

자원 제약 모바일 로봇을 위한 경량 장면변화감지 기반장면 분류 기법

신 동 훈*, 김 재 호^o

Scene Classification Based on Light Weight Scene Change Detection for Resource-Constrained Mobile Robots

Dong-Hun Shin*, Jae-Ho Kim^o

요 약

모바일 로봇은 정밀한 임무 수행을 위하여 공간의 컨텍스트에 대한 지속적인 인지가 요구된다. 최근 공간의 특징을 정확히 인식하기 위하여 카메라로부터 수집되는 이미지를 활용한 CNN기반 딥러닝 장면 분류기들이 사용되고 있다. 이러한 장면분류 기법은 이미지의 특징 추출을 기반으로 하기 때문에 많은 계산량을 요한다. 서비스 로봇의 활성화를 위해서 모바일 로봇의 경량화가 필수적이나, 고 연산을 요구하는 장면 분류기를 통한 환경인식은 높은 연산량과 많은 배터리 소모를 야기하고 있다. 본 논문에서 기존의 딥러닝 기반 장면 분류기와 유사한 성능을 가지면서도 낮은 연산량을 가지는 장면변화감지 기반 장면분류 기법을 제안하였다. 제안된 경량의 장면분류 기법은 적은 연산으로 동작 가능한 장면변화 감지기를 활용하여 이동하는 로봇의 공간이 변화되는 시점을 효과적으로 탐지해내고 공간의 변화가 감지된 경우에만 정밀한 딥러닝 장면분류기를 동작시킨다. 제안된 경량 장면분류기는 실내로봇을 대상으로 적용하기 위하여 실내 환경에서 촬영된 영상 데이터들을 활용하여 성능평가를 진행하였다. 성능 분석결과 무의미한 수준의 장면분류의 성능 저하에서도 약 45.49% 수준의 컴퓨팅 자원을 절약할 수 있었다. 개발된 장면분류기는 최근 급속히 확대되고 있는 모바일 로봇의 경량화에 기여할 것으로 기대한다.

키워드 : 장면분류, 자원 제약, 모바일 로봇, 장면변화감지, 히스토그램

Key Words : Scene Classification, Resource-Constrained, Mobile Robots, Scene Change Detection, Histogram

ABSTRACT

The mobile robot requires continuous awareness of the context of the spatial environment to carry out given missions precisely. Recently, CNN, a kind of deep learning technology, has been used to classify scenes from camera images to recognize the space. We need lightweight robots so that mobile robots can be used in various applications broadly, but the CNN-based scene classifier requires high computation resources and battery consumption. In this paper, we propose a scene classification technique based on scene change detection to reduce computational complexity. The proposed scene classifier shows a similar performance to the previous CNN-based scene classifier, but the computational complexity is about half. The proposed classifier uses a low-complexity scene change detector to reduce complexity. The compute-intensive CNN classifier is only

* 본 연구는 환경부의 재원으로 한국환경산업기술원의 지능형 도시수자원 관리사업(2019002950003)의 지원을 받아 연구되었습니다.

• First Author : Sejong University Department of Electrical Engineering, okeskr@gmail.com, 정희원

o Corresponding Author : Sejong University Department of Electrical Engineering, kimjh@sejong.ac.kr, 정희원

논문번호 : 202111-311-0-SE, Received October 30, 2021; Revised November 29, 2021; Accepted November 29, 2021

triggered when the scene change is detected. We evaluated the proposed scene classifier using indoor videos to verify whether it can be applied to indoor mobile robots. The analysis results show that the proposed classifier can achieve scene classification with similar accuracy with computing resources of about 45.49%.

1. 서론

최근공장 자동화 등 제조 현장 위주로 사용되던 로봇은 최근 인공지능 기술의 발전과 함께 응용 분야가 점점 다양해지고 있다. 특히, 택배를 배송하거나, 음식점 서빙, 병원 및 가정용 로봇 등 이동환경에서 자율성을 기반으로 서비스되는 모바일 로봇의 사례들이 늘어나고 있다.

