 수 있다.
- [0041] 프로세서(110)는 포트홀 영상 생성 장치(100)가 동작하도록 제어할 수 있다. 예컨대, 프로세서(110)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램(131)을 실행할 수 있다. 하나 이상의 프로그램(131)은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(110)에 의해 실행되는 경우 포트홀 영상 생성 장치(100)로 하여금 포트홀 영상 생성 네트워크를 이용하여 배경에 문맥적으로 적합한 포트홀 영상을 생성하기 위한 동작을 수행하도록 구성될 수 있다.
- [0042] 컴퓨터 판독 가능한 저장 매체(130)는 포트홀 영상 생성 네트워크를 이용하여 배경에 문맥적으로 적합한 포트홀 영상을 생성하기 위한 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(131)은 프로세서(110)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 포트홀 영상 생성 장치(100)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0043] 통신 버스(150)는 프로세서(110), 컴퓨터 판독 가능한 저장 매체(130)를 포함하여 포트홀 영상 생성 장치(100)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0044] 포트홀 영상 생성 장치(100)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(170) 및 하나 이상의 통신 인터페이스(190)를 포함할 수 있다. 입출력 인터페이스(170) 및 통신 인터페이스(190)는 통신 버스(150)에 연결된다. 입출력 장치(도시하지 않음)는 입출력 인터페이스(170)를 통해 포트홀 영상 생성 장치(100)의 다른 컴포넌트들에 연결될 수 있다.
- [0047] 그러면, 도 2를 참조하여 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법에 대하여 설명한다.
- [0048] 도 2는 본 발명의 바람직한 실시예에 따른 적대적 생성 신경망을 이용한 배경에 문맥적으로 적합한 포트홀 영상 생성 방법을 설명하기 흐름도이다.
- [0049] 도 2를 참조하면, 포트홀 영상 생성 장치(100)의 프로세서(110)는 적대적 생성 신경망 및 도메인 변환 네트워크를 포함하는 포트홀 영상 생성 네트워크를 학습한다(S110).
- [0050] 즉, 프로세서(110)는 포트홀이 포함된 제1 단일 영상 및 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득할 수 있다.
- [0051] 그런 다음, 프로세서(110)는 제1 단일 영상 및 제1 마스크와, 제2 단일 영상 및 제2 마스크를 기반으로, 포트홀 영상 생성 네트워크를 학습할 수 있다.
- [0052] 여기서, 포트홀 영상 생성 네트워크는 제1 생성자, 제2 생성자, 제1 판별자, 제2 판별자 및 제3 판별자를 포함할 수 있다.
- [0053] 제1 생성자는 입력된 영상 및 마스크를 제2 단일 영상의 도메인으로 변환할 수 있다.
- [0054] 제2 생성자는 입력된 영상 및 마스크를 제1 단일 영상의 도메인으로 변환할 수 있다.
- [0055] 제1 판별자는 제1 단일 영상 및 제1 마스크와, 제2 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득할 수 있다.
- [0056] 제2 판별자는 제2 단일 영상 및 제2 마스크와, 제1 생성자를 통해 출력되는 영상 및 마스크를 기반으로 판별 정보를 획득할 수 있다.

[0057] 제3 판별자는 제1 단일 영상에 포함된 포트홀 영역 영상과 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 기반으로 판별 정보를 획득할 수 있다. 이때, 제3 판별자에 입력되는 포트홀 영역 영상은 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상일 수 있다. 예컨대, 사각형 박스 영역은 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역일 수 있다.

[0058] 그리고, 포트홀 영상 생성 네트워크의 손실 함수는 제1 손실 함수, 제2 손실 함수, 제3 손실 함수, 제4 손실 함수 및 제5 손실 함수를 포함할 수 있다.

[0059] 제1 손실 함수는 제1 판별자와 제2 판별자에 대한 손실 함수로, 아래의 [수학식 1]과 같이 정의될 수 있다.

### 수학식 1

$$L_{advFull} = (D_X(I_X, M_A) - 1)^2 + D_X(F(I_Y, M_B))^2 + (D_Y(I_Y, M_B) - 1)^2 + D_Y(G(I_X, M_A))^2$$

[0061] 여기서,  $L_{advFull}$ 은 제1 손실 함수를 나타낸다.  $D_X$ 는 제2 판별자를 나타낸다.  $I_X$ 는 제2 단일 영상을 나타낸다.  $M_A$ 는 제2 마스크를 나타낸다.  $F$ 는 제1 생성자를 나타낸다.  $I_Y$ 는 제1 단일 영상을 나타낸다.  $M_B$ 는 제1 마스크를 나타낸다.  $G$ 는 제2 생성자를 나타낸다.

[0062] 제2 손실 함수는 제3 판별자에 대한 손실 함수로, 아래의 [수학식 2]와 같이 정의될 수 있다.

### 수학식 2

$$L_{advPot} = E_{I_{Y_{pot}} \sim P_{data}(I_{Y_{pot}})} [D_{Pot}(I_{Y_{Pot}}) - 1]^2 + E_{I_{X_{pot}} \sim P_{data}(I_{X_{pot}})} [(D(G(I_{X_{Pot}})) - 1)^2]$$

[0064] 여기서,  $L_{advPot}$ 는 제2 손실 함수를 나타낸다.  $D_{Pot}$ 은 제3 판별자를 나타낸다.  $I_{Y_{pot}}$ 은 제1 단일 영상에 포함된 포트홀 영역 영상을 나타낸다.  $G(I_{X_{pot}})$ 는 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상을 나타낸다.

[0065] 제3 손실 함수는 영상 및 마스크와 제1 생성자 및 제2 생성자를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 손실 함수로, 아래의 [수학식 3]과 같이 정의될 수 있다.

### 수학식 3

$$L_{cyc} = \|F(G(I_X, M_A)) - (I_X, M_A)\|_1 + \|G(F(I_Y, M_B)) - (I_Y, M_B)\|_1$$

[0067] 여기서,  $L_{cyc}$ 는 제3 손실 함수를 나타낸다.  $\| \cdot \|$ 은 L1 놈(norm)을 나타낸다.

[0068] 제4 손실 함수는 영상 및 마스크와 제1 생성자 및 제2 생성자 중 하나를 통해 생성된 영상 및 마스크와의 차이를 최소화하는 손실 함수로, 아래의 [수학식 4]와 같이 정의될 수 있다.

