

INTRODUCTION TO VIDEO AND IMAGE COMPUTING LAB



2018. 7.



Video and Image Computing Lab

❖ Location

- VIC Lab website: <http://www.viclab.kaist.ac.kr>
- Offices
 - Professor: Rm. 1107, Bldg. N24, KAIST
 - Students: Rms. 1106, 1108, Bldg. N24, KAIST

❖ Members

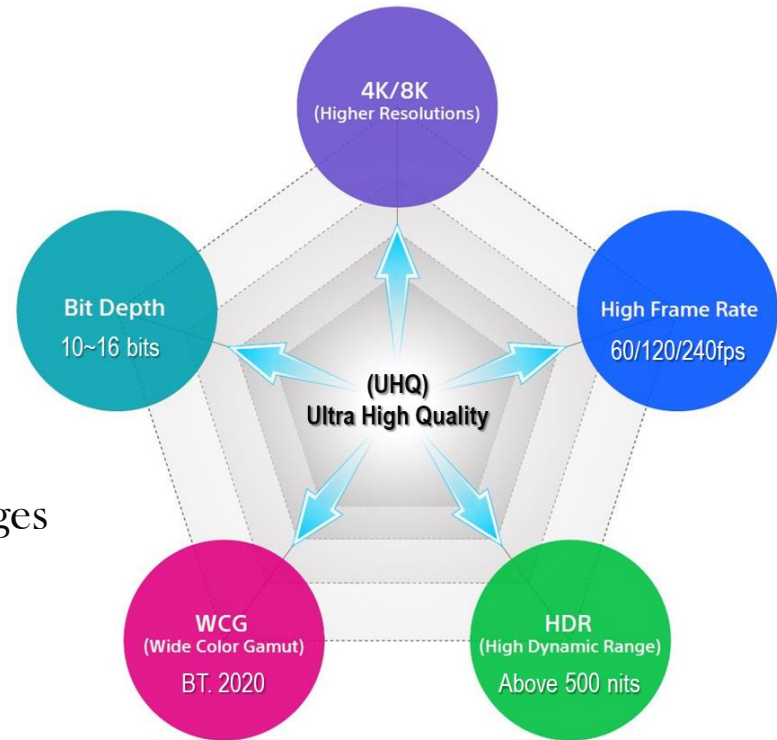
- Professor
 - Munchurl Kim (Full Professor)
- Ph. D. Students (9 persons)
- Master Students (8 persons)

Video and Image Computing Lab

❖ Research Areas

■ Computational Imaging and Image/Video Processing

- Machine/Deep Learning based single Image and Video Super Resolution
- (Reserve) Tone Mapping for (SDR-to-HDR) HDR-to-SDR Conversions
- Deep Learning based Image Denoising,
Motion Deblurring
- Frame Rate Up Conversion (FRUC)
- Image Dehazing
- Camera Image Signal Processing (ISP)
 - Denoising and Demosaicking for CFA Images
 - Denoising for Low-light Images
- Image Dehazing
- Quality Enhancement and Defect Detection for Semiconductor Wafer Images



Video and Image Computing Lab

❖ Research Areas

■ Image Quality Assessment

- Perceptual Visual Quality Distortion Metrics based on HVS's Perceptual Quality Modeling
- Perceptual Visual Quality Estimation for Full-Reference, Reduced-Reference, No-Reference

■ Deep-Learning based Object Recognition

- Automatic Target Detection and Recognition for SAR Imagery
- Classification of Ballistic Missile Fragments

■ 2D/3D Video Coding

- Deep Video Compression for High Compression and High Quality
- Perceptual Video Coding (PVC)
- HDR Video Coding Optimization

1. 초해상화 (Super-Resolution, SR)

❖ SR 기법

- 저해상도(Low-Resolution, LR) 영상으로부터 고해상도(High-Resolution, HR) 영상으로의 변환
- 업스케일링(Upscaling)으로도 불림