모바일 로봇은 주어진 임무를 더욱 정확하게 해결하기 위해 현재 위치하고 있는 공간을 인식할 필요가 있다. 예를 들어 실버케어 로봇은 공간에 대한 컨텍스트를 알고 있는 상황에서 사용자의 요청을 더욱 정확히 이해할 수 있다. 환경을 올바르게 예측할수록 정확한 임무 수행이 가능하므로, 올바른 환경 인지를 위한 장면분류(Scene Classification)의 성능을 높이는 것은 중요한 요소 중 하나이다. 장면분류의 정확도를 높이기 위해 합성곱 신경망(Convolutional Neural Network, 이하 “CNN”으로 표현)을 사용한 장면분류 기법이 제안되었으며 합성곱 신경망을 이용한 장면분류 기법은 이미지로부터 특징(feature)을 추출하고 추출된 특징을 기반으로 장면을 분류하게 된다. 추출된 특징을 활용하는 방법에 따라 이미지로부터 CNN을 활용하여 직접 장면분류 하는 방법 (Direct Scene Classification, D-SC)^[1]과 이미지의 객체를 먼저 인식하고 장면에 따라 객체가 나타나는 특성을 기반으로 장면분류하는 기법(Object Based Scene Classification, O-SC)^[2]이 제안되었다. 직접 장면분류 기법은 입력받은 이미지에 대해 CNN을 이용하여 특징을 추출하고, 추출된 특징을 기존의 장면의 특징과 직접 비교하여 장면을 분류한다. 객체기반의 장면분류 기법은 먼저 입력받은 이미지로부터 추출된 특징을 이용하여 물체를 검출한다. 이후 검출된 객체의 존재 여부와 출현빈도를 기반으로 장면을 분류한다. 합성곱 신경망을 사용한 장면분류 기법(CNN-based Scene Classification)은 이미지로부터 추출된 특징을 기반으로 장면을 예측한다. 따라서 추출되는 특징이 정교하고 다양할수록 장면분류를 위한 비교대상이 많아져 장면분류의 정확도가 높아지게 되며 이미지로부터 다양한 특징을 추출하기 위해서는 고연산을 지원하는 GPU장비 갖추고 있거나 자

원이 충분한 환경이 필요하다. 하지만, 모바일 로봇은 크기와 가격에 대한 제한으로 충분한 컴퓨팅 자원과 배터리 등의 확보가 어려운 문제점이 있다.

이러한 문제를 해결하기 위하여 본 논문에서는 간단한 연산을 통해서 장면변화를 감지하고 장면변화의 감지 시점에 장면분류를 수행하는 장면변화감지 기반 장면분류(CNN-based Scene Classification with Change Detection, 이하 “CD-SC 방식”으로 표현)를 제안한다. 본 논문에서 제안하는 CD-SC 방식은 고연산의 합성곱 신경망을 이용한 특징추출기반의 장면분류 횟수를 감소시켜 자원사용량을 줄이고 배터리 소모 문제를 해결한다. CD-SC 방식은 서로 다른 실내 공간이 공간 내에 있는 물체와 조명에 따라 각기 다른 색(hue)과 밝기(brightness)를 가지게 된다는 특성을 이용한다. 따라서 색과 밝기의 변화량을 통해 장면변화를 감지하고, 장면변화가 감지되면 합성곱 신경망기반의 장면분류를 수행하여 환경을 인지한다. CD-SC는 환경변화를 감지하는 장면변화 감지 모듈(Scene Change Detection module)과 장면분류를 수행하는 합성곱 신경망을 사용한 장면분류 모듈(CNN-based Scene Classification module)로 구성된다. 장면변화 감지 모듈은 과거 인식된 장면에 대해 현재 장면의 색상과 밝기의 유사도를 비교하여 장면변화를 감지한다. 장면분류 모듈에서는 기존에 제안된 CNN기반 장면분류 방법 중 D-SC를 사용하였다. CD-SC는 장면변화감지 모듈을 통해 장면변화 시점을 감지하고, 장면변화가 감지된 경우에만 장면분류를 수행한다. 따라서 모든 장면에 D-SC를 수행하는 방법과 유사한 장면 인식의 정확도를 가지면서도 고연산의 장면분류의 수행횟수 감소에 따라 연산량이 줄어들었다. 모바일 로봇이 실내 환경에서 공간을 인식하고 임무를 수행하는 시나리오에서 CD-SC의 성능을 평가하기 위해 실내공간을 이동하며 촬영된 영상을 기준으로 기존에 제안된 D-SC방법과 제안된 CD-SC 방식의 장면분류 정확도와 연산량을 분석, 평가하였다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 장면분류에 대한 기존 연구를 소개하고, 3장에서는 CD-SC의 상세 구성과 절차에 관해 기술한다. 그리고 4장에서는 CD-SC의 성능 검증을 위하여 기존의 장면분류기법과 비교 분석을 수행한다.

