제목

1st Author Name a, 2nd Author Name b, 3rd Author Name c

a 1st Author Affiliation, City, Country, e-mail address

b 2nd Author Affiliation, City, Country, e-mail address

c 3rd Author Affiliation, City, Country, e-mail address

**Abstract**

~~단일 이미지 스모크 제거 기술은 화재현장, 공사현장(dust)에서 중요한 역할을 한다. 기존의 desmoke 기술은 불완전 투명도를 연기에 대한 연구가 이루어져 contrast를 증가시키는 방법등으로 사실상 불투명한 연기로 인해 background의 정보를 잃은 이미지는 inpating과 비슷한 기술이 이루어져야한다. 허나 inpainting 기술과는 다르게 벡터와 모양이 일정치않고 불규칙적인 투명/불투명 noise는 background에 대한 정보를 일부 가지고 있어 기존의 Inpainting 알고리즘과 증가된 Contrast 알고리즘은 이에 맞지않아 새로운 알고리즘이 요구된다. 이 논문은 연기로 인해 손실된 background 정보를 재생성 하는 문제를 다룬다. 하나의 이미지에 Smoke와 Background 두 공간을 분리하기 위해 2-branch의 U-net 구조를 갖는 Interaction Generative Adversarial Network 를 제안한다. Generator의 2 Branch는 서로에게 정보를 공유하여 상호작용하고, Discriminator통해 L1 Loss와 Gan Loss로 feed-back하고, Identify loss 를 추가하여 Sharp연산과 groundtruth contrast 싱크한다. 실험 결과는 AI를 활용한 자율주행, SmartFactory, 공사 현장, 재난 현장등 다양한 산업분야에 benefit을 보여준다.~~

**Keywords**

Add about 5 keywords or phrases separated by semicolons for use in indexing this paper; these may also be used to identify appropriate reviewers. For the Keywords heading, use the Normal style and bold it.

1. Introduction

The renaissance of artificial intelligence has significantly benefited humans and computers in many areas. Especially image processing based on CNN is being utilized in various fields. For computers, noise poses great difficulties in image processing such as object recognition, image encoding/decoding, segmentation. Noise such as smoke,rain,foggy makes it difficult for both humans and computers to grasp information about images.[0] This causes a lot of human and property damage. In fact, according to the U.S. National Statistical Office, 28% of all accidents are affected by bad weather.[1] In addition, camera-based autonomous driving models do not function due to severe performance degradation when there is a lot of interference in the input image, and several image processes such as motion recognition are affected by image quality. Therefore, in image processing, an algorithm capable of corresponding to various noises is required.

Smoke는 불규칙적인(irregular vector) 벡터와 모양이 일정치(unconstant shape) 않고 투명/불투명도 또한 불규칙적으로 존재하는 기체의 특성을 갖는 smoke noise제거는 선행 연구가 부족하고 도전적인 문제이다. 기존의 smoke remove는 defogging, dehazing과 같은 기술로 대체 되었다.[a] Haze는 기본적으로 물체로부터 반사된 후 대기에 의해 일부 손실된 빛과 태양 등 주변 광원으로부터 대기에 반사되어 함께 시야에 들어오는 대기 산란광(Airlight)이 섞임으로 인해 발생한다. Kaiming He el al[2] 은 대부분 R,G,B 세 채널 중에서 적어도 한 채널의 명도의 값은 매우 낮은 경향이 있다는 것을 발견하여 dark channel prior 을 이용하여 다양한 아웃도어 이미지들에 존재하는 Haze를 제거하였으나 검은 smoke를 포함하고 거리와 Smoke의 투명도는 연관성이 적어 본 연구는 이에 적합하지 않다. 또한 MSE와 같은 pixel-wise average loss는 over-smooth를 발생시키고, Perceptual capture에 취약하다.[3,4,5,6] 일부 background의 정보를 모두 손실한 pixel은 복구 시키는데 어려움을 갖고 있기 때문에 smoke noise를 제거하는데 적합하지 않다. 이러한 Smoke의 특성은 불규칙적인 Background 정보를 손실 발생시켜 가시성을 심각하게 저하 시킬뿐만 아니라 smoke remove problem에 있어 어려운 문제이다.

기존의 CNN의 Dehazing, Defogging, deraining 등과 같은 기법과 대조적으로Small damaged pixel, Huge damaged pixel 두 가지 형태의 노이즈를 갖는 smoke의 특성 때문에 dehazing의 특성과 inpainting의 특성을 모두 cover 할 수 있는 새로운 네트워크가 요구된다. 네트워크는 Encoding,feature transformation, Decoding 아키텍쳐를 갖는다. 또한 Res-Block 모듈에 multe-scale dilated layer를 추가하여 복잡도와 성능을 개선하고 여러 스케일에서의 이미지의 contextual information을 종합하여 prediction 성능을 향상 시킨다[10][11]. pre-trained VGG 16 모델을 통한 perceptive loss로 feature domain을 유지시키고,[12] background image의 정보를 SSIM Loss를 통해 이미지의 contrast와 luminance, structure를 reconstruction 하여[13] 다양한 smoke noise에 대응 가능하도록 한다. 본 논문은 Dehaze 특성과 Inpainting 특성을 모두 갖는 네트워크 제안하고 여러 환경의 smoke noise를 테스트하여 Desmoke에 강력한 네트워크임을 증명한다.

**2. 관련 연구**

**2.1 GAN(Unet-GAN)**

생성적 적대 네트워크(GAN)에 의해 생성된 합성 이미지의 품질은 최근 크게 향상되었다.[21,22] 기존의 지도 학습 방식과 다르게 경쟁하는 과정을 통해 결과를 도출하는 새로운 프레음 워크 갖는 방식으로 adversarial nets 프레임 워크의 컨셉은 경쟁 으로, discriminator(D) 모델과 generative 모델(G)이 존재한다. D는 G가 실제 training data인지 허구의 생성된 데이터를 판별하고, G가 생성한 허구의 데이터와 training 데이터를 판별하지 못하도록 G는 training 데이터와 근사한 모집단으로 데이터를 생성한다.[18]

MSE와 같은 손실 함수를 사용하여 loss를 최소화 하면 pixel-wise average을 구하여 over-smooth 시키고 poor perceptual quality를 야기한다.[19] The authors of Denton et al. tackle this problem by employing generative adversarial networks (GANs) [18] for the application of image generation. Yu et al.augment pixel-wise MSE loss with a discriminator loss to train a super-resolution network.[20] 이를통해 fake이미지를 real 이미지와 근사하게 만들어 실제와 가깝게 reconstruction 하는데 benefit을 준다.

