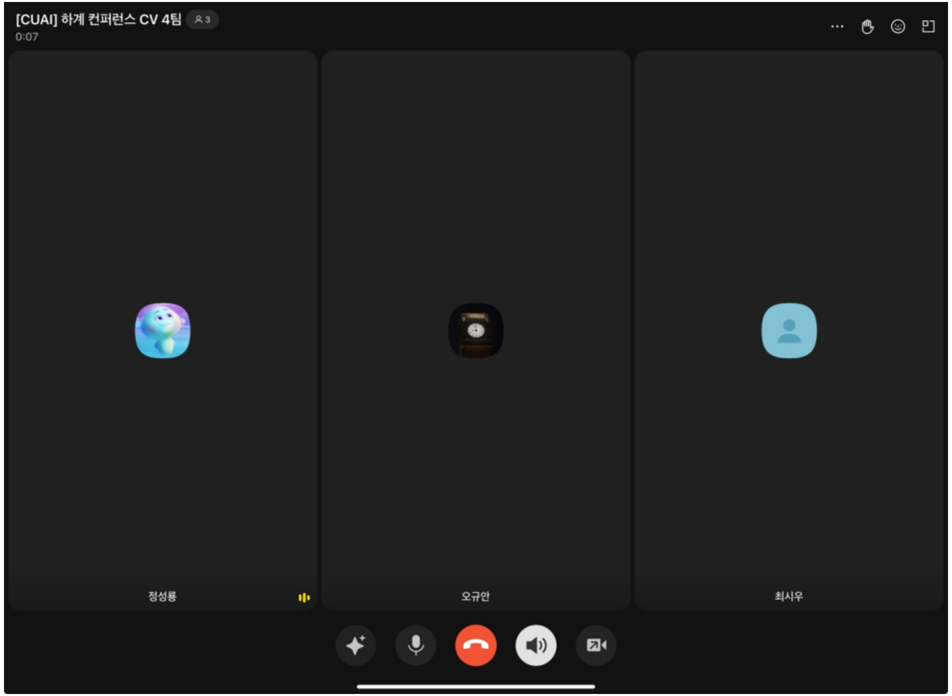
CUAI Advanced Multimodal 02팀

2024.11.05

발표자 : 최시우

스터디원 소개 및 만남 인증



스터디원 1 : 오규안 (AI)

스터디원 2 : 최시우 (AI)

스터디원 3 : 정성룡 (AI)

Introduction

낮은 조도에서 촬영된 이미지는 세부 정보가 손실되고

노이즈가 증가하여 성능이 저하될 수 있음.

Introduction

CLAHE와 같은 저조도 이미지 향상 기법이 사용되지만,

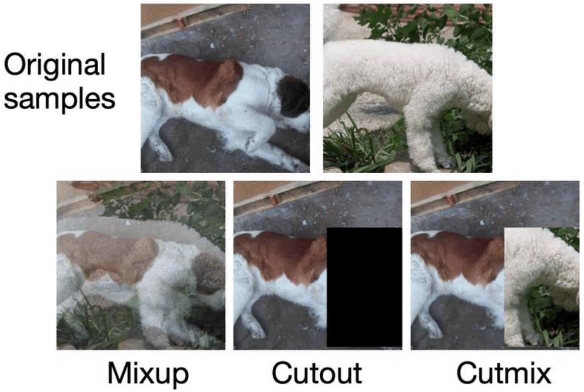
자율주행 등 실시간 처리가 중요한 응용에서는 속도 저하가 문제임.

Introduction

CLAHE + CutMix

Methodology

CutMix



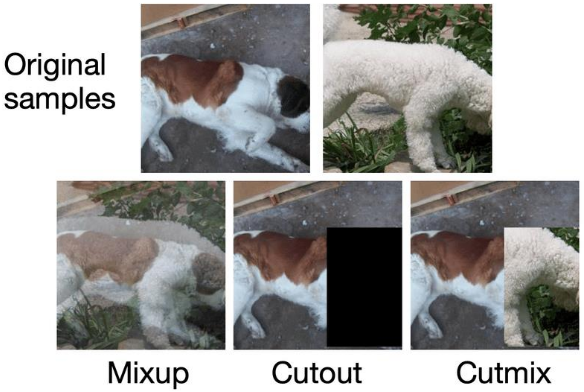
CLAHE

CutMix는 이미지를 랜덤하게 잘라 결합하여 모델의 일반화 성능을 높이는 기법이며,

CLAHE는 지역적인 대비 향상으로 저조도에서 중요한 세부 정보를 보존함.