마지막으로 5장에서는 결론과 향후 과제를 논한다.

II. 관련 연구

GIST^[3], SIFT^[4], Fisher Vector^[5]와 같이 초기에 제안된 장면분류 기법은 이미지가 가지고 있는 색상, 움직임, 밝기의 공간적 분포 속성을 이용하여 장면을 분류하였다. 그러나, 이러한 장면분류 기법은 단순한 특징(feature)를 기반으로 장면을 유추하기 때문에 이미지에 물체 간의 간섭이나 구조적 변형이 있을 경우 정확도가 낮아지는 한계점이 있다. 이러한 한계점을 극복하기 위해 이미지에 있는 다양한 특징 추출하기 위한 합성곱 신경망(Convolution Neural Network)이 제안되었다. CNN기반의 장면분류기법은 AlexNet^[6], GooLeNet^[7], VGGNet^[8] 등의 모델에 MIT Indoor 67^[9], SUN397^[10], Places365^[11]과 같은 데이터셋을 기반으로 모델을 학습하고, 학습된 모델을 기반으로 장면분류를 수행한다.

직접 장면분류기법(Direct Scene Classification, D-SC)은 입력된 이미지에 대해 기존의 학습된 장면의 특징과 직접 비교하였을 때 가장 유사한 장면으로 예측한다. 비교 대상이 되는 특징을 추출하기 위하여 입력 이미지 전체를 대상으로 특징을 추출하여 비교하는 방법^[1], 이미지를 작은 조각(patch)으로 나누어 각 조각의 특징을 추출하여 비교하는 방법^[12,13], CNN내의 특징 레이어 계층에서 특징을 추출하여 비교하는 방법^[14,15]이 사용되었다. 클래식기반의 장면분류 기법은 다양한 특징을 비교할 수 있도록 대용량의 학습데이터를 필요로 하며, 학습데이터의 클래스가 얼마나 잘 분류되어 있는지, 입력된 이미지로부터 추출된 특징이 기존에 학습된 데이터의 특징과 얼마나 유사한지에 따라 정확도에 영향을 받게 된다.

객체기반의 장면분류 기법(Object Based Scene Classification)은 CNN을 이용하여 추출된 이미지의 특징을 이용하여 객체를 검출하고, 장면과 객체간 관계성을 이용하여 장면을 분류한다. 이를 활용하여 장면에서 출현하는 객체의 빈도수와 장면을 대표하는 객체를 이용하여 장면을 분류하는 방법^[16], 이미지에서 검출된 객체 각각의 의미론적 관계를 기반으로 장면을 분류하는 방법^[17,18]이 제안되었다. 객체기반의 장면분류 기법은 검출된 객체를 기반으로 장면을 분류하기 때문에 이미지에서 추출된 특징 중 일부 불명확한 특징에 대해 강건하다는 장점을 지니지만, 객체간의 관계가 사전에 정의되어 있어야 한다는 한계점을 지닌다.

기존에 제안된 장면분류 방법들을 모바일 로봇에 적용하기 위해서는 카메라로부터 전달받은 신호를 이미지로 변환하여 저장하고 변환된 이미지에 대해 장면을 분류하는 과정을 반복해야 한다. 그러나 모바일 로봇에서 고연산의 CNN 기반 장면분류를 지속적으로 수행하게 되면 가용자원의 부족으로 정상적인 서비스를 제공하기 어려워지며, 과도한 전력 소모로 인해 배터리의 방전이 발생할 수 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 주기적으로 수집된 이미지를 장면분류 수행 이전에 장면변화를 감지하고 장면이 변화된 시점에만 장면분류를 수행하는 장면변화감지 기반 장면분류(CNN-based Scene Classification with Change Detection)를 제안한다. CD-SC는 주기적으로 수집되는 모든 이미지에 대해 기존에 제안된 직접 장면분류 기법(Direct Scene Classification)을 사용했을 때와 비교하여 유사한 정확도를 보이면서, 고연산의 CNN 기반 장면분류 횟수를 감소시켜 자원 제약적인 환경에서의 연산량 감소를 통한 여유 자원확보가 가능하다.