GAN의 문제는 판별자가 비정상 환경(non-stationary)에서 학습해야 할 때 증폭된다. 합성 샘플의 분포는 생성기가 훈련을 통해 지속적으로 변화함에 따라 변화하고 이전 작업을 잊어버리기 쉽다.[23] 이러한 Generator로 생성된 이미지는 종종 얼룩덜룩한 로컬 스트럭쳐를 생성하거나,[24] 일관성이 없는 기하학적 구조 패턴을 가진 이미지를 만들어 내기도 합니다.[25] 이러한 문제를 해결하기 위해 Anna Khoreva et al 은 GAN의 남아있는 가장 중요한 문제중 하나인 실제 이미지로부터 구별이 불가능하며 물체의 모양이나 texture가 globally, locally 일관성 있는 이미지를 만드는 것을 지적하였고, 이를 해결 하기 위해 U-Net의 기반으로한 Discriminator를 제안하였다. U-Net 모델은 전체적인 모양이 U형태를 이루는 네트워크로 Encoder-Decoder 구조를 이루고 있다. Encoder에서는 context를 캡처하고 이와 대칭되는 구조를 갖는 Decoder에서 localization을 가능케하는 구조이다. (유넷논문)은 deep convolution network는 많은 visual recognition 작업에서 매우 좋은 성능을 보였지만, training set의 크기와 고려할 네트워크의 크기 때문에 그 성공은 제한적임을 지적하였다. U-net의 핵심은 Encoder와 Decoder 사이에 Copy and Crop 하여 Concat하는 구조가 핵심적이다. 이러한 이점을 활용하여 Discriminator Encoder 에서 global classification을 피드백하고(이전의 GAN과 동일한), Decoder 과정을 통해 Local pixel을 피드백한다..[유넷간]

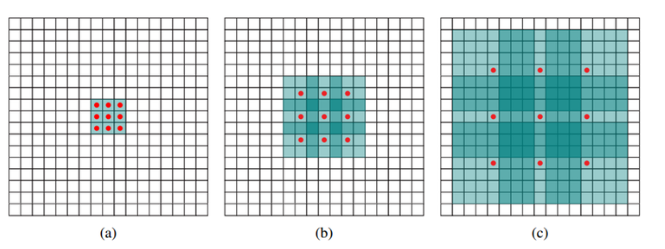
**2.3 Resdual Block**

**VGG16, 19는 3\*3 filter를 이용해 레이어의 깊이를 늘려 우수한 성능을 보였다.[35] 특히 CNN 계열에서 레이어가 깊어지면 특징이 더 풍부해지기 때문에, CV 분야에서의 네트워크 깊이는 많은 장점을 가져와주었다. 사람들은 따라서 VGG에 따라 깊이가 깊어지면 더 높은 성능을 낼 것이라고 생각하였다.**

**하지만 점차 레이어가 깊어지면 Vanishing Gradient와 같은 문제가 발생함을 인지하였으며, 레이어의 깊이가 깊어질수록, input층에 가까운 은닉층이 희미해지기에 학습에 도움이 되지 않을수도 있음을 주장하였다.[36]**

**ResNet is made up of residual blocks which contain additional skip connections. There is a direct link that bypasses certain layers in between. This is what known as the 'skip connection', and it is the base of residual blocks. The outcome of the layer is no anymore the very same because of this additional skip connection [33], [34]. ResNet's skip connections eliminate the issue of the vanishing gradient in deep neural network structures by permitting the gradient to travel along an additional bypass path [33], [34]. These connections also help the model by helping it to understand the identity functions and confirm that the upper layer performs at least as good as the bottom layer even in the worst case.**

**2.4 dilated Convolution**

****

FCN 개발자들이 발표한 fully convolutional network의 개념은 다른 많은 연구자들에게 큰 자극이 되었다.

그 중 Fisher Yu는 FCN에 대한 분석 및 구조를 dilated convolution의 형태로 변경으로 FCN의 성능을 향상 시켰다. 공간적 정보의 손실이 있는 것을 upsampling 하면 해상도가 떨어진다. 하지만 dilated convolution의 그림을 보면 receptive field를 크게 가져가면서 convolution을 하면 정보의 손실을 최대화하면서 해상도는 큰 output을 얻을 수 있다.

**2.5 Inpainting**

inpainting은 손상된 이미지를 복원 시키는 기술 중 하나로 sequential and machine-learning based을 포함한 여러 기술들이 존재한다.[7] Pioneering works on digital image inpainting proposed diffusion-based sequential methods where the missing region is filled by a smooth continuation of the image with information surrounding the missing area [8], [9]. Shift-net은 Copy-and-Paste 방식을 활용하여 이미지의 디테일을 살리고, 시프트 연산이 있는 제안된 시프트 연결 계층은 네트워크가 누락된 부분 이외의 가장 가까운 이웃이 제공하는 정보를 효과적으로 차용하여 생성된 부분의 글로벌 의미 구조와 국소 질감 세부사항을 모두 세분화할 수 있도록 한다.[14] DeepFill v1은 dilated convolution을 이용하여 inside missing area를 외부의 모든 feature 조합을 사용하여 누락된 feature을 generate하였다.[15] Generation Multi-column Convolutional Neural Networks 는 multi-scale dilated convolution을 제안하였다. 이는 네트워크와 누락된 영역 밖에서 가장 가까운 이웃을 찾는 Id-MRF loss를 추가하여 이웃을 참조하여 내부를 inpainting하는데 효과적임을 증명하였다.[16] 또한 Dilated Residual Networks 에서는 위의 dilated convolution과 residual block을 활용하여 segmentation을 진행하여 dilated Residual Network가 capturing feature에 효과적임을 증명했다.[17] 위와 같은 인페인팅적인 특징을 살려 본 논문 네트워크에 적용하였다.