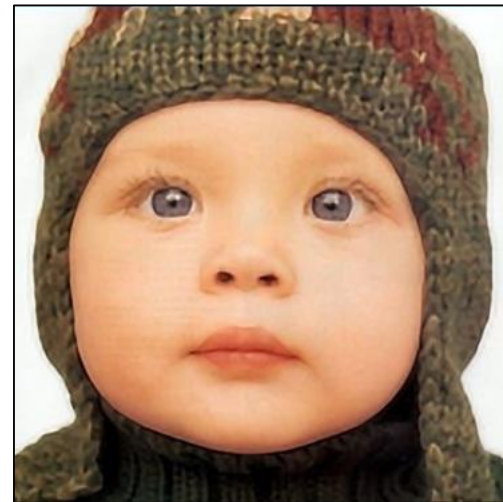


FHD: 1920×1080
4K: 3840×2160
8K: 7680×4320



저해상도 입력 영상

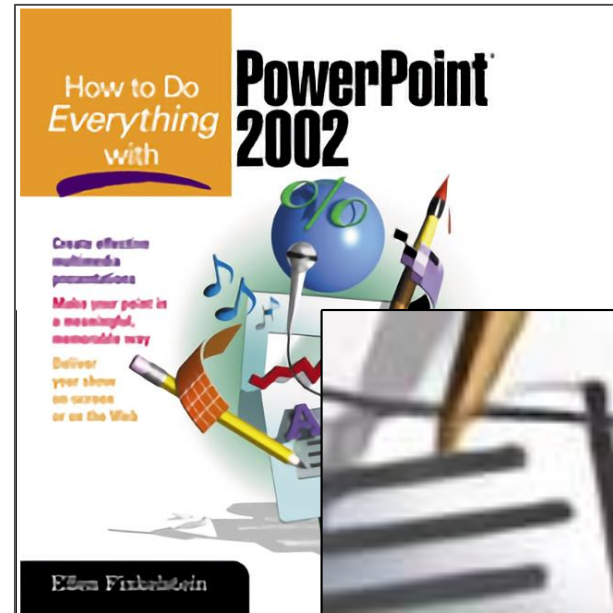
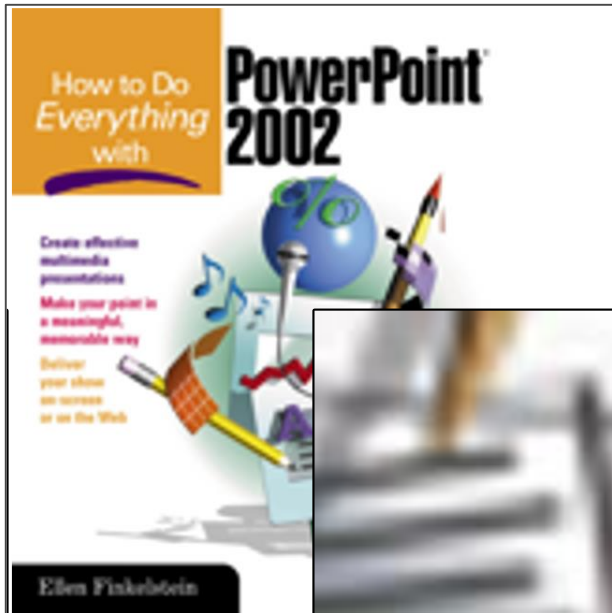
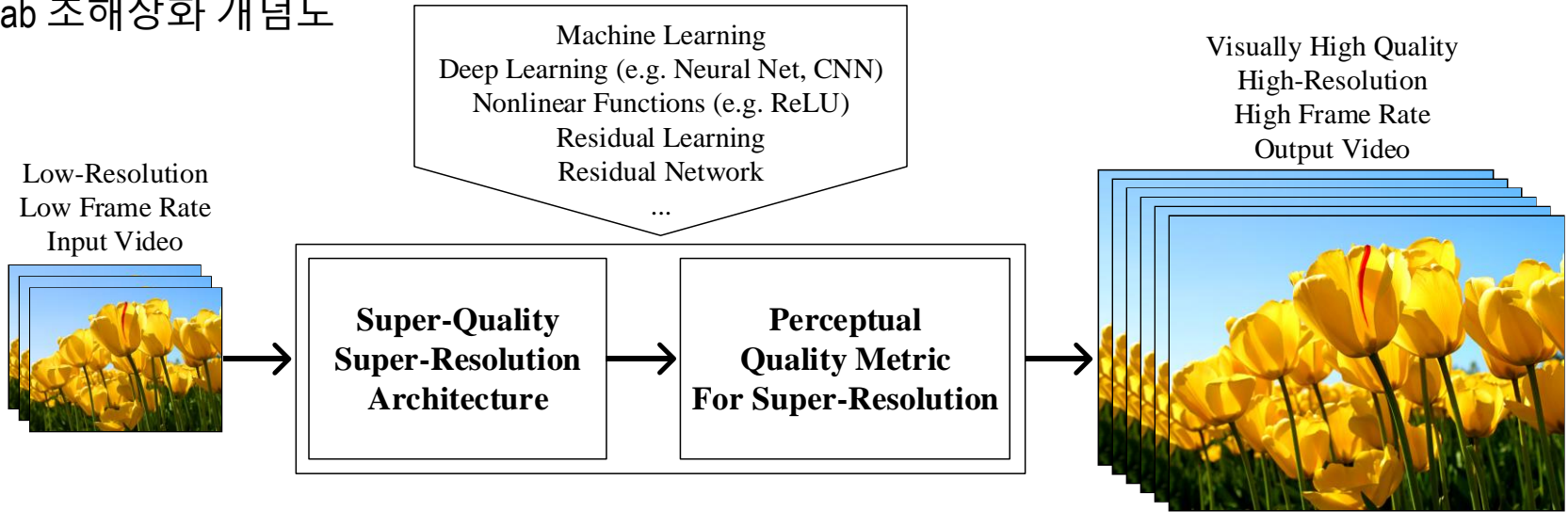
딥러닝 기반
저해상도에서 고해상도
매핑



