

AI 기초 연구 - 자유공모

인공 신경망을 이용한 가려진 객체 인식  
및 가림 영역 복원 기술

광주과학기술원

[붙임] 추가조정 서류 양식

## 1. (계속과제) (당해연도) 연차실적보고서

# 2018 GIST개발과제 계속과제 연차실적보고서

### □ 과제개요

연구사업명	AI 기초 연구 - 자유공모		
연구과제명	인공 신경망을 이용한 가려진 객체 인식 및 가림 영역 복원 기술		
연구책임자	고광희	연구책임자 소속	기계공학부
세부과제책임자 [참여교원]			
총 연구기간	2018.01.01. ~ 2018.12.31	총 연구비	50,000,000 원

## I. 연차실적

### 1. 연구 개발 목표 및 결과

#### 1-1. 최종 목표

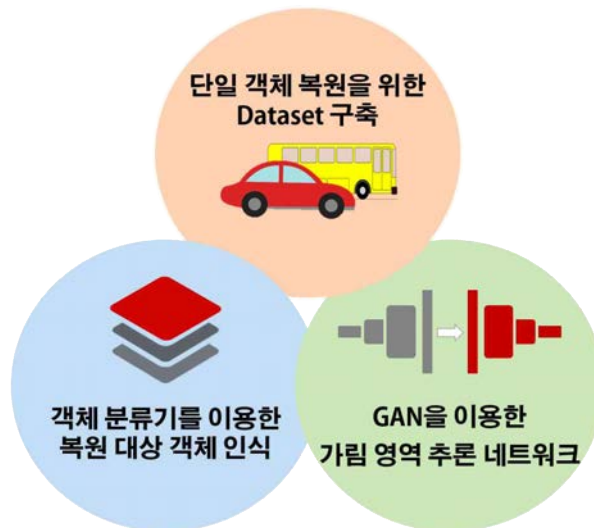


그림 1. 연구 개요

본 연구에서는 대상 객체에서 가려진 부분을 추정하고, 추정된 부분을 GAN을 통해 이질감 없이 복원하는 방법 및 네트워크를 개발한다. 이를 위해 1차 년도에는 관심 객체에서 픽셀 단위의 마스크를 추출하는 Mask R-CNN을 설계하고, 이를 상용 dataset을 사용하여 학습시켜 관심 객체에서 가려진 부분을 자동으로 알려주는 관심 객체 인식기를 개발한다. 또한 mask 영역에 해당하는 관심 객체의 부분을 복원하기 위해 Auto-encoding 구조를 포함한 GAN을 설계하고, 이를 통해 영상에서 가려진 부분을 이질감 없이 복원하는 방법을 제안한다. 본 연구의 2차 년도에서는 3D model과 가상의 시점정보, 이를 통해 생성된 영상들을 네트워크를 학습하고, 선행 연구에서 개발한 객체 인식기와 GAN을 개선함으로써 물체에 의해 가려진 영상의 가림 영역을 복원, 복원된 영상으로 촬영된 카메라의 위치를 파악할 수 있는 localization의 기술로써 확장하고자 한다.

## 1-2. 주요 연구 변경사항

변경된 연구사항은 영상에서 임의 형태의 mask 영역이 아닌 사각형 형태의 빈 부분을 추론하고, 영상의 임의의 부분을 사각형 형태로 지워낸 후 추론하였다. 이와 같은 변경의 사유는 다음과 같다.

첫째로, 본 연구에서 사용한 추론 네트워크의 경우 최대한의 추론 성능을 위하여 최신 수준의 추론 네트워크를 사용했으나, 해당 네트워크의 경우 사각형 형태의 빈 영역에 대한 복원만이 가능하다. 따라서 당초 계획했던 것과 같은 가림 객체 실루엣 형태의 영역에 대해서는 제대로 추론이 되지 않아 추론 성능이 떨어지게 된다. 또한, 당초 계획했던 객체 인식기를 이용한 가림 영역 설정의 경우 객체 인식기가 정확하게 가림 물체가 차지하는 영역을 인식할 필요가 있으나, 현존하는 최고 수준의 네트워크의 경우에도 mAP-50이 60% 가량으로 가림 영역을 정확히 특정해내는 것에는 한계점을 드러내었다.