Methodology

CutMix

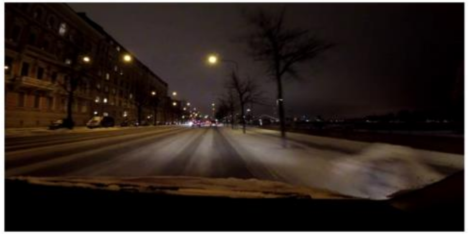


CLAHE

CLAHE는 추론 시점에서도 전처리 시간이 추가로 필요함.

-> 실시간 처리에 제약이 있음.

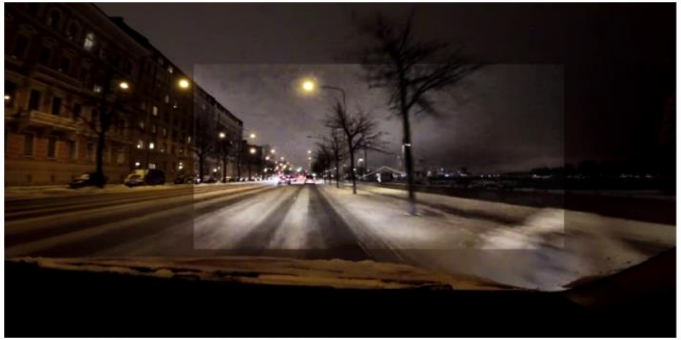
Methodology



원본 이미지

CLAHE

ClaheMIx

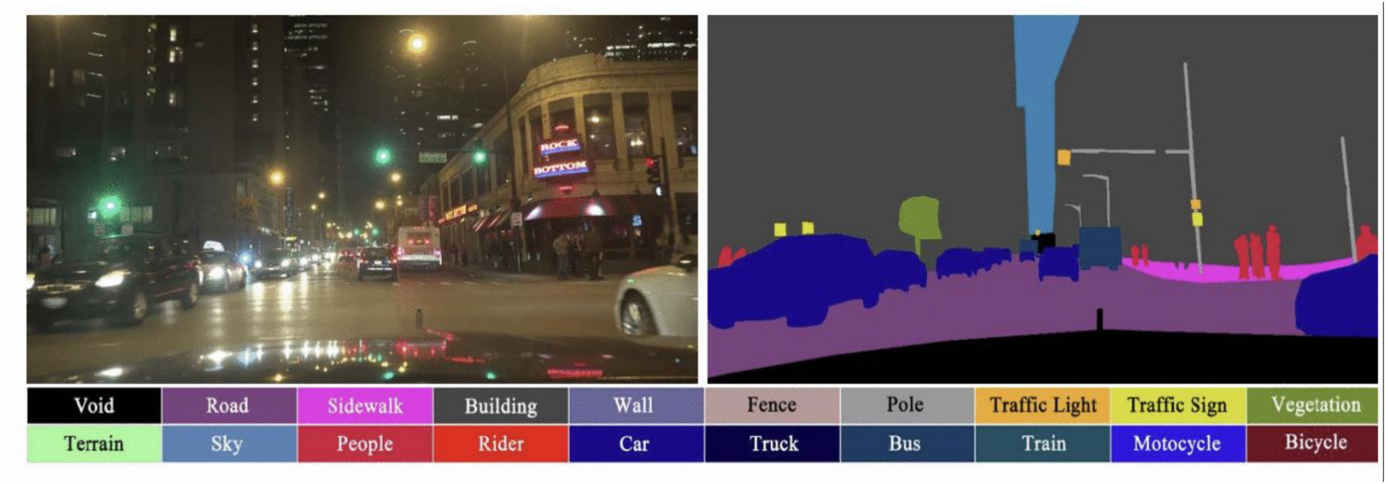


원본 이미지와 CLAHE를 적용한 이미지에서 동일한 위치에 있는 패치를 선택한 후,

원본 이미지의 해당 패치 부분을 CLAHE가 적용된 이미지의 패치로 대체함.