### 수학식 4

$$L_{idt} = \|G(I_Y, M_B) - (I_Y, M_B)\|_1 + \|F(I_X, M_A) - (I_X, M_A)\|_1$$

[0070] 여기서,  $L_{idt}$ 는 제4 손실 함수를 나타낸다.

[0071] 제5 손실 함수는 영상의 배경은 유지하고 포트홀 객체 부분만 변환되도록 하는 손실 함수로, 아래의 [수학식 5]와 같이 정의될 수 있다.

### 수학식 5

$$L_{ctx} = \|(M_A, M_B) \odot (I_X, I_Y)\|_1 + \|w(M_B, M_A) \odot (I_Y, I_X)\|_1$$

[0072] 여기서,  $L_{ctx}$ 는 제5 손실 함수를 나타낸다.  $I_X'$ 는 도메인이 변환된 제2 단일 영상을 나타낸다.  $M_A'$ 는 도메인이 변환된 제2 마스크를 나타낸다.  $I_Y'$ 는 도메인이 변환된 제1 단일 영상을 나타낸다.  $M_B'$ 는 도메인이 변환된 제1 마스크를 나타낸다.  $w$ 는 가중치로, 1에서 이진 마스크의 요소별 값을 뺀 값을 나타낸다.

[0075] 이후, 프로세서(110)는 포트홀 영상 생성 네트워크를 이용하여 소스 단일 영상의 배경 부분은 유지하면서 포트홀 객체 부분만 변환하여 소스 단일 영상에 대한 포트홀 영상을 생성한다(S130).

[0078] 그러면, 도 3 내지 도 6을 참조하여 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크에 대하여 보다 자세하게 설명한다.

[0079] 도 3은 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크의 구조를 설명하기 위한 도면이다.

[0080] 도 3을 참조하면, 본 발명은 구하기 쉽지 않은 포트홀 객체에 대한 영상 데이터의 부족 현상을 해결하기 위해, 소스 단일 영상에서 배경 부분은 그대로 유지한 채, 불규칙 잡음이 추가된 부분만을 타겟 단일 영상의 객체 형태로 변환하는 인페인팅 형식의 단일 영상간 도메인 변환 네트워크를 이용하여 영상 데이터를 생성하는 기술이다.

[0081] 증량하고자 하는 포트홀 객체가 포함된 단일 영상에서 주어진 포트홀의 위치 정보를 기반으로 포트홀 패치의 의미 있는 부분만 정교하게 이진 마스크로 추출한다. 추출된 마스크를 기반으로 포트홀 위치에 불규칙 잡음을 추가한다. 이렇게 하면 포트홀이 포함된 단일 영상 원본과 포트홀이 포함된 단일 영상에서 포트홀 영역에 불규칙 잡음이 있는 단일 영상이 한 쌍으로 존재하게 된다. 최종적으로 두 단일 영상이 서로의 도메인으로 변환이 되도록 학습한다. 이때, 포트홀의 위치 좌표 정보와 앞서 생성했던 포트홀 부분의 마스크를 함께 추가로 입력하여 배경은 최대한 유지되 불규칙 잡음 부분만 변환이 될 수 있도록 유도한다.

[0083] 도 4는 본 발명의 바람직한 실시예에 따른 불규칙 잡음 추가 과정을 설명하기 위한 도면으로, 도 4의 (a)는 포트홀이 포함된 영상을 나타내고, 도 4의 (b)는 포트홀 객체 영상을 나타내며, 도 4의 (c)는 포트홀 마스크 추출을 위한 의미론적 분할 네트워크를 나타내고, 도 4의 (d)는 포트홀 마스크를 나타낸다.

[0084] 본 발명의 포트홀 영상 생성 네트워크에서는 단일 영상과 마스크로 이루어진 제1 영상 세트, 및 불규칙 잡음을 포함하는 단일 영상과 마스크로 이루어진 제2 영상 세트가 모두 입력된다.

[0085]  $Y$ 는 포트홀이 포함된 원본 단일 영상(즉, 제1 단일 영상)( $I_Y$ ) 및 해당 포트홀의 위치가 표시된 마스크(즉, 제1 마스크)( $M_b$ )이다.  $X$ 는  $Y$ 의 원본 단일 영상에서 포트홀 부분을 불규칙 잡음으로 채워 넣은 단일 영상(즉, 제2 단일 영상)( $I_X$ ) 및 포트홀의 위치가 표시된 마스크(즉, 제2 마스크)( $M_a$ )이다.

[0086] 도 4를 참조하면, 포트홀이 포함된 영상에서 포트홀 객체 영상을 추출한 후, 미리 설정된 사이즈(128\*128 사이즈 등)로 크기 조정을 할 수 있다. 이후, 사각형으로 추출된 포트홀 객체 영상에서 실제 포트홀 모양만 추출할 수 있도록 사전에 학습된 의미론적 분할(Semantic Segmentation) 네트워크에 입력할 수 있다. 여기서, 의미론적 분할 네트워크는 포트홀 영상 일부를 레이블링하여 미리 학습되어 있다. 이후, 의미론적 분할 네트워크를 사용하여 추출된 포트홀 마스크를 본래 위치에 본래 크기로 부착할 수 있다. 해당 마스크를 이용하여 포트홀이 포함된 원본 영상에서 포트홀 부분에 불규칙 잡음을 추가할 수 있다. 이렇게 생성된 영상과 마스크를 본 발명의 포트홀 영상 생성 네트워크의 입력 영상/마스크( $X$ )로 사용할 수 있다.

[0088] 생성자  $G$ (즉, 제2 생성자)는 도메인  $X$ 를 갖는 단일 영상  $I_X$  및 마스크  $M_a$ 를 입력 받아 도메인  $Y$ 를 갖는 단일 영

상  $I_Y$  및 마스크  $M_B$ 와 유사한 도메인의 단일 영상 및 마스크를 생성하도록 학습할 수 있다.