위에 기술한 연구 문제에 따라, 다중 객체 영상으로부터 객체를 인식해내는 객체 인식기(Object detector) 대신 단일 객체 영상의 객체를 분류하는 객체 분류기(Object classifier)를 사용하였다. 또한 충분한 dataset의 확보를 위하여, 당초 제시하였던 COCO[Lin et al. 2014]를 대신해 Open Images Dataset[Kuznetsova et al. 2018]을 기반으로 하여 객체별 dataset을 구축하였다. 변경된 연구 내용에 대한 개략적인 그림은 다음의 그림 2와 같다.

## 1-3. 당해 연도 개발 목표 및 결과

본 연구에서는 기존의 GAN(Generative Adversarial Network) 기반의 가려짐 추론 네트워크를 개선하여, 객체 영상을 대상으로 한 추론 성능을 향상시키고자 그림 2와 같은 형태의 추론 네트워크 시스템을 제안하였다.

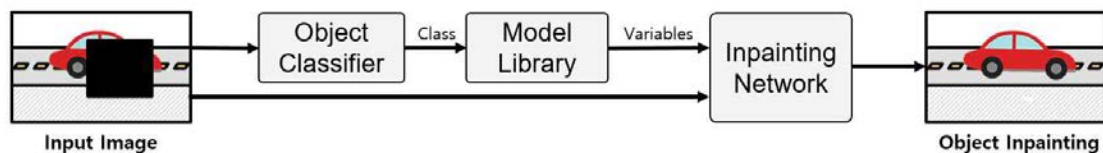


그림 2. 전체 시스템 구성도

제시된 추론 네트워크는 입력 영상 내의 객체를 분류하는 객체 분류기와, 입력 영상 및 객체 분류 결과를 기반으로 알맞은 모델을 통해 가려진 영역을 추론하는 추론 네트워크로 나뉜다. 기존의 추론 네트워크의 경우 다양한 종류의 객체 영상이 혼합된 dataset을 이용해 학습되어 여러 객체 영상에 대해 적용될 수 있으나, 추론 결과물의 완성도에서는 부족한 모습을 보였다. 반면, 본 연구에서 제시하는 추론 네트워크 시스템의 경우 목표 객체에 대한 단일 dataset으로 추론 네트워크를 각각 학습시켜 각 객체에 특화된 네트워크를 구성함으로써 단일 객체에 대한 추론 성능을 크게 향상시켰다. 추론 네트워크의 입력 및 출력 영상의 크기는  $256 \times 256$ 으로, 현존하는 최고 수준의 네트워크와 동일한 수준이다. 또한 객체 분류기를 이용해 입력 영상 내의 가려진 객체의 종류를 분류하도록 함으로써 입력 영상에 대한 추가적인 정보(객체의 종류) 없이 기존 방식과 동일한 입력 조건 하에서 추론 시스템이 작동할 수 있도록 한다. 아래 그림 3은 기존의 방법과 본 연구에서 제시하는 방법을 비교한 것이다.

본 연구에서는 개발하고자 하는 시스템을 도로 교통 상황에 한정하여 적용하기 위해 해당 상황에서 주로 등장하는 객체인 자전거, 버스, 자동차, 사람 및 트럭을 복원 대상으로 하였다. 이들 객체에 대한 영상 dataset을 구축하기 위하여 Stanford Cars dataset[Krause et al. 2013] 및 Open Images dataset을 사용하고자 하였으나, 해당 dataset의 경우 다수의 객체가 하나의 영상에 등장하거나, 이미지

및 객체의 크기가 서로 달라 그대로 사용하기에는 본 연구의 방향성과 맞지 않는 부분이 있었다. 따라서 본 연구에서는 dataset이 제공하는 bounding box 정보를 이용해 대상 객체 영역만을 잘라내고, 그 중에서도 가려짐이 없고 일정 크기(256×256) 이상인 영상만을 사용해 개별 객체 dataset을 구축하였다. 아래의 표 1은 구축된 dataset의 소스와 각 dataset 별 영상의 수이다.