III. 장면변화감지 기반 장면분류(Change Detection based Scene Classification)

3.1 장면변화감지 알고리즘

장면변화감지 모듈은 카메라로부터 수집된 이미지에 대해 노이즈 감쇄와 이미지 축소를 수행하는 이미지처리 과정(Image-Processing)과 기존에 인식된 장면과 현재 장면의 색(hue)과 밝기(brightness)의 유사도를 비교하여 장면변화를 감지(Scene Change Detection)를 수행하는 두 가지 과정으로 나뉘며, 상세 구조와 동작은 그림 1과 같다.

이미지처리 과정은 모바일 로봇에 장착된 카메라가 외부 환경을 인지하기 위하여 주기적으로 이미지를 수집하면서 시작된다. 카메라로부터 촬영 및 수집된 이미지는 전기적 신호의 영향으로 노이즈가 포함될 수 있다. 노이즈는 장면변화감지 및 장면분류에 영향을 끼칠 수 있으므로 이를 해결하기 위해 가우시안 블러(Gaussian Blur)를 적용하여 노이즈를 감쇄한다. 이때, 가우시안 블러에서 사용되는 커널(kernel)은 커널의 중심일수록 가중치를 높게 주고 멀어질수록 가중치를 낮게 주어 원래 이미지의 형태를 최대한 유지하면서도 노이즈를 완화하게 된다. 모바일 로봇에 장착된 카메라의 설정에 따라 수집되는 이미지는 다양한 크기를 가지게 되는데, 이미지의 크기가 크면 연산량이 많아지게 된다. 이러한 문제점을 해결하기 위