고해상도 결과 영상

1. 초해상화 (Super-Resolution, SR)

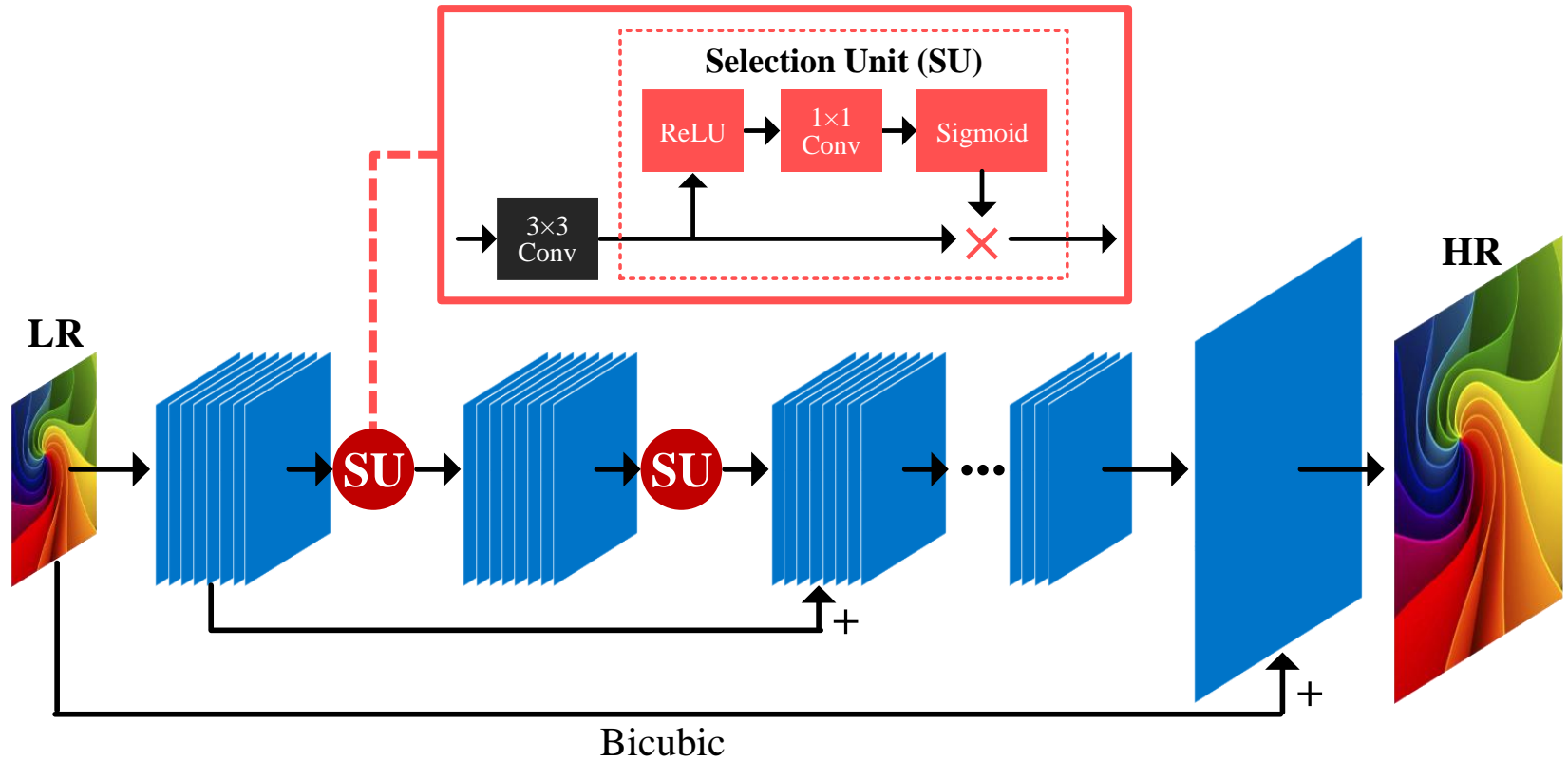
VIC Lab 초해상화 개념도



기존 보간법

VIC Lab
초해상화 결과

1.1 Proposed SR Network using SU (SelNet)

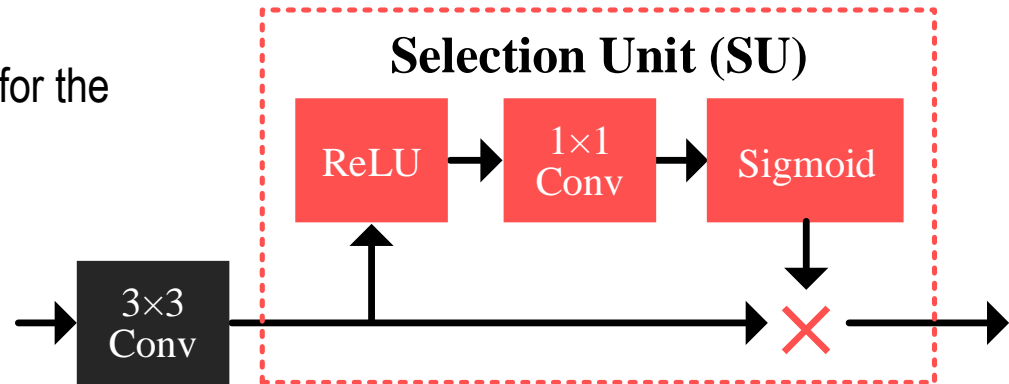


- + 22 conv layers