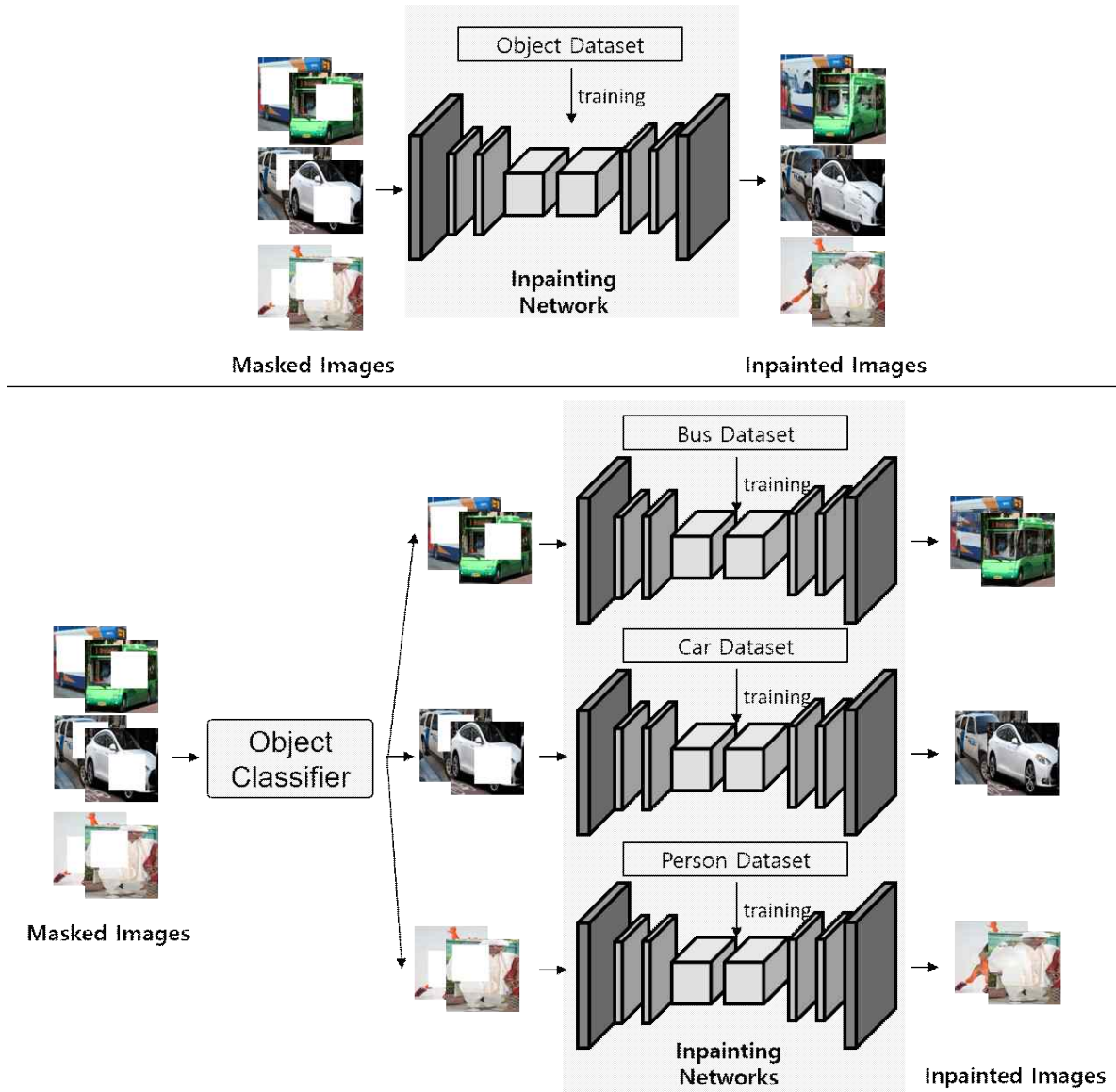


그림 3. 기존의 추론 네트워크(위)와 본 연구에서 제시하는 추론 네트워크 시스템(아래)의 비교

	자전거	버스	자동차	사람	트럭
Train	OpenImages (1334)	OpenImages (3110)	Stanford Cars (8144)	OpenImages (2803)	OpenImages (2524)
Test	OpenImages (167)	OpenImages (167)	OpenImages (167)	OpenImages (167)	OpenImages (167)

표 1. 객체별 dataset의 소스와 영상의 수

또한 복원 결과물에 대한 정량적인 평가를 위하여 Inception-ResNet Score를 도입하였다. Inception-ResNet score는 최신의 객체 분류기인 Inception-ResNet-v2[ Szegedy et al. 2016]를 이용한 GAN 생성물 평가 지표로, Inception-ResNet-v2의 true positive 정확도를 점수로 하여 분류기가 해당 영상을 얼마나 잘 인식하는가를 평가한다. 