- [0089] 생성자 F(즉, 제1 생성자)는 도메인 Y를 갖는 단일 영상  $I_Y$  및 마스크  $M_B$ 를 입력 받아 도메인 X를 갖는 단일 영상  $I_X$  및 마스크  $M_A$ 와 유사한 도메인의 단일 영상 및 마스크를 생성하도록 학습할 수 있다.
- [0090] 생성자 G로부터 생성된  $G(I_X)$ ,  $G(M_A)$ 는 반대 도메인의 원본 단일 영상  $I_Y$  및 원본 마스크  $M_B$ 와 함께 판별자  $D_Y$ (즉, 제1 판별자)에 입력될 수 있다.
- [0091] 판별자  $D_Y$ 는 입력 받은 단일 영상 및 마스크가 각각 생성된 단일 영상 및 마스크인지 원본 단일 영상 및 마스크인지 판단할 수 있다. 여기서, 단순한 적대적 생성 신경망(GAN)이 아닌 LSGAN(Least Squares GAN)의 손실 구조를 사용하면, 더욱 안정적인 학습이 되므로 LSGAN의 손실 함수를 따를 수 있다.
- [0092] 이를 반대 도메인, 판별자  $D_X$ (즉, 제2 판별자)에도 똑같이 적용하면 위의 [수학식 1]과 같이 표현될 수 있다.
- [0094] 도 5는 도 3에 도시한 판별자를 설명하기 위한 도면이다.
- [0095] 생성된 단일 영상 및 원본 단일 영상에서 포트홀 부분만 사각형 박스로 잘라내어 판별자  $D_{Pot}$ (즉, 제3 판별자)에 입력할 수 있다.
- [0096] 이때, 포트홀 부분의 주변 환경까지 고려하여 문맥적으로 자연스럽게 포트홀이 생성되었는지를 판단하기 위해 실제 포트홀 부분보다 상하좌우 각 미리 설정된 픽셀(10픽셀 등)씩 넓은 부분을 잘라내어  $D_{Pot}$ 에 입력할 수 있다.
- [0097] 판별자  $D_{Pot}$ 은 입력 받은 단일 영상 일부가 각각 생성된 단일 영상의 일부인지 원본 단일 영상의 일부인지 판단할 수 있다. 단, 판별자  $D_{Pot}$ 에는 마스크는 제외하고 단일 영상만 입력되며 X 도메인 단일 영상  $I_X$ 를 Y 도메인으로 변환하는 과정에만 사용될 수 있다. 판별자  $D_{Pot}$ 의 손실 함수는 위의 [수학식 2]와 같다.
- [0099] 생성자 G로부터 생성된 단일 영상 및 마스크를 다시 생성자 F에 입력하고, 생성자 F로부터 생성된 단일 영상 및 마스크를 다시 생성자 G에 입력하게 되면, 아래의 [수학식 6]과 같이  $F(G(I_X, M_A))$ 는  $I_X$ ,  $M_A$ 와 유사한 형태로,  $G(F(I_Y, M_B))$ 는  $I_Y$ ,  $M_B$ 와 유사한 형태로 재생성하게 될 수 있다.

## 수학식 6

[0100] 
$$X \rightarrow G(X) \rightarrow F(G(X)) \approx X, \quad Y \rightarrow F(Y) \rightarrow G(F(Y)) \approx Y$$

- [0101] 원본 단일 영상 및 마스크와 두 개의 생성자를 통해 재생성된 단일 영상 및 마스크 차이를 L1 norm으로 계산한 값을 본 발명의 포트홀 영상 생성 네트워크의 손실 함수로 사용할 수 있다. 두 단일 영상 및 마스크 사이의 도메인 차이가 적을수록 손실 함수는 감소하게 되고, 이를 cycle-consistency(cyc) 손실 함수라고 하며, 위의 [수학식 3]와 같이 표현된다.
- [0103] 도 6은 도 3에 도시한 생성자를 설명하기 위한 도면이다.
- [0104] 원본 단일 영상 및 마스크  $\{I_X, M_A\}$ 와,  $\{I_X, M_A\}$ 를 생성자 F에 입력하고, 원본 단일 영상 및 마스크  $\{I_Y, M_B\}$ 와,  $\{I_Y, M_B\}$ 를 생성자 G에 입력했을 때, 원본 단일 영상과 마스크의 정보를 손실하지 않도록, 생성되는 단일 영상과 마스크  $\{\hat{I}_X, \hat{M}_A\}$ (혹은  $\{\hat{I}_Y, \hat{M}_B\}$ ) 사이의 차이를 L1 norm으로 계산한 값을 손실 함수로 추가하며, 위의 [수학식 4]와 같다.
- [0106] 마지막으로, 생성자 G를 통하여 X 도메인의 단일 영상 및 마스크  $\{I_X, M_A\}$ 에서 Y 도메인의 단일 영상 및 마스크  $\{I_Y, M_B\}$ 로 도메인을 변환하고, 생성자 F를 통하여 Y 도메인의 단일 영상 및 마스크  $\{I_Y, M_B\}$ 에서 X 도메인의 단일 영상 및 마스크  $\{I_X, M_A\}$ 로 도메인을 변환할 때, 전역 문맥 도메인은 유지하고 인스턴스(즉, 포트홀 부분)만 변환하는, 즉, 배경을 유지하는 손실 함수를 추가할 수 있다.

- [0107] 이진 마스크  $M_A$ (혹은  $M_B$ )에서 배경은 1, 인스턴스는 0의 가중치를 두고, 도메인 변환된  $G(M_A)$ (혹은  $F(M_B)$ )와의 차이를 L1 norm으로 계산한 값을 손실 함수로 추가하며 context-preserving(ctx) 손실 함수라고 정의한다.
- [0108] 본 발명의 포트홀 영상 생성 네트워크에는 서로 간의 도메인을 변환하려는 두 단일 영상이 한 쌍으로 입력되며, 두 단일 영상의 배경은 같으므로 ctx 손실 함수를 이용하여 더욱 강인하게 배경을 유지할 수 있다.
- [0109] 즉, 원본 견본을  $\{I_X, M_A\}$  및  $\{I_Y, M_B\}$ , 변환된 견본을  $\{I_X', M_A'\}$  및  $\{I_Y', M_B'\}$ 이라고 할 때, 가중치  $w(M_A, M_B')$ ,  $w(M_B, M_A')$ 는 1에서 이진 인스턴스 마스크의 요소별 값을 뺀 값을 말한다. ctx 손실 함수는 위의 [수학식 5]와 같다.
- [0110] 본 발명의 포트홀 영상 생성 네트워크의 전체 손실 함수는 아래의 [수학식 7]로 표현할 수 있다.