- + Improved residual units
- + Residual learning
- + Sub-pixel convolutions
- + Gradient switching for training

1.1 Proposed SR Network using SU (SelNet)

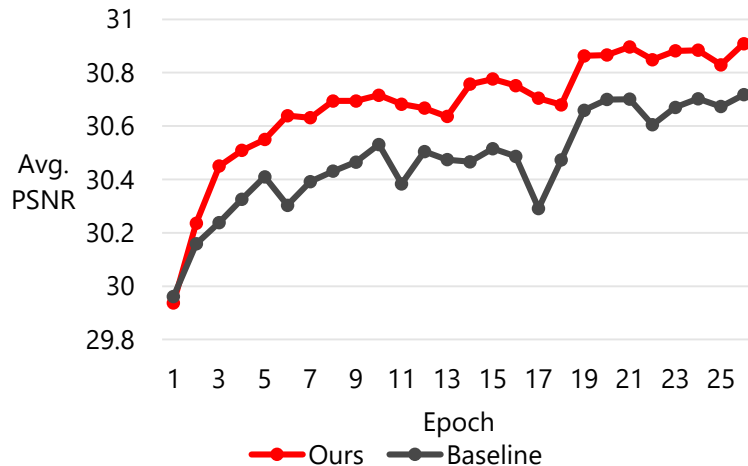
■ Selection Unit (SU)