Experimental Setup

NightCity

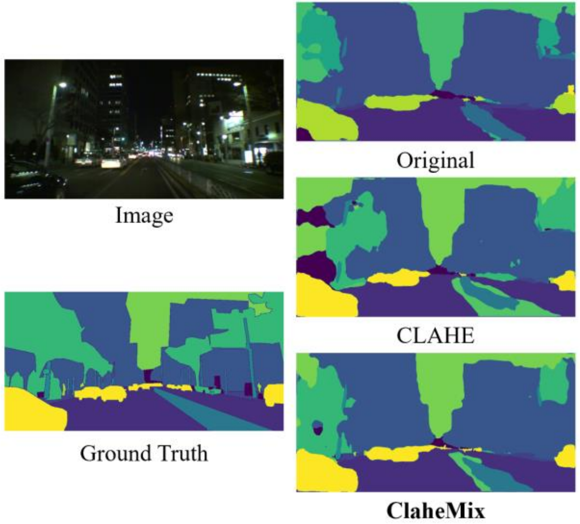


NightCity 데이터셋은 다양한 조명 조건에서 촬영된 야간 도시 이미지로,

저조도 환경에서 이미지 인식 및 분할 성능을 평가하기에 적합함.

Results

		mIoU	mAP	Pixel Accuracy	Preprocessing Time
Original		33.7	20.2	78.1	0.13ms
CLAHE		36.4 (+2.7)	21.6 (+1.4)	80.3 (+2.2)	167.82ms
ClaheMix		**37.1** (+3.4)	**22.0** (+1.8)	**80.9** (+2.8)	0.13ms



ClaheMix를 활용했을 때 mIoU, mAP, Pixel Accuracy 모두 향상되었으며,

추론 시 추가 전처리가 필요하지 않아 전처리 시간을 줄일 수 있었음.

References

◊ VIII.5

Contrast Limited Adaptive Histogram Equalization

Karel Zuiderveld
Computer Vision Research Group
Utrecht University
Utrecht, The Netherlands
karel@cv.ruu.nl

This Gem describes a contrast enhancement technique called *adaptive histogram equalization*, AHE for short, and an improved version of AHE, named *contrast limited adaptive histogram equalization*, CLAHE, that both overcome the limitations of standard histogram equalization. CLAHE was originally developed for medical imaging and has proven to be successful for enhancement of low-contrast images such as portal films (Rosenman *et al.* 1993).

◊ Introduction ◊

Probably the most used image processing function is contrast enhancement with a lookup table, a 1-to-1 pixel transform as described in (Jain 1989). When an image has poor contrast, the use of an appropriate mapping function (usually a linear ramp) often results in an improved image.

The mapping function can also be non-linear; a well-known example is gamma correction. Another non-linear technique is *histogram equalization*; it is based on the assumption that a good gray-level assignment scheme should depend on the frequency distribution (histogram) of image gray levels. As the number of pixels in a certain class of gray levels increases, one likes to assign a larger part of the available output gray

CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features

Sangdoo Yun[1] Dongyoon Han[1] Seong Joon Oh[2] Sanghyuk Chun[1]
Junsuk Choe[1,3] Youngjoon Yoo[1]

[1]Clova AI Research, NAVER Corp.
[2]Clova AI Research, LINE Plus Corp.
[3]Yonsei University

arXiv:1905.04899v2 [cs.CV] 7 Aug 2019

Abstract

Regional dropout strategies have been proposed to enhance the performance of convolutional neural network classifiers. They have proved to be effective for guiding the model to attend on less discriminative parts of objects (e.g. leg as opposed to head of a person), thereby letting the network generalize better and have better object localization capabilities. On the other hand, current methods for regional dropout remove informative pixels on training images by overlaying a patch of either black pixels or random noise. Such removal is not desirable because it leads to information loss and inefficiency during training. We therefore propose the **CutMix** *augmentation strategy: patches are cut and pasted among training images where the ground truth labels are also mixed proportionally to the area of the patches. By making efficient use of training pixels and retaining the regulariza-*