### 수학식 7

$$L_{POTGAN} = L_{advFull} + L_{advPot} + \lambda_{cyc} L_{cyc} + \lambda_{idt} L_{idt} + \lambda_{ctx} L_{ctx}$$

- [0112] 여기서,  $L_{POTGAN}$ 은 전체 손실 함수를 나타낸다.  $\lambda_{cyc}$ ,  $\lambda_{idt}$ ,  $\lambda_{ctx} > 0$ 은 각각 손실의 균형을 맞추는 하이퍼 파라미터의 일부를 나타낸다.
- [0115] 그러면, 도 7 및 도 8을 참조하여 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크의 성능에 대하여 설명한다.
- [0116] 도 7은 본 발명의 바람직한 실시예에 따른 테스트 영상 생성 과정을 설명하기 위한 도면이다.
- [0117] 도 7을 참조하면, 학습 과정에서 추출한 사각형 형태의 포트홀 영상들을 이용하여 적대적 생성 신경망(GAN)을 학습시킨 후, 생성된 적대적 생성 신경망(GAN)을 이용하여 다양한 포트홀 영상을 생성한다. 여기서, 적대적 생성 신경망(GAN)은 PG-GAN(Karras, T., Aila, T., Laine, S., & Lehtinen, J., "Progressive growing of gans for improved quality, stability, and variation." arXiv preprint arXiv:1710.10196. 참조)을 이용할 수 있다.
- [0118] 이후, 생성된 포트홀 영상을 학습 과정에서 이용한 실제 포트홀 모양만 추출하는 의미론적 분할 네트워크에 입력하여 포트홀 마스크를 추출한다.
- [0119] 이와 동시에, 도로상에 포트홀이 없는 주행 환경 배경 영상을 준비한다. 이를 의미론적 분할 네트워크 기반의 도로 영역 검출 네트워크를 이용하여 도로 영역 마스크를 추출한다. 여기서, 도로 영역 검출 네트워크는 DeepLab에 City Scapes Dataset을 입력하여 사전에 학습되어 있다.
- [0120] 앞서 생성된 가상의 포트홀 모양 기반으로 주행 환경 배경의 도로 영역 내에만 붙여넣는다. 영상을 직접 붙여넣지는 않고 포트홀 영역 마스크를 기반으로 가상의 포트홀 모양의 불규칙 잡음을 생성한다. 여기서, 크기는 무작위로 조절될 수 있다. 이와 동시에, 포트홀이 붙여진 위치와 크기가 동일한 이진 마스크도 생성한다.
- [0121] 이렇게 생성된 영상/마스크를 학습 과정에서 학습된 본 발명의 포트홀 영상 생성 네트워크에 입력하면 주행 환경에서의 가상의 포트홀 영상이 생성되게 된다.
- [0123] 도 8은 본 발명의 바람직한 실시예에 따른 포트홀 영상 생성 네트워크의 성능을 설명하기 위한 도면으로, 도 6의 (a)는 소스 단일 영상을 나타내고, 도 6의 (b)는 포트홀 객체 부분에 대한 판별자를 포함하지 않는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성된 포트홀 영상을 나타내며, 도 6의 (c)는 포트홀 객체 부분에 대한 판별자를 포함하는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성된 포트홀 영상을 나타낸다.
- [0124] 도 8의 (a)는 깨끗한 주행 환경 단일 영상이고, 도 8의 (b)는 주행 환경 단일 영상에 불규칙 잡음을 추가한 뒤 INSTA GAN(S. Mo, M. Cho, and J. Shin, "InstaGAN: Instance-aware Image-to-Image Translation," CoRR, vol. abs/1812.1, 2018. 참조)을 기반으로 포트홀 객체 부분에 대한 판별자를 포함하지 않는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성한 포트홀 단일 영상이며, 도 8의 (c)는 깨끗한 주행 환경 단일 영상에 불규칙 잡음을 추가한 뒤 포트홀 객체 부분에 대한 판별자를 포함하는 본 발명에 따른 포트홀 영상 생성 네트워크를 이용하여 생성한 포트홀 단일 영상이다.

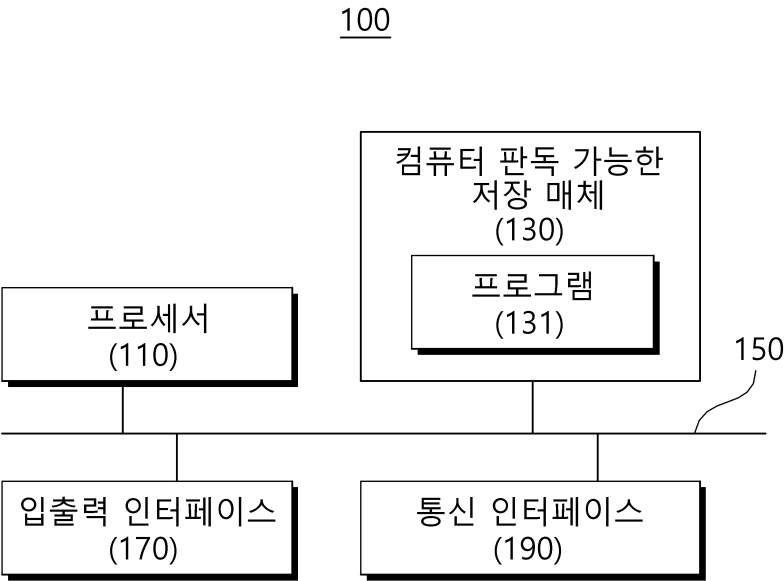
- [0125] 본 발명은 적대적 생성 신경망을 이용하여 포트홀이 포함된 가상의 단일 영상 데이터를 생성하는 방법을 제안한다. 실제 도로 환경에서의 포트홀 객체는 흔하지 않으며 주행 환경에서 취득하기가 매우 어려운 문제가 있다. 이를 해결하기 위해 적대적 생성 신경망과 도메인 변환 방법을 이용하여 가상의 포트홀 단일 영상 데이터를 생성하며, 생성된 영상들은 실제 도로 환경에서의 포트홀과 유사한 모습을 보이는 것을 확인할 수 있다. 이는 객체 검출 및 의미론적 영상 분할 네트워크 등을 학습시키는데 사용할 수 있다. 따라서, 포트홀 데이터를 직접 취득하고 레이블링하는 방법보다 손쉽게 적은 비용으로 다양한 포트홀 영상을 획득할 수 있다.
- [0128] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 저장 매체에 기록될 수 있다. 컴퓨터 판독 가능한 저장 매체는 실행을 위해 프로세서에 명령어를 제공하는데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 저장 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예컨대, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술 분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.
- [0129] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