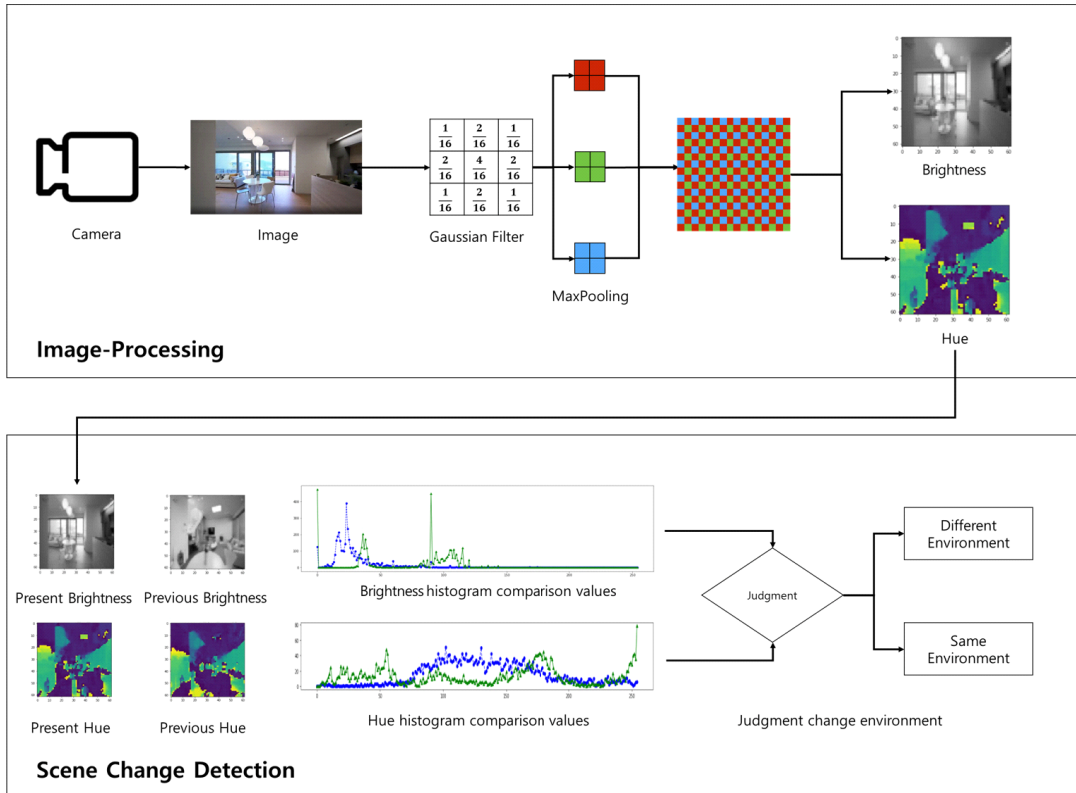


그림 1. 장면변화감지 모듈 처리 흐름도
Fig. 1. Scene change Detection Module Process Flow

하여 이미지의 크기는 감소시키고 이미지가 가지고 있는 특성을 유지하기 위해 맥스 풀링(Max Pooling)을 사용했다. 맥스 풀링을 통해 커널은 이미지 내부를 순회하면서 커널이 위치한 곳에서의 가장 크기가 큰 값을 대푯값으로 추출하게 된다. 따라서 맥스 풀링을 적용한 이미지는 이미지가 가지고 있는 특성은 유지하면서도 이미지가 크기가 감소하게 된다. 이미지처리 과정이 완료되면 카메라로부터 수집된 이미지는 노이즈가 감소되어 있고 특성을 유지하면서도

$$Gray = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \quad (1)$$

$$V = \max(R, G, B)$$

$$H = \begin{cases} \frac{60(G-B)}{V - \min(R, G, B)} & \text{if } V = R \\ 120 + \frac{60(B-R)}{V - \min(R, G, B)} & \text{if } V = G \\ 240 + \frac{60(R-G)}{V - \min(R, G, B)} & \text{if } V = B \end{cases} \quad (2)$$

$$\text{if } H < 0, H = H + 360$$

크기가 줄어든 상태가 된다. 이후 이미지는 색상(Hue)와 밝기(Brightness)에 대한 비교를 위하여 RGB 이미지 데이터를 색상과 명도에 대한 행렬(Matrix)로 변환한다. 명도를 나타내는 행렬로 변환하기 위해 ITU에서 제안된 방법^[9]인 수식 1을 사용하였으며, 색상에 대한 표현으로 변환하기 위해 수식 2를 사용하였다. 장면변화 감지는 현재 장면과 이전 장면의 색과 밝기에 대한 유사도를 비교하여 판단한다. 이미지처리 결과 현재 장면의 밝기와 색상을 나타내는 행렬은 숫자의 분포로 표현 가능하며, 하나의 행렬은 하나의 히스토그램으로 나타낼 수 있다. 따라서 히스토그램 비교 방법을 통해 이전장면과 현재 장면의 색과 밝기에 대한 각각의 유사도를 계산할 수 있으며, 서로 다른 공간은 색 또는 밝기가 다르므로 색과 밝기 중 하나라도 유사도의 값이 임계치 미만이면 장면이 변화되었다고 판단한다. 또한, 이전 장면에 대한 정보가 없는 경우에도 장면변화가 감지되었다고 판단한다. 장면변화가 감지되면, 장면분류 모듈로 현재 장면에 대한 장면분류가 필요함을 알린다. 이에 따른 장면변화감지의 상세 알고리즘은 표 1과 같다.

표 1. 장면분류감지 방법 알고리즘
Table 1. Scene Change Detection Algorithm

```

1: Input:
2:   It ← Image captured by the camera at "t" time
3:   Tb, Th ← Threshold for the amount of change
      for brightness and hue
4: Output:
5:   If the scene has changed (True or False)
6: Initial:
7:   Cp, Bp, Hp ← Class, Brightness, Hue of
      Previous Scene
8:   Hm() : Compare to Histogram function
9: While It != null :
10:  | G ← Gaussain_Blur(It)
11:  | M ← Max_Pooling(G)
12:  | Bc ← matrix brightness of Current Scene
13:  | Hc ← matrix hue of Current Scene
14:  | if Cp == null:
15:  |   | (Bp, Hp) ← (Bc, Hc)
16:  |   | return True;
17:  | else:
18:  |   | Vb ← Hm(Bp, Bc): Brightness histogram
19:  |   | comparison value
20:  |   | Vh ← Hm(Hp, Hc): Hue histogram
21:  |   | comparison value
22:  |   | if Vb < Tb or Vh < Th :
23:  |   |   | (Bp, Hp) ← (Bc, Hc)
24:  |   |   | return True;
25:  |   | else:
26:  |   |   | return False;
27:  | end
28: end