**2.5 Dehaze**

Single image dehazing means restoring the corresponding clean image from the given hazy image. Usually, hazy images exhibit blurring, low contrast, color distortion, and other poor visual quality issues. Taking those degraded images as inputs impedes the real performance of many subsequent visual tasks. Generally, image dehazing acts as a prerequisite and can affect a series of subsequent applications, such as object detection [40], content-based image retrieval [41], [42], and so on [43], [44], [45]. Currently, image dehazing has become a hot research topic and has received increasing attention. 최근 이러한 디헤이즈를 위해 많은 연구들이 진행되었다.[2][46][47] These methods are generally segregated into two classes: prior-based methods and learning-based methods.

These priors can be applied to draw a clear distinction between hazy images and their corresponding haze-free images, such as dark-channel prior [2], haze-line prior [47]. Although prior-based methods work well in some scenarios, there are others in which they are ineffective. For example, the dark channels prior does not work for some areas closer to atmospheric light, such as white scenes, where it is difficult to obtain an accurate transmission map.[48]

**3. Method**

이 절의 First part는 데이터 설명과 모델의 아키텍쳐에 대해 설명하고, 이후 Loss Function에 대해 설명한다.

3.1데이터 설명

Hsiang-Yin Cheng at al 은 Smoke100K 데이터셋을 제안하였다.[28] 이 데이터셋은 LabelMe dataset, NYU dataset[26][27]을 기반으로 100k의 synthesized smoke image, smoke-free image, smoke mask 이미지를 224x224x3 크기로 변환 하였다. 스모크 마스크 이미지의 density 3가지의 단계로 나누어 Low, Middle, High로 나누었으며, The amount of each base smoke mask image made from phothoshap is 220, 400 and 300. 이러한 smoke mask는 smoke-free 이미지에 다른 앵글, 모양 등으로 변화를 주어 합성되는데 수식은 아래를 따른다.



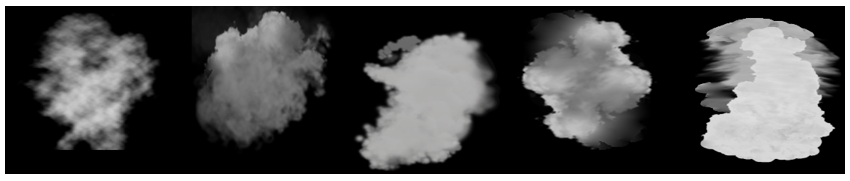
where xi,c, yi,c, and zi,c denote the ith synthesized smokeimage, smoke-free image, and smoke mask, respectively; andzi,α denotes the alpha channel for transparency of smokemask.