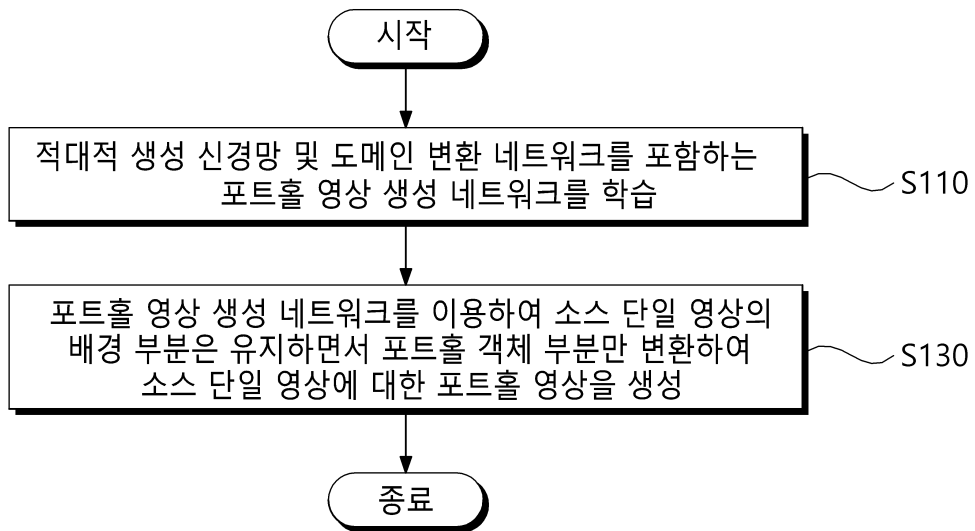
- [0130] 100 : 포트홀 영상 생성 장치,  
110 : 프로세서,  
130 : 컴퓨터 판독 가능한 저장 매체,  
131 : 프로그램,  
150 : 통신 버스,  
170 : 입출력 인터페이스,  
190 : 통신 인터페이스