```

* C_p ← Class of Current Scene(In Scene Classification module)

3.2 장면분류 방법

장면변화가 감지되었을 경우 장면분류 모듈은 전달 받은 이미지에 대한 장면을 CNN 기반으로 분류하여 환경을 인지한다. 본 논문에서는 기존에 제안된 다양한 장면분류 방법 중에서 직접 장면분류기법(Direct Scene Classification)을 사용하였다. 객체 비교기반의 장면분류를 수행하는 모델을 만들기 위해 작은 학습

데이터를 이용하여 모델을 학습시키기 위하여 전이 학습을 통한 학습을 진행하였다. 이렇게 학습된 CNN 모델을 기반으로 입력된 이미지의 장면분류가 진행된다.

IV. 실험 및 분석

제안하는 CD-SC의 성능 검증을 위해서 실내를 이동하여 촬영된 영상을 이용하여 기존에 제안된 방법 중 직접 장면분류기법(Direct Scene Classification, D-SC)과 장면변화 감지 기반 장면분류 기법(CNN-based Scene Classification with Change Detection, 이하 "CD-SC 방식"으로 표현)의 정확도와 CPU 사용률을 비교하였다. 실험에서 실내를 이동하는 모바일 로봇의 이동속도에서 장면 변화를 충분히 인지해 낼 수 있도록 영상에서 이미지 샘플링 주기를 1초로 설정하였다. 장면변화감지 모듈의 입력 이미지 크기는 125 X 125 픽셀로 실험을 진행하였다. 수집된 이미지는 장면변화감지 모듈로 입력되기 전에 125 X 125의 크기로 재조정(resize)하였다. 재조정된 이미지에 대해 노이즈를 제거하기 위하여 가우시안 블러를 적용하였으며, 이때 사용되는 커널(kernel)은 3 X 3으로 OpenCV에서 제공하는 cv2.GaussianBlur 함수를 사용하였다. 그리고 이미지의 특성을 유지하면서 이미지 크기를 줄이기 위하여 2 X 2 커널을 이용한 맥스 풀링(Max Pooling)을 수행하였다. 장면분류에 사용되는 모델은 적은 데이터로 높은 정확도를 가질 수 있도록 전이학습을 수행하였다.

객체 비교기반의 장면분류를 수행하는 모델을 만들기 위해 작은 학습데이터를 이용하여 모델을 학습시키기 위하여 전이학습을 통한 학습을 진행하였다. 전이학습은 MobileNetV3_Large 네트워크에 ImageNet을 학습데이터로 사용한 사전학습모델을 이용하였으며, 전이학습을 통해 사전학습모델의 마지막 계층인 FC(Fully Connected Layer)만 학습하였다. D-SC와 CD-SC는 모두 동일한 CNN 장면분류 모델을 사용하며, 신뢰도가 가장 높은 장면 5개를 유추하였다.

정확도를 분석하기 위해 영상의 장면마다 라벨링된 데이터를 기준으로 장면분류 결과와 일치하는지 판별하였다. 장면분류 결과 유추된 5개의 장면 중 신뢰도가 가장 높은 장면이 라벨링된 데이터와 일치하는 경우의 정확도를 Top1으로, 유추된 5개의 장면 중 하나라도 일치하는 경우를 Top5로 정의하였다. CD-SC는 다양한 영상에 대해 D-SC방법 대비 Top1, Top5 모두 0.8%이내의 오차율을 보였다. 이는