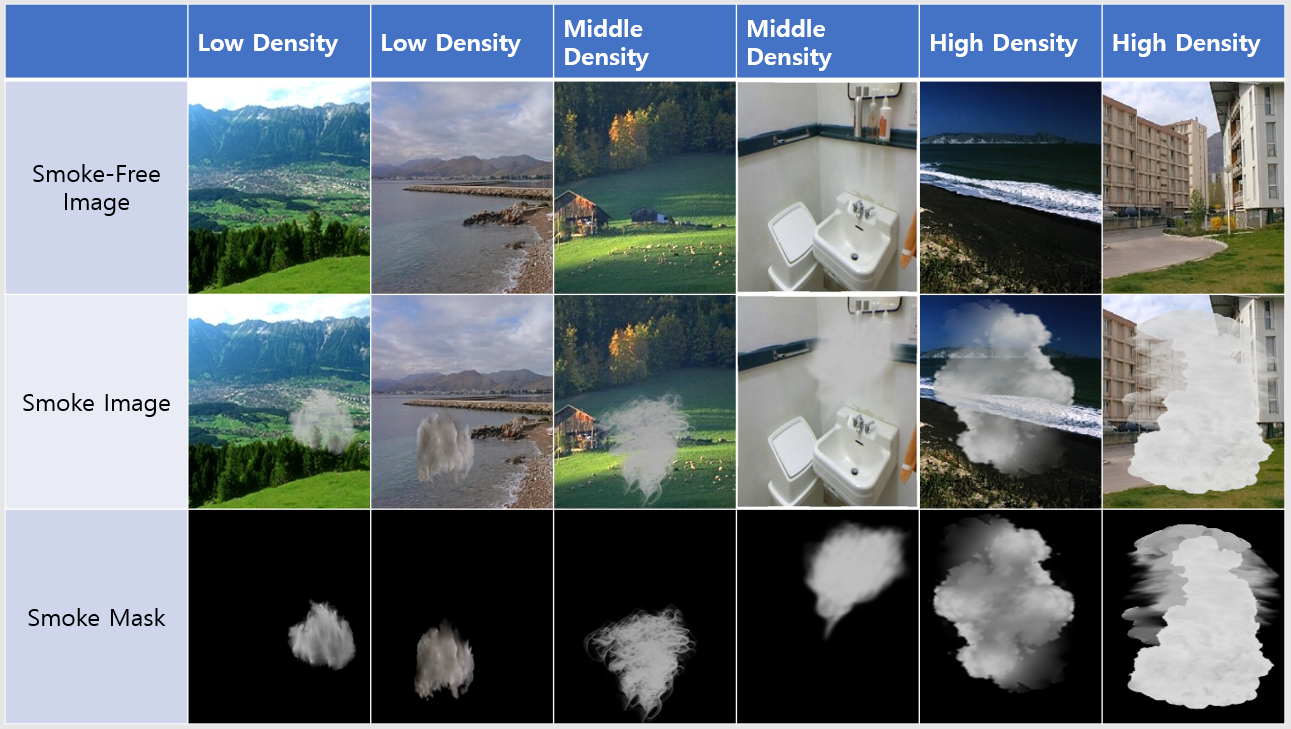
[그림] a. Smoke-Free Images



[그림] b. Smoke Images



[그림] c. Smoke Mask Images

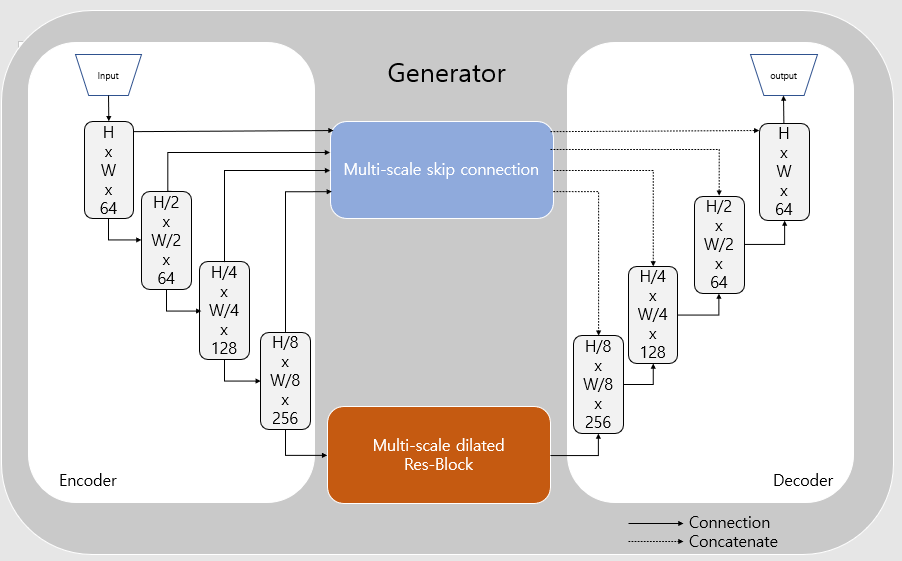


[그림] Smoke-Free Image, Smoke Image and Smoke Mask for each Density

위 데이터 중 High density smoke image를 활용하여 MobilenetV2,ResNet50을 훈련시켜 각각 0.88, 0.92 IoU스코어 갖는 Bounding box를 예측하여 의미 있는 데이터 셋임을 증명하였다(Smoke100k dataset is able to be exploited to train a discriminative model for more accurate smoke detection).[28]

**3.2모델설명**

**Generator**

****

**[그림] Generator Model Architecture**

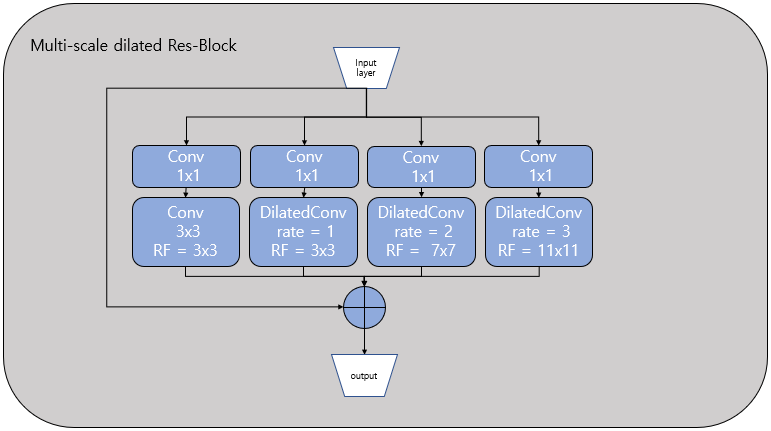
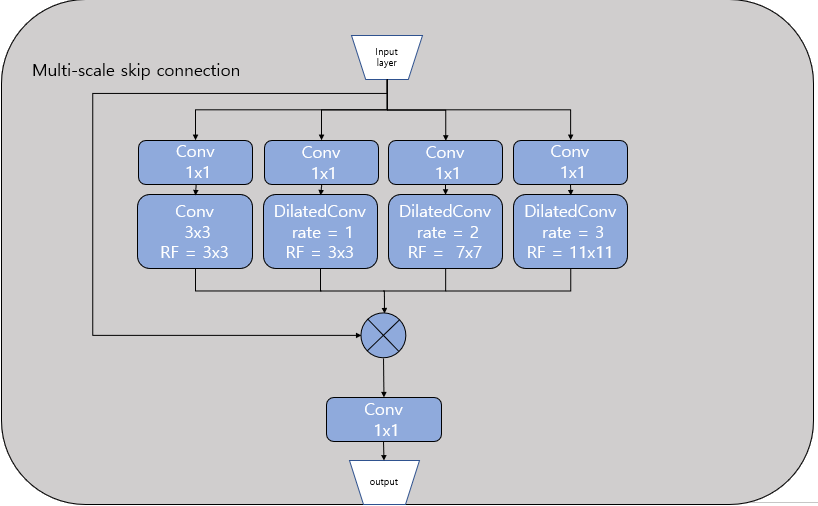
As Shown Fig Generator consists of Encoder-Decoder Structure involving Skip-Connections as U-net.

each encoder, decoder consist of 3 down(up)sampling, 4 feature extracting layer. common skip connection replaced Multi-scale skip connection(MSSC), common Residual Block replaced Multi-scale dilated Residual Block(MSDRB).

MSDRB reconstucture feature, MSSC benefit to using before reconstructuction feature Details generator model as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Generator | | |  |
| Layer | Filter | kernel | Input  Layer |
| Input  (256\*256\*3) |  |  |  |
| Conv1 (strides = 0)  (256\*256\*32) | 32 | 7\*7 | input |
| Conv2  (128\*128\*64) | 64 | 3\*3 | conv1 |
| Conv3  (64\*64\*128) | 128 | 3\*3 | conv2 |
| Conv4  (32\*32\*256) | 256 | 3\*3 | conv3 |
| MSDRB \* 8  (32\*32\*256) | 256 | multi-scale | - |
| Conv Transpose1 (strides = 0)  (32\*32\*256) | 256 | 3\*3 | MSDRB |
| Concatenate1 | - | - | Conv Transpose1,  Conv4 |
| Conv Transpose2  (64\*64\*128) | 128 | 3\*3 | Concatenate1 |
| Concatenate2 | - | - | Conv Transpose2,  Conv3 |
| Conv Transpose3  (128\*128\*64) | - | - | Concatenate2 |
| Concatenate3 | - | - | Conv Transpose3,  Conv2 |
| Conv Transpose4  (256\*256\*32) | 32 | 3\*3 | Concatenate2 |
| Concatenate3 | - | - | Conv Transpose4,  Conv1 |

모든 Conv,ConvTranspose 는 Reflaction padding, batch normalization과 relu activation이 생략 되었다.