	ResNet-50	Mixup [48]	Cutout [3]	CutMix
Image				
Label	Dog 1.0	Dog 0.5 Cat 0.5	Dog 1.0	Dog 0.6 Cat 0.4
ImageNet Cls (%)	76.3 (+0.0)	77.4 (+1.1)	77.1 (+0.8)	**78.6** (+2.3)
ImageNet Loc (%)	46.3 (+0.0)	45.8 (-0.5)	46.7 (+0.4)	**47.3** (+1.0)
Pascal VOC Det (mAP)	75.6 (+0.0)	73.9 (-1.7)	75.1 (-0.5)	**76.7** (+1.1)

Table 1: Overview of the results of Mixup, Cutout, and our CutMix on ImageNet classification, ImageNet localization, and Pascal VOC 07 detection (transfer learning with

Night-time Scene Parsing with a Large Real Dataset

Xin Tan, Ke Xu, Ying Cao, Yiheng Zhang, Lizhuang Ma, and Rynson W.H. Lau

arXiv:2003.06883v3 [cs.CV] 1 Apr 2022

***Abstract*—Although huge progress has been made on scene analysis in recent years, most existing works assume the input images to be in day-time with good lighting conditions. In this work, we aim to address the night-time scene parsing (NTSP) problem, which has two main challenges: 1) labeled night-time data are scarce, and 2) over- and under-exposures may co-occur in the input night-time images and are not explicitly modeled in existing pipelines. To tackle the scarcity of night-time data, we collect a novel labeled dataset, named *NightCity*, of 4,297 real night-time images with ground truth pixel-level semantic annotations. To our knowledge, NightCity is the largest dataset for NTSP. In addition, we also propose an exposure-aware framework to address the NTSP problem through augmenting the segmentation process with explicitly learned exposure features. Extensive experiments show that training on NightCity can significantly improve NTSP performances and that our exposure-aware model outperforms the state-of-the-art methods, yielding top performances on our dataset as well as existing datasets.**

***Index Terms*—Autonomous Driving, Night-time Vision, Scene Analysis, Adverse Conditions.**

I. Introduction

SCENE Parsing is an important computer vision task for many applications, such as human parsing [1], image editing [2] and autonomous driving [3]. Although a lot of methods have been proposed, they mainly focus on day-time scenes. However, as night time and day time cover roughly about 50% of the time each (averaged over a year), it is equally important to build vision systems that perform well at night time, particularly for autonomous driving at night. In this paper, we address the night-time scene parsing (NTSP) problem.

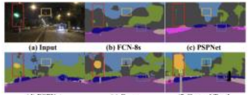


Fig. 1. Night-time scene analysis. (a) shows an input night-time image, and parsing results from state-of-the-art methods: (b) FCN-8s [4], (c) PSPNet [5], and (d) ESPNet [6]. We also show the result by our model (e), and the ground truth (f). The yellow and blue boxes highlight under- and over-exposed regions, respectively. The red box highlights the region with a mixture of under- and over-exposures.

due to under-exposure, causing it to be difficult to detect. Second, the texture and structure of the cars highlighted by the blue box are corrupted due to over-exposure, causing them to be difficult to segment correctly. Third, the traffic light highlighted by the red box is difficult to be detected or segmented correctly due to a mixture of over-/under-exposures.

There are two main challenges to the NTSP problem. First, large-scale *labeled* datasets of *night-time* scenes are not available. Existing large datasets for scene parsing mainly contain day-time images, with few or no night-time images [7], [8], [9]. Models trained on these datasets do not generalize well to the complexity of night-time scenes. Second, existing methods do not explicitly model over- and under-exposures, as they are primarily developed for day-time scenes. However,

감사합니다