CD-SC가 이미지의 색과 밝기의 유사도를 기반으로 장면의 변화가 발생한 지점을 올바르게 특정하여 장면분류 하였기 때문이며 그 결과는 그림 2와 같다. 장면분류를 위해서는 이미지로부터 특징을 추출하고 추출된 특징을 기반으로 만들어진 객체를 기반으로 장면을 분류하기 때문에 고연산이 필요하다. CD-SC는 장면변화가 감지되었을 때만 장면분류를 수행하므로 장면변화가 없는 이미지에 대해서는 장면분류를 수행하지 않는다. 그 결과 CNN기반 고연산의 장면분류의 수행횟수가 감소하여 그림 3과 같이 D-SC대비 CPU의 사용률이 45.49% 감소하였다. 우리는 본 실험을 통하여 CD-SC는 주기적으로 수집된 이미지에 대해 장면변화를 감지하고 장면변화가 감지된 시점에만 장면분류를 수행하며 장면변화가 감지되지 않으면 장면분류를 수행하지 않음으로서 절반수준의 계산량으로 유사한 수준의 오차율을 가지는 것을 확인하였다. 표 2는 실험결과를 요약하였다.

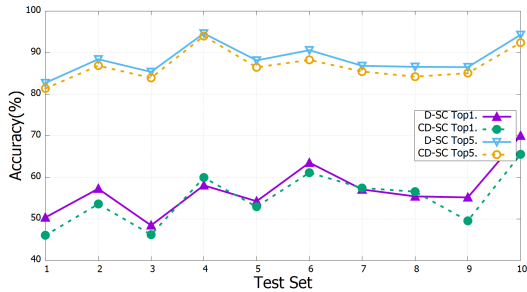


그림 2. D-SC 방법 대비 CD-SC 의 Top1. Top5. 정확도 비교
Fig. 2. Top 1. Top 5. Accuracy Comparison of CD-SC vs D-SC

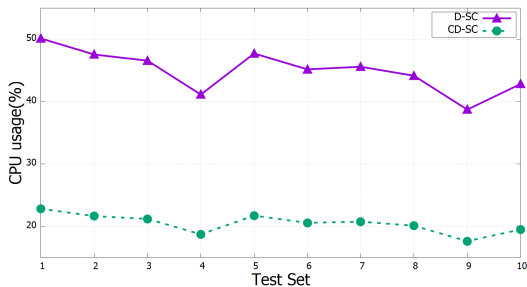


그림 3. D-SC 방법 대비 CD-SC 의 CPU 사용률 비교
Fig. 3. CPU usage Comparison of CD-SC vs D-SC

표 2. D-SC와 CD-SC의 평균 정확도 & 평균 CPU사용률
Table 2. Avg accuracy & Avg CPU usage of D-SC and CD-SC

	CPU usage	Accuracy	
		Top1.	Top5.
D-SC	44.95%	56.54%	88.08%
CD-SC	20.45%	55.79%	87.29%

V. 결 론

본 논문에서는 장면변화감지 기반 장면분류 기법 (CNN-based Scene Classification with Change Detection)을 제안하였다. CD-SC 는 주기적으로 수집된 이미지의 색(hue)와 밝기(brightness)를 기반으로 장면변화를 감지하고 장면이 변화된 시점에 장면분류를 수행한다. 그 결과 주기적으로 수집된 이미지를 모두 고 연산의 CNN 기반 직접 장면분류 하는 방법 (Direct Scene Classification, D-SC)대비 유사한 정확도를 유지하면서 고연산의 장면분류 횟수 감소를 통해 연산량이 크게 감소하였다. 따라서 CD-SC를 적용한 모바일 로봇은 기존에 제안된 방법대비 유사한 정확도를 가지며 적은 연산량을 통해 자원 소모를 줄이고 배터리 사용량도 줄어들게 되어 임무 수행에 필요한 자원을 확보할 수 있다.

References

- [1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Int. Conf. CVPR*, pp. 248-255, Miami, FL, USA, Jun. 2010, doi: 10.1109/cvpr.2009.5206848.
- [2] R. Wu, B. Wang, W. Wang, and Y. Yu, "Harvesting discriminative meta objects with deep CNN features for scene classification," in *Proc. IEEE Int. Conf. CVPR*, vol. 2015, pp. 1287-1295, Oct. 2015, doi: 10.1109/ICCV.2015.152.
- [3] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145-175, 2001, doi:10.1023/A:1011139631724.
- [4] A. M. Vogel, D. R. Lennon, S. Gray, E. Farrell, and P. Anderson, "Registered nurse