본 연구에서는 복원 영상의 평가를 위하여 Inception-ResNet-v2를 복원 대상인 자전거, 버스, 자동차, 사람, 트럭에 대해 학습시켜 사용하였다.

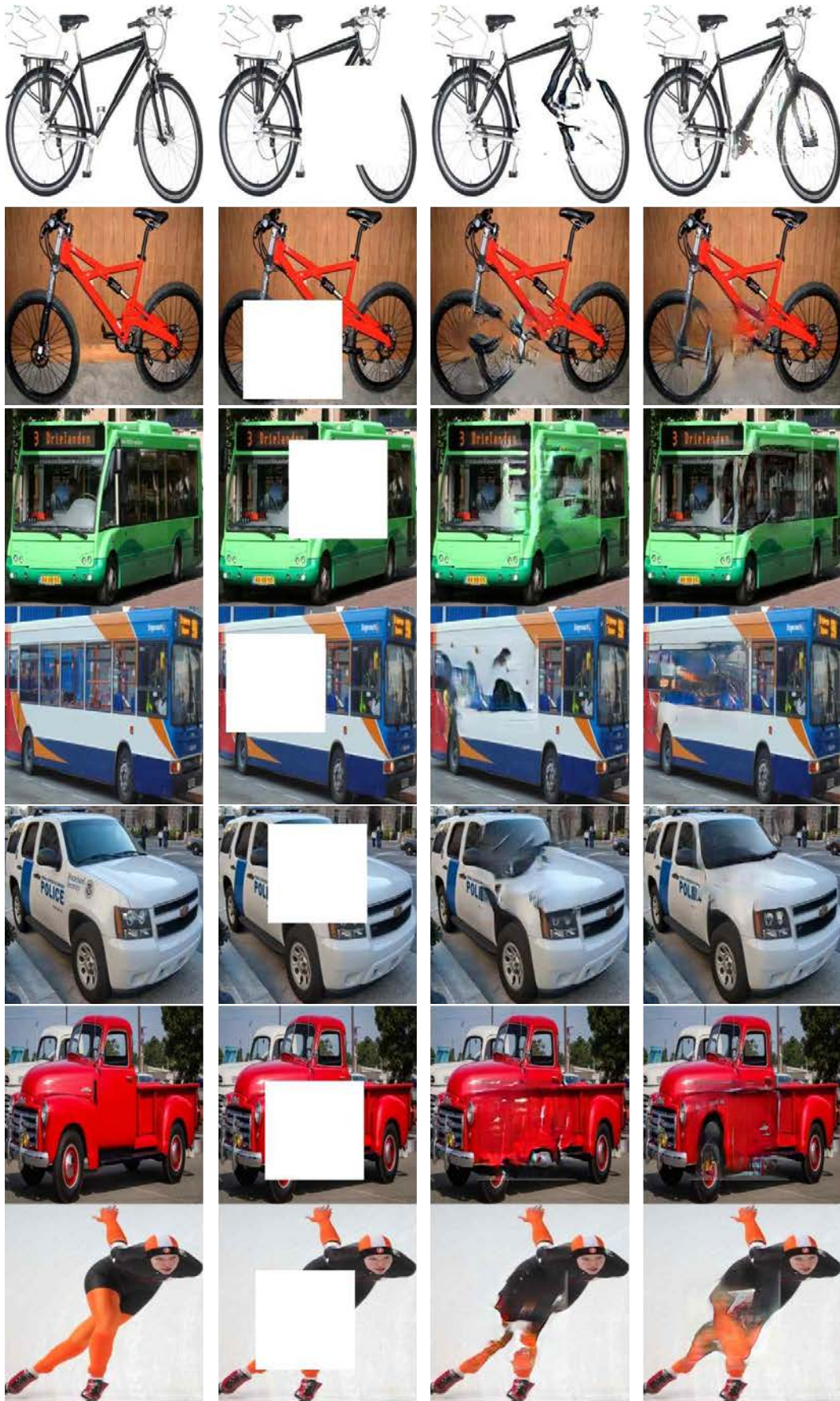
추론 네트워크의 경우 DeepFillv1[Yu et al. 2018]을 기반으로 하여 개별 객체 dataset으로 fine-tuning 했다. DeepFillv1은 최신 수준의 영상 추론 네트워크로 256×256 크기의 영상에서 최대 128×128 크기의 사각형 빈 영역을 추론할 수 있다. 본 연구에서는 ImageNet[Russakovsky et al. 2015] dataset으로 학습된 네트워크를 개별 객체 dataset으로 fine-tuning하여 각각 자전거, 버스, 자동차, 사람, 트럭 영상의 빈 공간을 추론하는 네트워크를 구축하였다. Fine-tuning을 통해 구축한 개별 객체 추론 네트워크는 기존의 ImageNet으로 학습된 네트워크에 비하여 각각의 객체 영상에 대해 아래 표 2와 같이 향상된 추론 성능을 보였다.