- A novel nonlinear activation unit for the SR problem.



✓ **Filter A:** 3×3 Conv

✓ **Filter B:** 3×3 Conv – ReLU – 1×1 Conv



- A PSNR performance curve for our toy network with SU and a baseline network with ReLU.
- The basic architecture for the both networks are the same (6 conv layers).
- The number of channels for two networks is adjusted, so that the total number of filter parameters are the same.

1.1 NTIRE Challenge 2017



- 4th Place
 - The lowest computational complexity
 - The highest SSIM value for a scale factor of 4

Scale	Metric	1st SNU_CVLab ²	2nd SNU_CVLab ²	3rd Lab402	4th Our SelNet ³
2	PSNR	34.93	34.83	34.66	34.29
	SSIM	0.948	0.947	0.946	0.943
	Time	67.24	14.07	4.08	0.54
	Complexity	x125	x26	x8	-
3	PSNR	31.13	31.04	30.83	30.52
	SSIM	0.889	0.888	0.882	0.880
	Time	28.72	7.34	5.12	0.27
	Complexity	x106	x27	x19	-
4	PSNR	26.91	29.04	28.83	28.55
	SSIM	0.752	0.836	0.830	0.845
	Time	20.05	5.24	5.22	0.19
	Complexity	x106	x28	x27	-



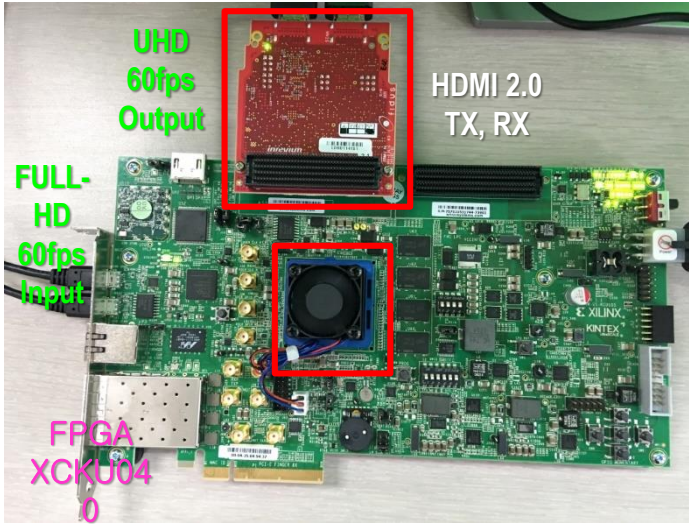
¹R. Timofte, et al. NTIRE 2017 challenge on single image super-resolution: methods and results. IEEE CVPRW: NTIRE Challenge, Hawaii, USA, 2017.

²B. Lim et al. Enhanced deep residual networks for single image super-resolution. IEEE CVPRW: NTIRE Challenge, Hawaii, USA, 2017.

³J.-S. Choi and M. Kim. A deep convolutional neural network with selection units for super-resolution. IEEE CVPRW: NTIRE Challenge, Hawaii, USA, 2017.