- assessment and treatment of skin sepsis in New Zealand schools: The development of protocols," *N. Z. Med. J.*, vol. 126, no. 1380, pp. 27-38, 2013.
- [5] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the Fisher kernel for large-scale image classification," *LNCS (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6314 LNCS, no. PART 4, pp. 143-156, 2010, doi: 10.1007/978-3-642-15561-1_11.
- [6] T. F. Gonzalez, "ImageNet classification with deep convolutional neural networks," *Handb. Approx. Algorithms Metaheuristics*, pp. 1-1432, 2007, doi: 10.1201/9781420010749.
- [7] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, and L. Zhang, "Preparation of novel high copper ions removal membranes by embedding organosilane-functionalized multi-walled carbon nanotube," *J. Chem. Technol. Biotechnol.*, vol. 91, no. 8, pp. 2322-2330, 2016, doi: 10.1002/jctb.4820.
- [8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1-14, 2015.
- [9] A. Quattoni and A. Torralba, "Recognizing indoor scenes," *2009 IEEE Conf. CVPR*, pp. 413-420, 2010, doi: 10.1109/cvpr.2009.5206537.
- [10] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, "SUN database: Large-scale scene recognition from abbey to zoo," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3485-3492, 2010, doi: 10.1109/CVPR.2010.5539970.
- [11] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1452-1464, 2018, doi:10.1109/TPAMI.2017.2723009.
- [12] Z. Zuo, et al., "Learning contextual dependence with convolutional hierarchical recurrent neural networks," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 2983-2996, 2016, doi:10.1109/TIP.2016.2548241.
- [13] D. Yoo, S. Park, J. Y. Lee, and K. Inso, "Multi-scale pyramid pooling for deep convolutional representation," *IEEE Comput. Soc. Conf. CVPRW*, vol. 2015-Oct., pp. 71-80, 2015, doi: 10.1109/CVPRW.2015.7301274.
- [14] S. Liu, G. Tian, and Y. Xu, "A novel scene classification model combining ResNet based transfer learning and data augmentation with a filter," *Neurocomputing*, vol. 338, pp. 191-206, 2019, doi:10.1016/j.neucom.2019.01.090.
- [15] P. Tang, H. Wang, and S. Kwong, "G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition," *Neurocomputing*, vol. 225, pp. 188-197, 2017, doi:10.1016/j.neucom.2016.11.023.
- [16] X. Cheng, J. Lu, J. Feng, B. Yuan, and J. Zhou, "Scene recognition with objectness," *Pattern Recognit.*, vol. 74, pp. 474-487, 2018, doi:10.1016/j.patcog.2017.09.025.
- [17] C. Y. Xu, D. Yang, and Y. Zhang, "Image semantics recognition based on ontology and description logics," *1st Int. Conf. Multimed. Inf. Netw. Secur. MINES 2009*, vol. 2, pp. 33-36, 2009, doi:10.1109/MINES.2009.19.
- [18] S. Dasiopoulou, I. Kompatsiaris, and M. G. Strintzis, "Applying fuzzy DLs in the extraction of image semantics," *LNCS (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5880 LNCS, pp. 105-132, 2009, doi: 10.1007/978-3-642-10562-3_4.
- [19] ITU, "Recommendation BT.601-7," vol. 7, 2011, [Online]. Available: <http://www.itu.int/ITU-R/go/patents/en>.

신 등 훈 (Dong-Hun Shin)



2016년 2월 : 단국대학교 소프트
웨어공학과 학사
2021년 2월~현재 : 세종대학교
정보통신공학과 석사과정
<관심분야> 사물인터넷, 자율지
능시스템, 스마트시티 플랫폼,
디지털트윈 및 메타버스

[ORCID:0000-0002-0685-6067]

김 재 호 (Jae-Ho Kim)



2017년 : 연세대학교 전기전자
공학과 박사
2000년~2020년 : 한국전자기술
연구소 자율지능IoT연구센터
장
2020년~현재 : 세종대학교 전 자
정보통신공학과 교수

2019년~현재 : TTA 사물인터넷/스마트시티 플랫폼 포
준그룹(PG1001)의장

2021년~현재 : ITRC 메타버스 자율트윈 연구센터 센
터장

<관심분야> 사물인터넷, 자율지능시스템, 스마트시티
플랫폼, 디지털트윈 및 메타버스

[ORCID:0000-0001-6597-7988]