[그림] MSSC, MSDRB Module Architecture

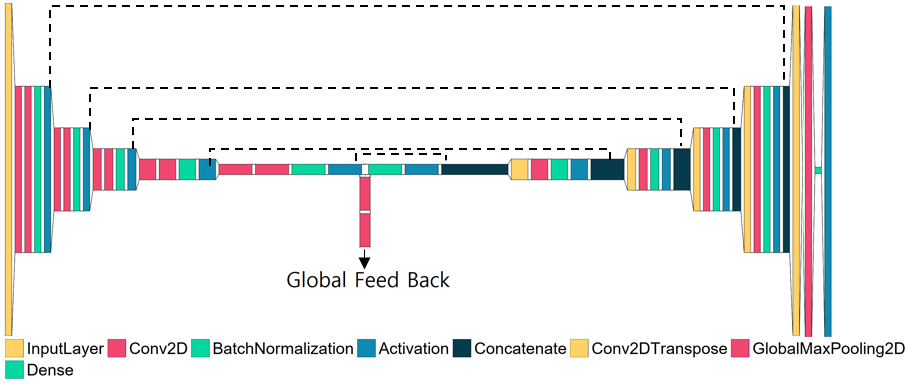
Figure Shown MSSC,MSDRB Module Architecture

두 모델은 Previous layer 뒤에 1x1 convolution이 위치하는데 구글에서 발표한 Inception 계통의 Network에서는 1x1 Convolution을 통해 유의미하게 연산량을 줄였습니다. 또한 채널(Channel) 또는 공간적(Spatial) 특성 파악을 위해 1x1 Convolution 방법 사용이 유의미하다는 것을 증명한 바 있다.[21]

이러한 특성을 통해 통해 연산량을 줄이고 그 후에 각 kernel size별로 연산 후 concatenate, Add layer를 통해 다음 레이어로 출력된다. 두 모듈은 multi-scale로 damaged area reconstruction.

VishalVenkatesh et al은 Multi-Scale residual block(MSRB)를 통해 복강경 수술 사진에서 연기를 제거하는데 탁월한 효과가 있음을 증명하였다. MSRB는 의미분할[[30],](https://www.sciencedirect.com/science/article/pii/S0010482520302274#b33)영상 초해상도[[31]](https://www.sciencedirect.com/science/article/pii/S0010482520302274#b34)및 영상 디블러링[[32]](https://www.sciencedirect.com/science/article/pii/S0010482520302274#b35)과 같은 작업에서 멀티 스케일 특징 추출의 능력에 영감을 받아 만들어진 블럭으로 MSRB의 접근법은 여러 커널 크기를 사용하여 순차적이고 병렬적인 일련의 컨볼루션에 의해 기능 추출을 수행하고 잔여 학습을 수행함으로써 더 효율적이다.(The approach presented in block performs feature extraction by sequential and parallel set of convolutions using multiple kernel sizes and makes it more efficient by performing residual learning.)[29]. 본 연구에서는 MSRB와 비슷하나 순차적인 과정을 생략하고 멀티 스케일의 특징점은 갖되 내부적인 메모리 사용량이 적은 dilated convolution을 적용하여 Multi-Scale-Skip-Connetion(MSSC)로 활용하고 인코딩 과정에서 생성된 이미지가 보다 넓은 스케일로 디코딩 과정에 영향을 미칠 수 있도록 하였다.