도면

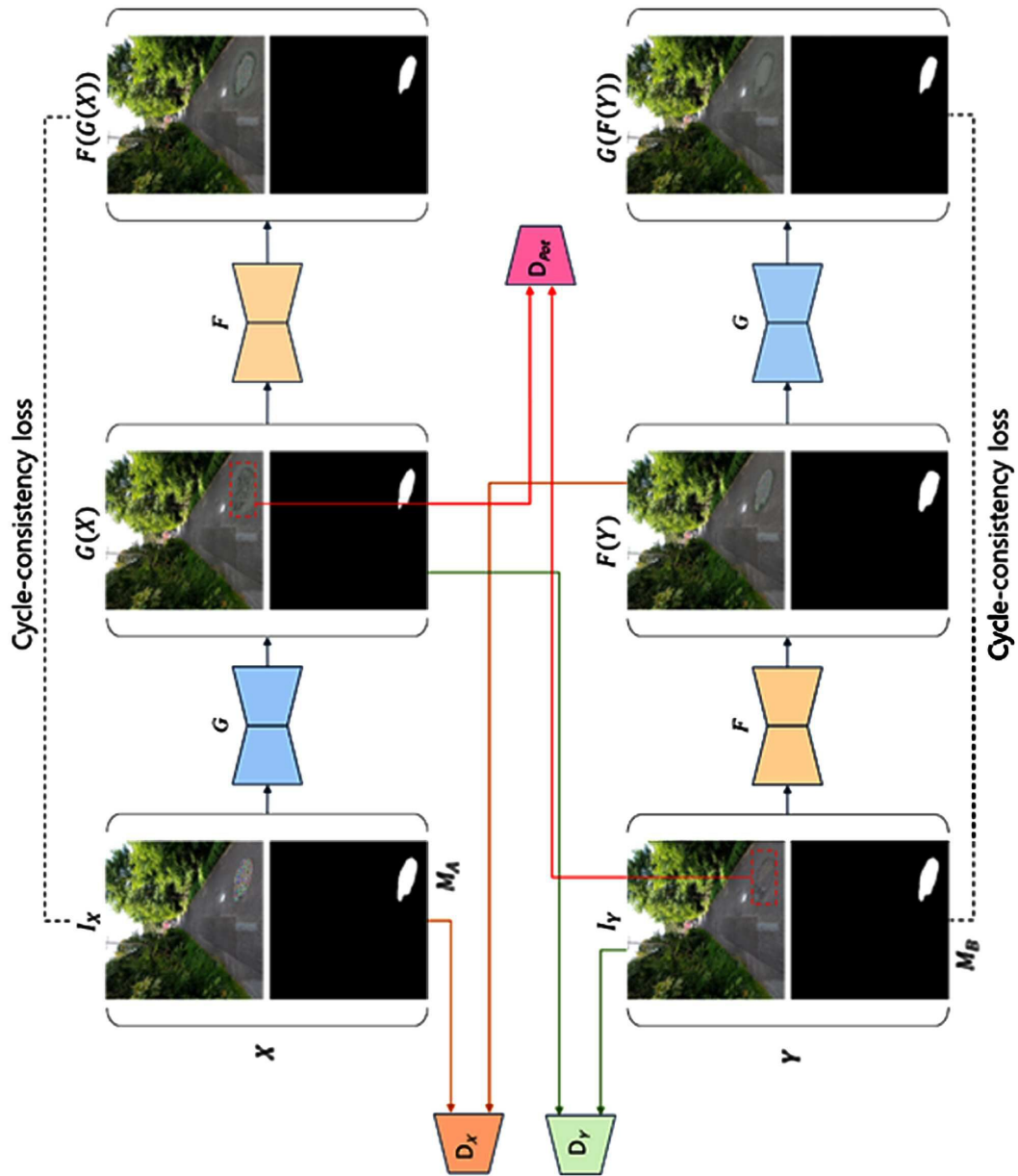
도면1



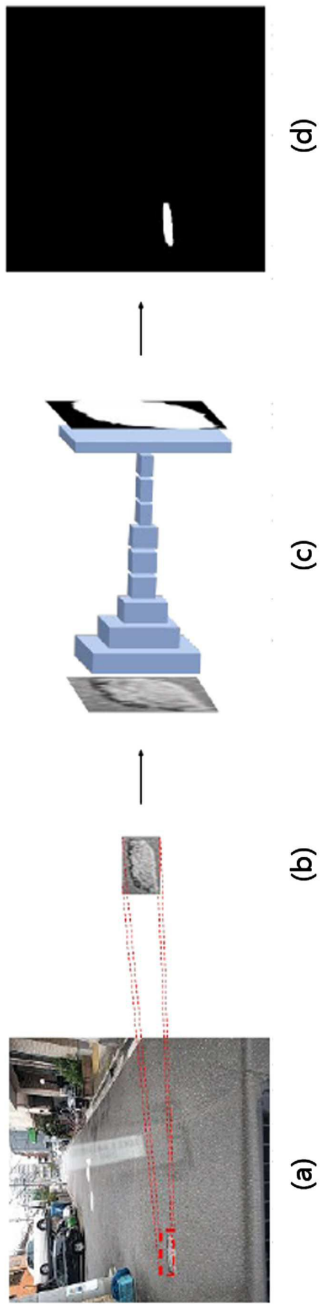
도면2



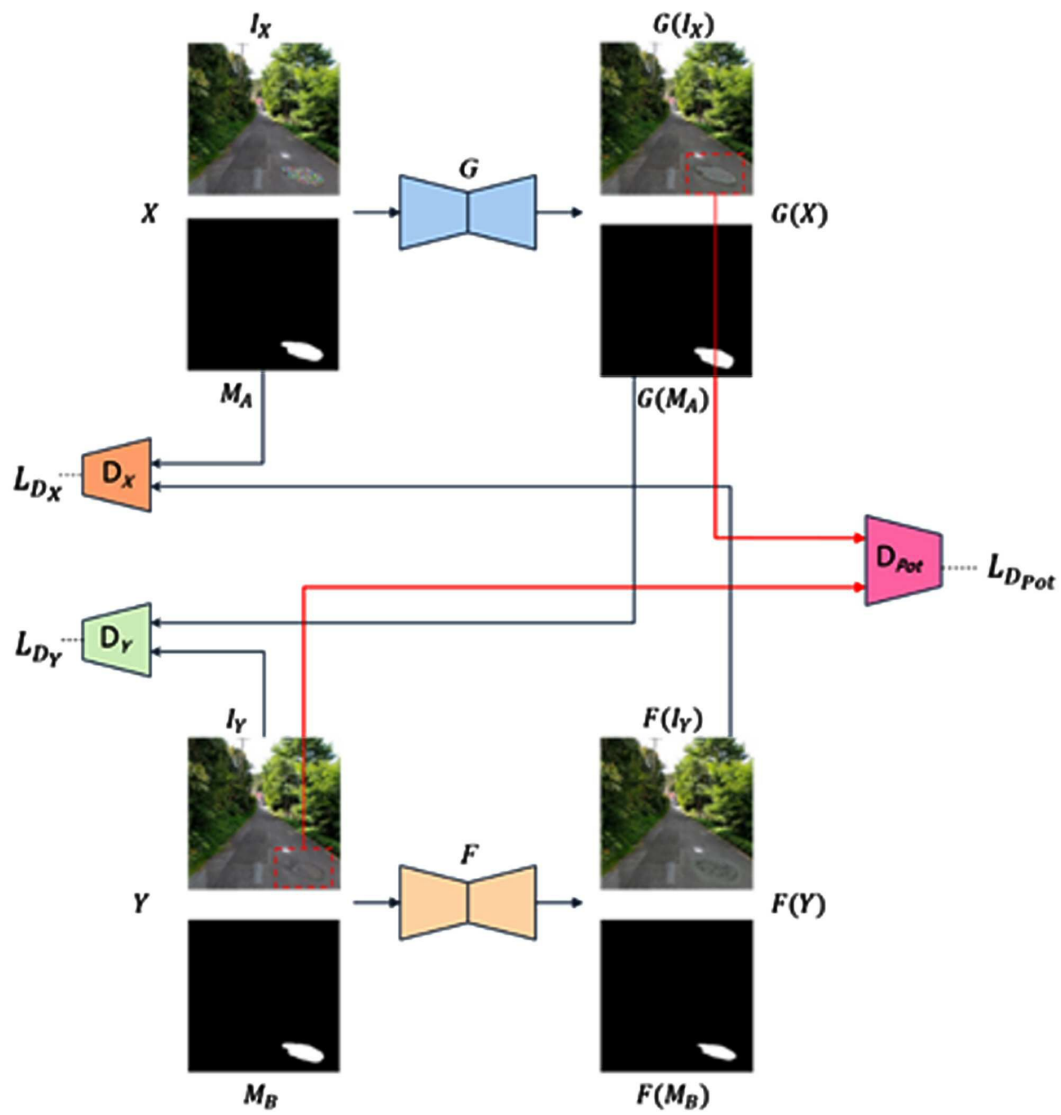
도면3



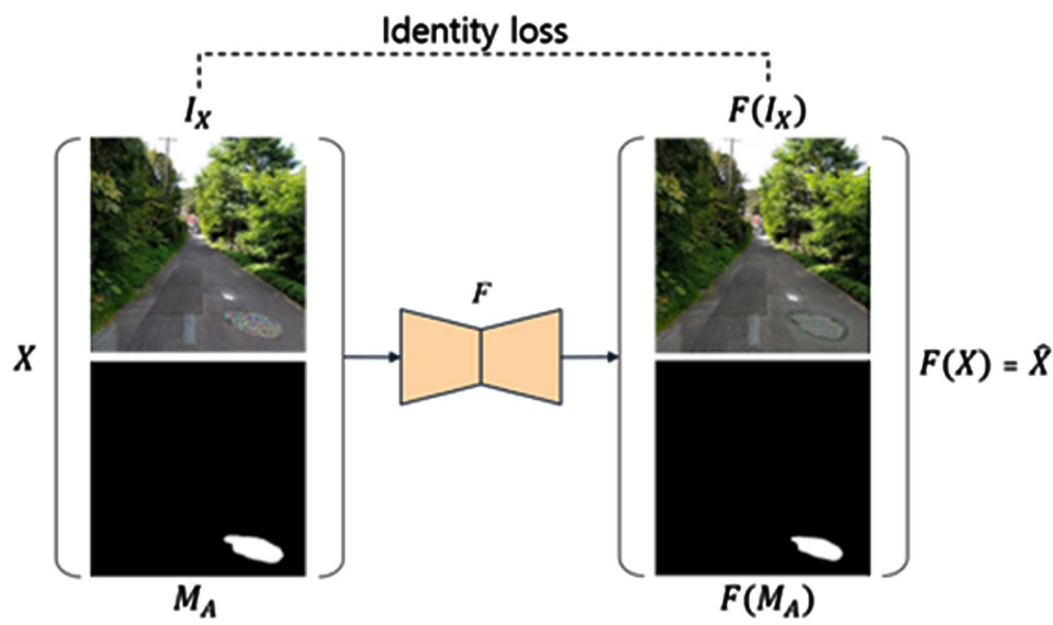
도면4



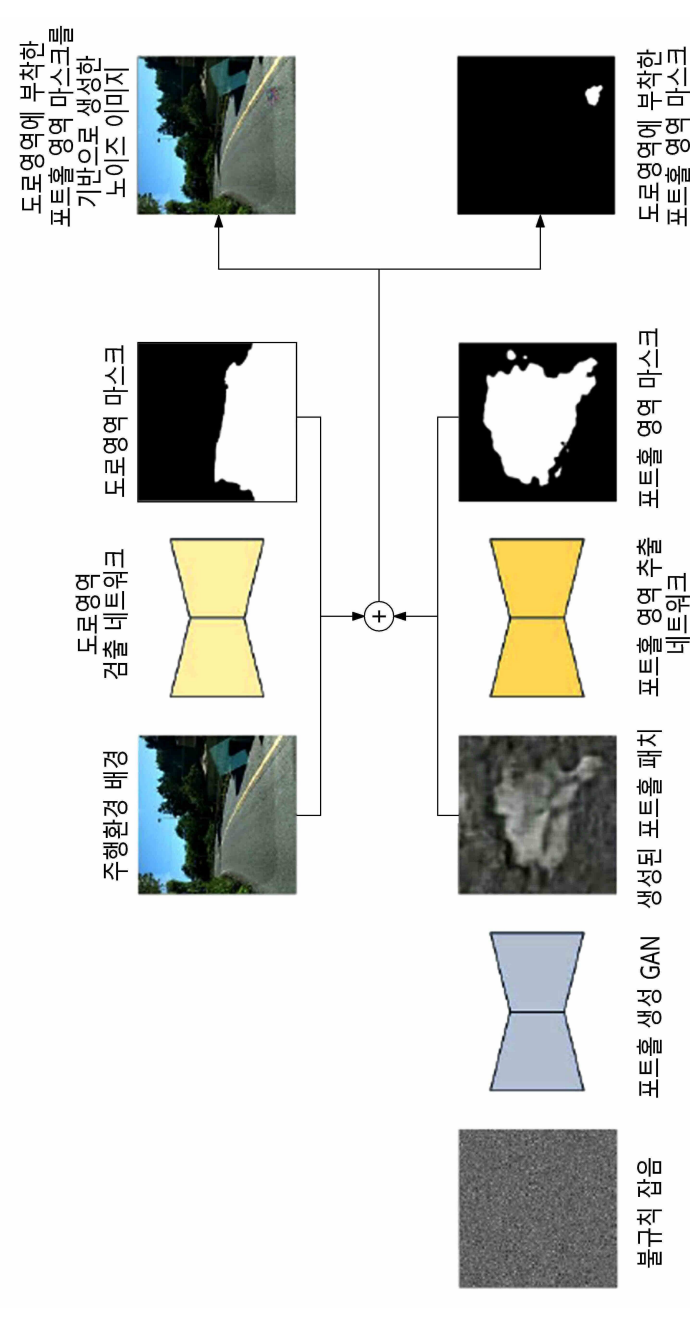
도면5



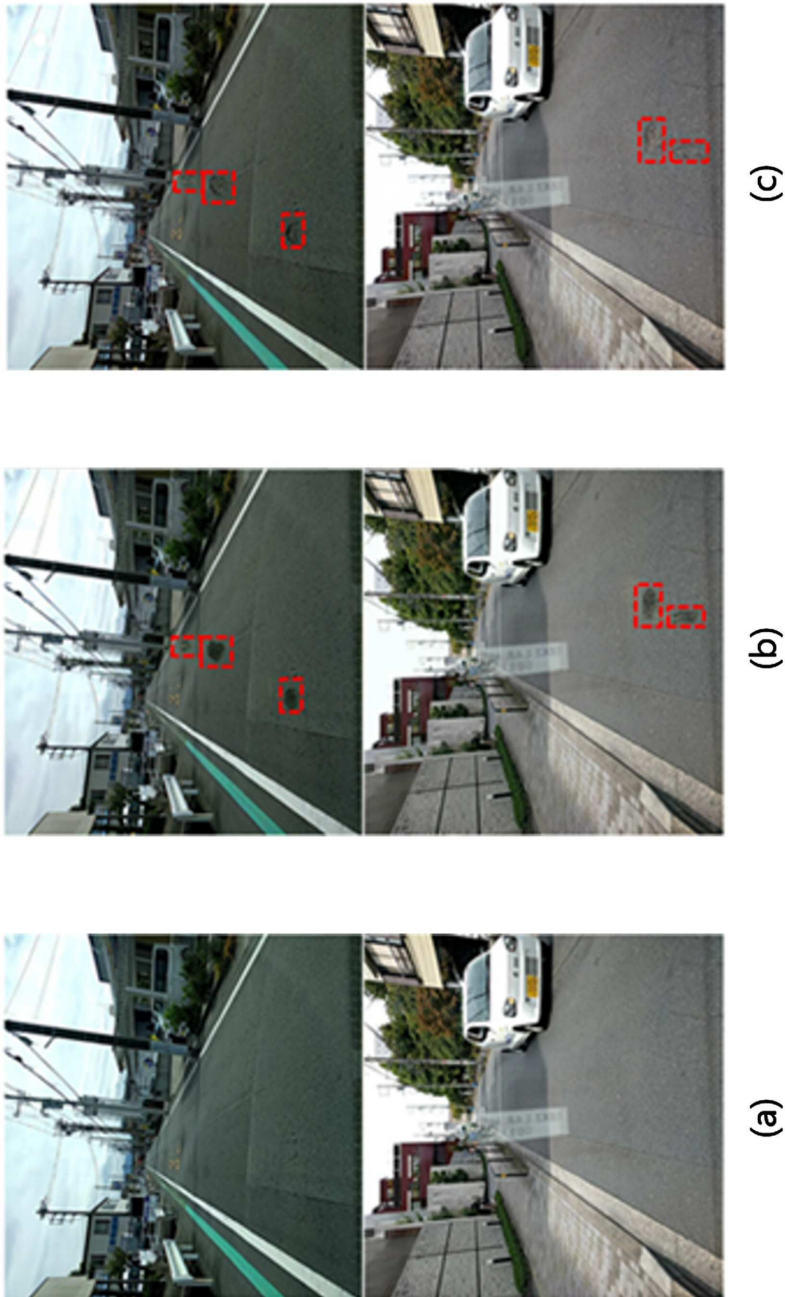
도면6



도면7



도면8



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

프로세서가, 포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하는 단계;

상기 프로세서가, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하는 단계; 및

상기 프로세서가, 상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경

부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하는 단계를 포함하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 포트홀 영상을 생성하는 단계는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하는 단계;

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하는 단계; 및

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는 단계를 포함하는,

포트홀 영상 생성 방법.