	Conventional	Ours
# of True Positive	729	751
Score	82.68	83.61

표 2. 기존 네트워크와 본 연구에서 제시한 네트워크의 성능 비교  
(True Positive의 수, Inception-ResNet Score)

입력된 가림 영상의 경우 해당 영상 내의 객체에 대한 정보가 주어지지 않는데, 개별 객체 추론 네트워크를 적용하기 위해서는 영상 내의 객체에 대한 정보가 필요하다. 따라서 추론 네트워크의 앞부분에 객체 분류기를 적용하여 가려진 객체 종류 정보를 구하고, 이를 기반으로 영상에 적용할 개별 객체 추론 네트워크를 결정한다. 적용된 객체 분류기는 목표 객체 dataset을 통해 학습된 네트워크로, 목표 객체 5종에 대한 분류를 수행하도록 학습되었다. 본 연구에서 제시한 객체 가림 영역 추론 시스템의 최종 결과는 그림 4에 나타나있다.





&lt;Ground Truth&gt;

&lt;Masked&gt;

&lt;Conventional&gt;

&lt;Ours&gt;

그림 4. 본 연구에서 제시한 시스템을 이용한 객체 영상의 빈 영역 추론 결과 및 기존 네트워크와의 비교

## 2. 연구 성과 (※연구성과실적표 제출 불필요)

### 2-1. 정량 실적

성과지표		당초 계획	추진 실적	달성률(%) [=B/A*100]
국외 논문*	SCI	-	-	-
	비SCI	1	0	0
	계(단위:편)	1	0	0
국내 논문*	SCI	-	-	-
	비SCI	1	0	0
	계(단위:편)	1	0	0
총 계		2	0	0

※논문 성과는 카테고리별 상위 20%이내 논문에 한하며 GIST개발과제 사사표기 완료된 논문에 한하여 인정함

▪ Computers & Graphics 저널에 “Unsupervised scene reconstruction using deep embedded clustering” 제목으로 논문 심사 중

## 3. 연구 수행에 따른 문제점 및 개선 방향

본 연구에서 가림영역 추론 네트워크의 경우 사각형 형태의 빈 영역에 대한 복원만 할 수 있으나 이를 개선한 임의 형태의 빈 영역을 복원 할 수 있는 네트워크로 최근 공개되었으며, 기존의 추론 네트워크에 적용 하여 확장성을 높일 것이다.