**Discriminator**



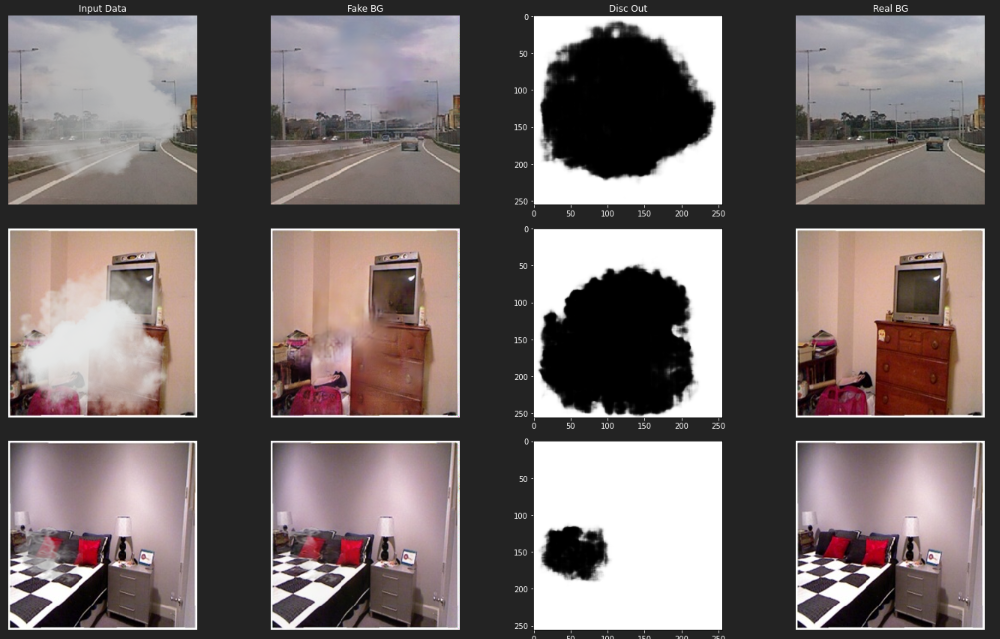
**[그림] Unet-GAN Discrimintor**

[그림] 은 본 연구에서 사용된 discriminator의 아키텍쳐이다. 전체적인 아키텍쳐는 Unet-Gan의 Discriminator 아키텍쳐를 가지고 있으며, 다운샘플링 되는 인코더 부분을 거쳐 차원을 증가시키고 맥스풀링으로 Global Feed back을 진행한다. 이는 기존의 GAN의 discriminator와 동일한 구조를 갖는다.(스킨 커넥션을 제외한) 동일한 구조를 갖는 만큼 역할 또한 동일하게 전체적인 이미지가 실제와 얼마나 유사한 분포를 갖는지에 대해 feed back한다. 이후 디코더 부분을 거쳐 다시 차원을 축소 시키고 이미지를 복원한다. 이때에 스킵 커넥션을 통해 GT를 반영하여 기존의 이미지와 유사한 형태를 복원 시키고 최종적으로 각각의 pixel (per pixel)을 local feed back을 함으로 디테일하게 피드백을 진행한다. Global/Local Feed back 모두 activation function으로 sigmoid를 통해 출력된다. 각각의 해당하는 y값은 실제 이미지를 넣었을 때의 Global/Local feed back으로 1의 값이 주어지고, 가짜 이미지를 넣었을 때 Global feed back은 0, Local Feed back은 연기를 제외한 이미지에 1, 연기에 해당하는 Local에 0을 주어 Binary-crossentropy를 Loss Function으로 진행한다. 연기를 제외한 이미지에 1 이 주어지는데 이때 초기 ADVLoss의 초기 알파 값을 적게 주어 generator에 주는 영향을 적게 만들어 다른 손실함수로 기존의 이미지를 복원하게 하고 안정화가 진행 되면 Discriminator의 Local 부분이 영향을 미치게 한다. 이는 Discirminator가 Global하게는 연기가 있는 사진을 가짜로 판별하고, Local하게는 실제 이미지를 제외한 연기 부분만을 Feed back 하게 만든다.

이에 따른 수식과 상세 네트워크는 아래를 따른다.

[수식] Adverserial Loss랑 Local Feed back 수식

[표] discriminator 상세 (1페이지)

****

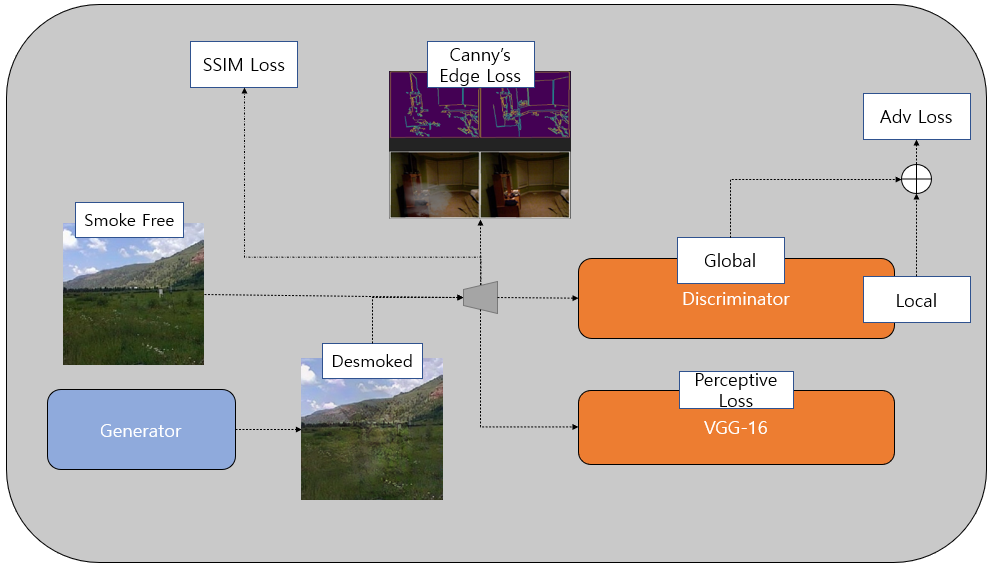
**[그림] 주어진 real/fake 이미지와 Smoke Mask GT and Discriminator 출력값 (임시,수정요구)**

**[표] GT, Discriminator 출력값 IoU**

[그림]은 real/fake 이미지와 그에 fake 이미지의 Smoke Mask GT 그리고 Discriminator의 출력 값이다.

[표]는 Test data set의 GT와 출력값의 IoU를 나타낸다. IoU는 ~~ 값을 나타나고, 블라블라. 이곳에 TP,TN,FP,FN 표를 기입 예정

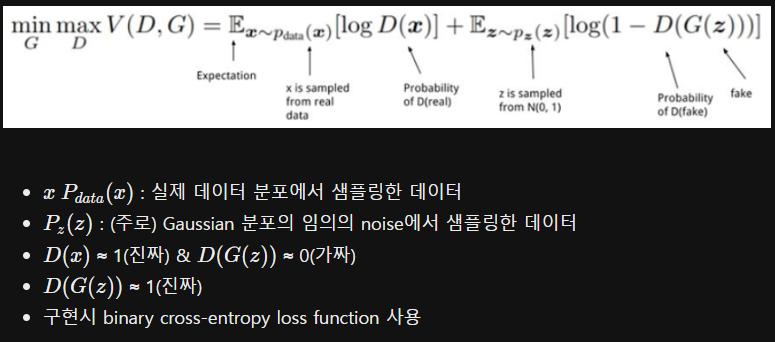
**전체 구조도**



**3.3 손실함수**

(Unet-Gan Adverserial Loss +Edge Loss+ SSIM Loss + Perceptive Loss)

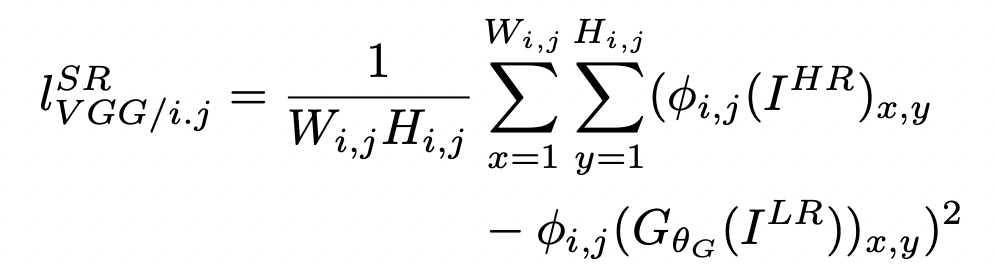
~~-~~ Adverserial Loss

****

D는 training data의 sample과 G의 sampl에 진짜인지 가짜인지 올바른 라벨을 지정할 확률을 최대화하기 위해 학습하고, (V(D,G) 최대화) G는 log(1-D(G(z))를 최소화(D(G(z))를 최대화)하기 위해 학습되도록하는 것이 (V(D,G) 최소화) GAn minmax loss의 핵심이다. ( 추후 adv Loss rate와 함께 수정)