#### 【변경후】

프로세서가, 포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하는 단계;

상기 프로세서가, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하는 단계; 및

상기 프로세서가, 상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경 부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하는 단계를 포함하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에

는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 생성 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 포트홀 영상을 생성하는 단계는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하는 단계;

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하는 단계; 및

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는 단계를 포함하는,

포트홀 영상 생성 방법.

## 【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 9

## 【변경전】

포트홀 영상 생성 네트워크를 이용하여 포트홀 영상을 생성하는 포트홀 영상 생성 장치로서,

상기 포트홀 영상 생성 네트워크를 이용하여 상기 포트홀 영상을 생성하기 위한 하나 이상의 프로그램을 저장하는 메모리; 및

상기 메모리에 저장된 상기 하나 이상의 프로그램에 따라 상기 포트홀 영상을 생성하기 위한 동작을 수행하는 하나 이상의 프로세서;

를 포함하고,

상기 프로세서는,

포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에

포함된 포트홀의 위치가 표시된 제2 마스크를 획득하고,

상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하고,

상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경 부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 프로세서는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하고,

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하고,

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는,

포트홀 영상 생성 장치.

#### 【변경후】

포트홀 영상 생성 네트워크를 이용하여 포트홀 영상을 생성하는 포트홀 영상 생성 장치로서,

상기 포트홀 영상 생성 네트워크를 이용하여 상기 포트홀 영상을 생성하기 위한 하나 이상의 프로그램을 저장하는 메모리; 및

상기 메모리에 저장된 상기 하나 이상의 프로그램에 따라 상기 포트홀 영상을 생성하기 위한 동작을 수행하는 하나 이상의 프로세서;

를 포함하고,

상기 프로세서는,

포트홀이 포함된 제1 단일 영상 및 상기 제1 단일 영상에 포함된 포트홀의 위치가 표시된 제1 마스크를 기반으로, 상기 제1 단일 영상에 포함된 포트홀 부분에 불규칙 잡음을 추가한 제2 단일 영상 및 상기 제2 단일 영상에 포함된 포트홀의 위치가 표시된 제2 마스크를 획득하고,

상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 학습 데이터로 하여, 포트홀 영상 생성 네트워크를 학습하고,

상기 포트홀 영상 생성 네트워크를 이용하여, 포트홀이 없는 소스 단일 영상에 대하여, 배경 부분은 유지하면서 포트홀 부분만 변환하여 포트홀 영상을 생성하고,

상기 포트홀 영상 생성 네트워크는, 입력되는 영상 및 마스크를 상기 제2 단일 영상의 도메인으로 변환하는 제1 생성자와, 입력되는 영상 및 마스크를 상기 제1 단일 영상의 도메인으로 변환하는 제2 생성자를 포함하고,

첫 번째 상기 제2 생성자에는, 상기 제2 단일 영상 및 상기 제2 마스크가 입력되고, 첫 번째 상기 제1 생성자에는, 상기 제1 단일 영상 및 상기 제1 마스크가 입력되고, 첫 번째 상기 제2 생성자의 출력은 두 번째 상기 제1 생성자의 입력이 되고, 첫 번째 상기 제1 생성자의 출력은 두 번째 상기 제2 생성자의 입력이 되고,

상기 포트홀 영상 생성 네트워크는, 제1 판별자, 제2 판별자 및 제3 판별자를 더 포함하고,

상기 제1 판별자는, 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 첫 번째 제2 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제2 판별자는, 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 첫 번째 제1 생성자를 통해 출력되는 영상 및 마스크에 대한 판별 정보를 획득하며,

상기 제3 판별자는, 상기 제1 단일 영상에 포함된 포트홀 영역 영상과, 상기 첫 번째 제2 생성자를 통해 출력되는 영상에 포함된 포트홀 영역 영상에 대한 판별 정보를 획득하며,

상기 포트홀 영상 생성 네트워크의 손실 함수는, 상기 제1 판별자와 상기 제2 판별자에 대한 제1 손실 함수, 상기 제3 판별자에 대한 제2 손실 함수, 상기 제2 단일 영상 및 상기 제2 마스크와 상기 두 번째 상기 제1 생성자를 통해 출력되는 영상 및 마스크 간의 차이, 및 상기 제1 단일 영상 및 상기 제1 마스크와 상기 두 번째 상기 제2 생성자를 통해 출력되는 영상 및 마스크 간의 차이에 대한 제3 손실 함수, 및 상기 제1 단일 영상 및 상기 제1 마스크와, 상기 제1 단일 영상 및 상기 제1 마스크를 상기 제2 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이, 및 상기 제2 단일 영상 및 상기 제2 마스크와, 상기 제2 단일 영상 및 상기 제2 마스크를 상기 제1 생성자에 입력하여 얻어지는 영상 및 마스크 간의 차이에 대한 제4 손실 함수의 합으로 구성되고,

상기 포트홀 영상 생성 네트워크는 상기 손실함수를 최소화하도록 학습되며,

상기 제3 판별자에 입력되는 포트홀 영역 영상은, 영상의 전체 영역에서 추출된 포트홀 객체 부분을 포함하는 사각형 박스 영역에 대한 영상이고,

상기 사각형 박스 영역은, 포트홀 객체 부분을 기준으로 하여 미리 설정된 픽셀 크기만큼 상하좌우로 더 넓은 영역이고,

상기 프로세서는,

상기 소스 단일 영상에 대하여, 도로 영역을 검출하고, 상기 도로 영역 내에 가상의 포트홀 모양의 불규칙 잡음을 생성하고,

상기 가상의 포트홀 모양에 대응하는 마스크를 생성하고,

상기 불규칙 잡음이 생성된 소스 단일 영상 및 상기 가상의 포트홀 모양에 대응하는 마스크를, 상기 제2 생성자에 입력하여, 상기 제2 생성자로부터 포트홀 영상을 획득하는,

포트홀 영상 생성 장치.