객체 인식을 통한 가림 영역 설정에서는 사각형 형태의 빈 영역을 복원하는 추론 네트워크로 연구를 수행하여 임의 형태의 빈 영역을 생성할 Mask R-CNN을 사용할 필요가 없었으나 차년도 연구를 위해 객체 분류기를 인식기로 바꾸어 적용할 것이다.

## 참고 문헌

[Lin et al. 2014] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740–755). Springer, Cham.

[Kuznetsova et al. 2018] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., ... & Ferrari, V. (2018). The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale. arXiv preprint arXiv:1811.00982.

[Krause et al. 2013] Krause, J., Stark, M., Deng, J., & Fei-Fei, L. (2013). 3d object representations for fine-grained categorization. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 554–561).

[Szegedy et al. 2016] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In AAAI (Vol. 4, p. 12).

[Russakovsky et al. 2015] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 115(3), 211–252.

## 2. (계속과제) (차년도) 연차계획서(안)

## 2018 GIST개발과제 계속과제 연차계획서

## □ 과제개요

연구사업명	AI 기초 연구 - 자유공모		
연구과제명	인공 신경망을 이용한 가려진 객체 인식 및 가림 영역 복원 기술		
연구책임자	고광희	연구책임자 소속	기계공학부
세부과제책임자 [참여교원]	-		
당해 연도 연구기간	2019.01.01. ~ 2019.12.31	당해 연도 연구비	35,000,000 원

## I. 차년도 연구개발 계획

본 연구의 1차년도에는 관심 객체를 가리고 있는 대상을 제거하고 가려진 부분을 추론하기 위해 GAN 기반의 가림 영역 추론 네트워크를 개발하고, GAN의 입력으로 요구되는 가림 영역에 대한 마스크를 입력 영상으로부터 도출하는 네트워크를 제안하였다. 2차년도에는 이를 적용하여 해양플랜트와 같은 가림 현상이 많은 산업 현장 환경에서 영상을 통해 카메라의 위치를 파악할 수 있는 localization 문제에 적용하고자 한다.

## 가려짐 영역 복원을 통한 이미지 기반 Localization

산업 현장은 굉장히 다양한 물체들로 구성되기 때문에 환경이 복잡하고 움직이는 물체가 많아 물체 간의 가림이 빈번히 발생한다. 자동 로봇이나 자율주행에 있어 localization의 정확도는 이러한 가림 현상에 매우 취약하며, 이러한 문제로 객체 검출기 기반의 localization은 대부분 도로 및 일반 실내 환경 데이터에 특화되어 왔다. 본 연구에서는 설계데이터인 3D model로부터 가상의 시점의 영상(rendering image)들을 생성하고 광원 변화를 적용한 data augmentation을 통해 수렴성을 향상시킬 수 있는 학습 데이터를 확보한다. 카메라의 위치 정보와 해당 시점의 영상, 3D model의 정보들을 네트워크에 학습하고 실제 촬영한 영상을 입력으로 하여 촬영한 영상의 카메라의 위치를 추정할 수 있는 네트워크를 설계함으로써 localization 문제를 해결하고자한다. 또한 물체의 가림이 빈번한 산업 현장에서 촬영한 영상은 localization을 위해 네트워크에 학습한 데이터와 다르며 이러한 차이는 네트워크의 localization 정확도를 떨어뜨린다. 촬영한 영상과 학습한 데이터 사이의 차이점을 줄이기 위해 객체 인식기 네트워크를 통해 가림 영역을 설정하고 임의 형상의 가림영역의 mask로 기존의 가림 영역 추론 네트워크를 개선하여 물체의 가림에 의한 영향을 줄인 복원 영상을 생성한다. 가려짐 영역 복원을 통한 이미지 기반 localization의 전체 시스템 구성은 아래의 그림 5와 같다.