**- VGG Loss**

VGG Loss는 SRGAN( [CVPR2017] photo-realistic single image super-resolution using a generative adversarial network )에서 처음 등장한 Loss이다. MSE loss의 문제점(smooting, texture 반영 실패 등) 을 해결하기 위해 'VGG Loss'를 새로 정의하는데 VGG Loss는 다음과 같은 식으로 정의된다.



perceptual similarity에 더 가까운 loss로, 이 loss는 pre-train된 19 layer VGG network의 ReLU 활성화 계층을 기반으로 정의한다.

pre-trained된 VGG net을 이용해서 feature map(VGG19 network 속 i번째 maxpooling layer 전 activation 후 j번째 convolution에 의해 얻어진 feature map)에서의 유클리디안 거리를 계산하는 방법을 사용한다. reconstructed image와 reference image(high resolution image)의 feature map사이 거리를 구하는 것이다.

따라서 pixel 각각의 값이 아닌 'perceptual similarity'에 집중하였기 때문에 좀 더 detail한 부분을 잘 잡아낼 수 있게 되었다.[CVPR 2018] The unreasonable effectiveness of deep features as a perceptual metric

**3.4 결과 분석**

인페인팅의 특성을 가지고 있으나, 실제로는 연기이기에 Dehaze 알고리즘들과 본 연구의 네트워크를 비교한다. 우선 본 네트워크에 의 각각 loss 별 score를 비교하고 다른 dehaze model과 비교 분석한다. 스코어는 각각 일반적으로 널리 알려지고 사용되어지는 L1, L2을 사용하고 추가적으로 SSIM, PSNR, LPIPS를 비교한다.The SSIM is a well-known quality metric used to measure the similarity between two images. The structural similarity index (SSIM) (Wang et al., 2004)[37] is an FR-IQA measure inspired by the theory that HVS is highly adapted for extracting structural information from the scenes. The incorporation of this characteristic as the intrinsic component of an Image Quality Assessment (IQA) measure allowed authors to outperform not only the MSE-based measures but also the existing state-of-the-art perceptual image quality measures. The PSNR value approaches infinity as the MSE approaches zero. this shows that a higher PSNR value provides a higher image quality. At the other end of the scale, a small value of the PSNR implies high numerical differences between images.[38]. LPIPS는 비교적 초기의 ImageNet classsification 모델인 AlexNet, VGG, SqueezeNet을 사용합니다. LPIPS는 기존의 IS나 FID와는 다르게 유사도를 사람의 인식에 기반하여 측정하려 시도했습니다. 그 과정에서 AlexNet, VGG, SqueezeNet의 feature map이 사람의 인식과 유사하기 때문에 이를 활용하고자 하였습니다. 낮은 LPIPS는 두 이미지가 더욱 유사함을 의미합니다.[39]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Loss 비교표 | | | | | |
|  | L1 | L2 | SSIM | PSNR | LPIPS |
| Adverserial (norm) ,SSIM,Perspective Loss |  |  |  |  |  |
| Adverserial Loss(norm) + SSIM + Edge+Perspective Loss | 0.07918 | 0.0180 | 0.9215 | 24.5040 | 0.1008 |
| Adverserial (unet) ,SSIM,Perspective Loss | 0.0660 | 0.0141 | 0.93677 | 25.9828 | 0.08667 |
| ours  Adverserial Loss(unet) + SSIM + Edge+Perspective Loss | **0.0659** | **0.0137** | **0.9347** | **26.1005** | **0.0846** |

표[]은 본 연구의 네트워크에 각각 다른 Loss를 적용 하였을때 IQA Score이다.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 타 모델과 비교 | | | | | | |
|  | L1 | L2 | SSIM | PSNR | LPIPS | FID |
| DarkChannelPrior | 0.1160 | 0.0284 | 0.8419 | 16.4343 | 0.2256 | 43.8674 |
| DehazeNet | 0.1088 | 0.0217 | 0.7954 | 17.4828 | 0.2005 | 42.1189 |
| Non Local Image Dehazing | 0.1352 | 0.0367 | 0.7527 | 15.1274 | 0.2447 | 15.2004 |
| ours | **0.0280** | **0.0028** | **0.9428** | **27.1647** | **0.0744** | **10.3016** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| (a) | (b) | (c) | (d) | (e) | (f) |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**(a) randomly selected test input images, (b) Dark Channel Prior,**

**(c) DehazeNet, (d) Non-Local Image Dehazing (e)ours (f) Ground Truth**

**3.5 discussion**

**ㅇ**

**결론**

3Please review the [Author Guidelines](https://www.elsevier.com/journals/information-processing-and-management/0306-4573/guide-for-authors).

1. ACKNOWLEDGMENTS

This is just an example. This section is optional.

1. REFERENCES

[0] Visibility in bad weather from a single image 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2008), pp. 1-8

[1] https://www.nhtsa.gov/ 미국통계청

[2] Single Image Haze Reduction Removal Using Dark Channel Prior

[3] J. Bruna, P. Sprechmann, and Y. LeCun. Super-resolution with deep convolutional sufficient statistics. In International Conference on Learning Representations (ICLR), 2016.

[4] A. Dosovitskiy and T. Brox. Generating images with perceptual similarity metrics based on deep networks. In Advances in Neural Information Processing Systems (NIPS), pages 658–666, 2016.

[5] M. Mathieu, C. Couprie, and Y. LeCun. Deep multi-scale video prediction beyond mean square error. In International Conference on Learning Representations (ICLR), 2016.

[6] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deepnetwork training by reducing internal covariate shift. In Proceedingsof The 32nd International Conference on Machine Learning (ICML),pages 448–456, 2015.

[7] O. Elharrouss, N. Almaadeed, S. Al-Maadeed, Y. Akbari Image inpainting: a reviewNeural Process. Lett., 51 (2) (2020), pp. 2007-2028

[8] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proc. ACM SIGGRAPH Conf. Comput. Graph., 2000: pp. 417–424.

[9] Variational image inpainting technique based on nonlinear second-order diffusions Comput. Electr. Eng., 54 (2016), pp. 345-353

[10] DilatedNet(2016), Multi-Scale Context Aggregation

[11] ResNet - Residual Connection

[12] K. Simonyan, A. Zisserman Very Deep Convolutional Networks for Large-scale Image Recognition

[13] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli Image quality assessment: from error visibility to structural similarity IEEE Trans. Image Process., 13 (4) (2004), pp. 600-612

[14] Shift-net

[15] DeepFill v1

[16] Generation Multi-column Convolutional Neural Networks

[17] Dilated Residual Networks

[18] Generative Adversarial Networks  
[19]J. Johnson, A. Alahi, L. Fei-Fei Perceptual losses for real-time style transfer and super-resolution ECCV, Springer (2016), pp. 694-711

[20] X. Yu, F. Porikli Ultra-resolving face images by discriminative generative networks ECCV, Springer (2016), pp. 318-333

[a] Tchaka K., Pawar V.M., Stoyanov D. Chromaticity based smoke removal in endoscopic images Medical Imaging 2017: Image Processing, Vol. 10133, International Society for Optics and Photonics (2017)

[21] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis. In International Conference on Learning Representations (ICLR), 2019. 1, 2, 5, 7, 8, 9, 12, 16, 19, 23

[22] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In Conference on Computer Vision and Pattern Recognition (CVPR), 2019. 1, 2, 7, 12

[23] Ting Chen, Xiaohua Zhai, Marvin Ritter, Mario Lucic, andNeil Houlsby. Self-supervised gans via auxiliary rotationloss. In Conference on Computer Vision and Pattern Recognition (CVPR), 2019. 2

[24] Chieh Hubert Lin, Chia-Che Chang, Yu-Sheng Chen, DaCheng Juan, Wei Wei, and Hwann-Tzong Chen. Coco-gan: Generation by parts via conditional coordinating. In International Conference on Computer Vision (ICCV), 2019. 1,2,8, 12

[25] Chieh Hubert Lin, Chia-Che Chang, Yu-Sheng Chen, DaCheng Juan, Wei Wei, and Hwann-Tzong Chen. Coco-gan: Generation by parts via conditional coordinating. In International Conference on Computer Vision (ICCV), 2019. 1,2,8, 12

[26] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. “Labelme: A database and web-based tool for image annotation”. International Journal of Computer Vision, vol. 77, no. 1, pp. 157–173, May 2008.

[27] P. K. Nathan Silberman, Derek Hoiem and R. Fergus. “Indoor segmentation and support inference from rgbd images”. ECCV, 2012.

[28] Hsiang-Yin Cheng et al. “Smoke100k: A Database for Smoke Detection”. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)

[29] VishalVenkatesh et al. “Unsupervised smoke to desmoked laparoscopic surgery images using contrast driven Cyclic-DesmokeGAN”. Computers in Biology and Medicine Volume 123, August 2020

[30] Mahendra Khened et al. “Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers”. Med. Image Anal., 51 (2019), pp. 21-45

[31] J. Li, F. Fang, K. Mei, G. Zhang, Multi-scale residual network for image super-resolution, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 517–532.

[32] S. Nah, T. Hyun Kim, K. Mu Lee, Deep multi-scale convolutional neural network for dynamic scene deblurring, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3883–3891.

[33] Ş.Kiliç, İ.Askerzade, Y.Kaya Using ResNet Transfer Deep Learning Methods in Person Identification According to Physical Actions IEEE Access, 8 (2020), pp. 220364-220373,

[34] S.Al-Abri, T.X.Lin, M.Tao, F.Zhang A Derivative-Free Optimization Method With Application to Functions With Exploding and Vanishing Gradients IEEE Control Syst. Lett., 5 (2) (April 2021), pp. 587-592

[35] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition.

[36] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[37] Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: From error visibility to structural similarity IEEE Transactions on Image Processing, 13 (4) (2004), pp. 600-612

[38] Hore, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." 2010 20th international conference on pattern recognition. IEEE, 2010.

[39] Zhang, Richard, et al. "The unreasonable effectiveness of deep features as a perceptual metric." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[40] LiuZ., TanY., HeQ., XiaoY. SwinNet: Swin transformer drives edge-aware RGB-D and RGB-T salient object detection IEEE Trans. Circuits Syst. Video Technol. (2021), 10.1109/TCSVT.2021.3127149

[41]Yu E., Sun J., Li J., Chang X., Han X.-H., Hauptmann A.G. Adaptive semi-supervised feature selection for cross-modal retrieval IEEE Trans. Multimed., 21 (5) (2018), pp. 1276-1288, 10.1109/TMM.2018.2877127

[42] Min W., Mei S., Li Z., Jiang S. A two-stage triplet network training framework for image retrieval IEEE Trans. Multimed., 22 (12) (2020), pp. 3128-3138, 10.1109/TMM.2020.2974326

[43]Chang X., Ren P., Xu P., Li Z., Chen X., Hauptmann A.G. A comprehensive survey of scene graphs: Generation and applicationIEEE Trans. Pattern Anal. Mach. Intell. (2021),

[44] Pi Z., Shao Y., Gao C., Sang N. Instance-based feature pyramid for visual object tracking IEEE Trans. Circuits Syst. Video Technol. (2021),

[45] Cheng D., Zhou J., Wang N., Gao X. Hybrid dynamic contrast and probability distillation for unsupervised person Re-Id IEEE Trans. Image Process., 31 (2022), pp. 3334-3346,

[46] Berman D., Avidan S., et al. Non-local image dehazing IEEE Conference on Computer Vision and Pattern Recognition

[47] Cai B., Xu X., Jia K., Qing C., Tao D. Dehazenet: An end-to-end system for single image haze removal IEEE Trans. Image Process., 25 (11) (2016), pp. 5187-5198,

[48] YanLi et al. Single image dehazing with an independent Detail-Recovery Network. Knowledge-Based Systems

Volume 254, 27 October 2022, 109579