그림 5. 가려짐 영역 복원을 통한 이미지 기반 Localization



## II. 연구비 소요명세서

### 2-1. 연구비 소요명세서(총괄표)

(단위 : 천원)

비목	세목					합계
직 접 비	인건 비	내부인 건비	미지급		8,750	8,750
			지급	현금		
				현물		
		외부인 건비	미지급			
			지급	현금		
				현물		
	학생인건비				13,100	13,100
	인건비소계				13,100	13,100
	연구장비·재 료비	현금		10,000	10,000	
		현물				
	연구활동비				7,000	7,000
	연구과제추진비				4,900	4,900
	연구비소계				21,900	21,900
	위탁연구개발비					
간접비						
연구비 총액				35,000	35,000	

### 2-2. 내부인건비

부서명	성명	직급	월급여	참여기간 (개월)	참여율 (%)	총액	비고
기계공학부	고광희	교수	*****	2019.01.01.-2019.12.31. (12개월)	6.3	****	미지급
합 계						****	

### 2-3. 외부인건비

부서명	성명	직급	월급여	참여기간 [개월]	참여율 [%]	총액	비고
기계공학부	최준호	박사과정	*****	2019.01.01.-2019.12.31. (12개월)	21.9	*****	
기계공학부	손문구	박사과정		2019.01.01.-2019.12.31. (12개월)	21.9		
합 계							

## 2-4. 연구장비재료비

구분	품명	규격	단위	수량	단가	금액	비고
재료	병렬처리용 그래픽카드			3	1,000	3,000	
재료	RGB-D 센서 모듈			3	1,000	3,000	
재료	Dataset 구축 실험용 부품			4	1,000	4,000	
합 계						10,000	

\*구분 란에는 장비/재료 기재

## 2-5. 연구활동비

구 분	산 정 기 준	금 액	비 고
국외여비	·책임급 : 1인×4,500천원×1회=	4,500	
인쇄, 복사, 인화, 슬라이드제작비	50천원×8회=	400	
공공요금	25천원×2회=	50	
제세공과금, 수수료	25천원×2회=	50	
전문가활용비	200천원/회×1회×1인=	200	
문헌구입비	100천원/권×5회	500	
학화·세미나 참가비	50천원×5회×5인=	1,250	
원고료, 통역료, 번역료, 속기료	50천원×1회=	50	
합 계		7,000	

## 2-6. 연구과제추진비

구 분	산 정 기 준	금 액	비 고
국내 출장비	·책임급 : 1인×200천원×5회=1,000 ·원 급 : 2인×150천원×5회=1,500	2,500	
사무용품비	250천원×2회=	500	
회의비	30천원×9회×7인=	1,900	
합 계		4,900	

## Ⅲ. 연차별 연구성과 계획

(단위 : 편, 건, 천원)

구 분		당해연도(2018)	차년도(2019)	합계
대형과제기획	연구과제명 [연구수행기관]	-	-	
	계	-	-	
국외논문 게재 목표	SCI(상위20%)	-	1	
	계	-	1	
특허출원 목표	국내	-	-	
	국외	-	-	
	계	-	-	
특허등록 목표	국내	-	-	
	국외	-	-	
	계	-	-	
기술료/매출액 달성 계획	계약건수	-	-	
	계약액	-	-	
	징수액	-	-	
언론홍보 계획	TV	-	-	
	중앙지	-	-	
	계	-	-	
기타				

※ 1차년도 계획서상 제출한 연구성과 계획에 변경이 있는 경우에 한하여 작성

※ 기타 : 과제 특성에 맞는 성과지표 설정

※ 논문 성과는 카테고리별 상위 20% 이내 논문에 한함