1.2 CNN 기반 실시간 SR HW 설계 및 FPGA 구현

HW Design and Real-Time (4K60P) Implementation of CNN-based Super-Resolution



전자신문(2018.01.17)

딥러닝 활용... '초고화질 영상' 변환



YTN News - "AI 기술로 '초고화질 TV 영상' 만든다" (2018.01.18)



동아일보(2018.01.17)

풀HD TV로 화질 4배 좋은 UHD 영상 본다

KAIST 연구진, AI활용 첫 개발
셋톱박스 크기 장치로 고화질 감성
360도 VR 콘텐츠에도 적용 가능



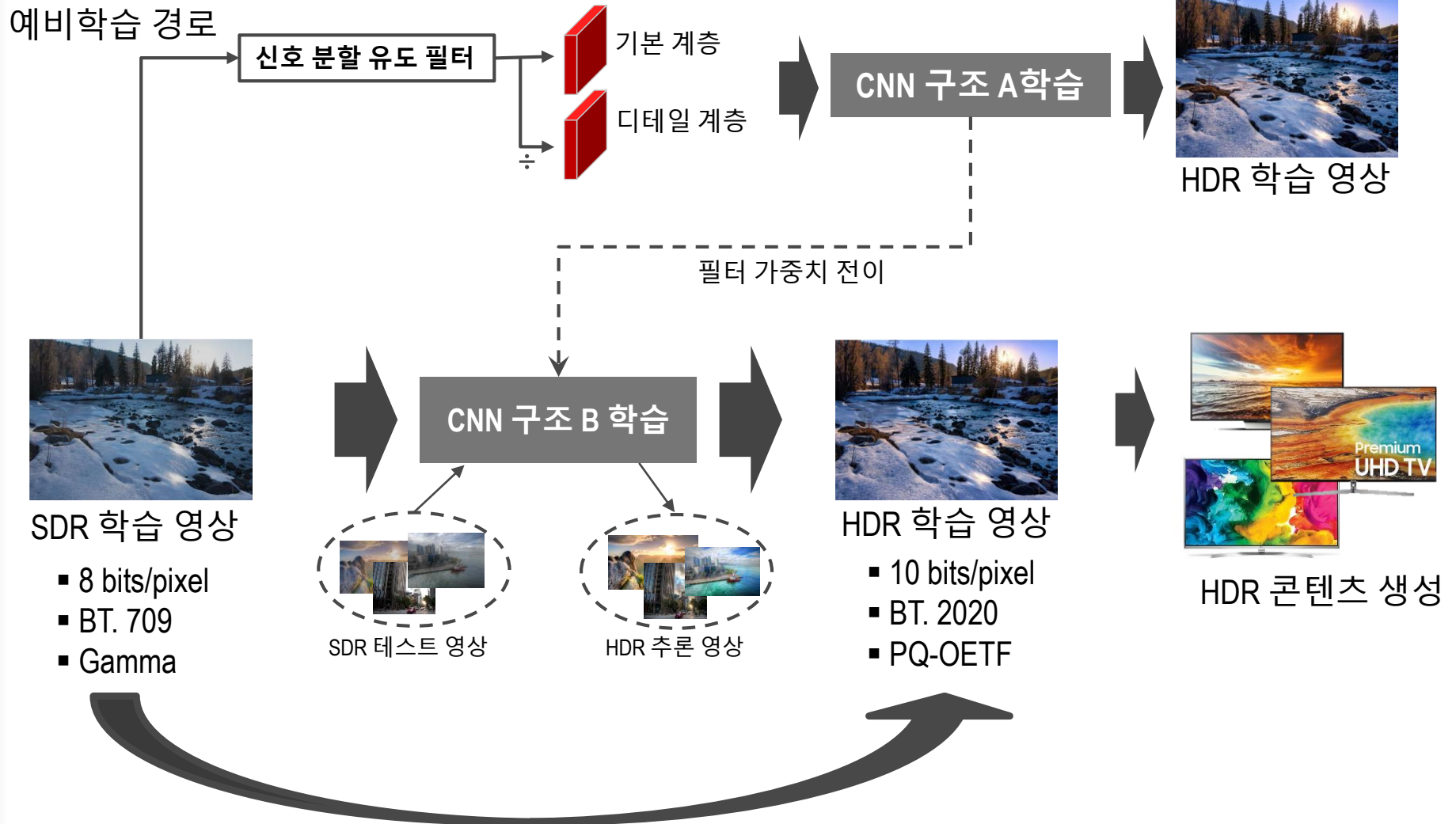
화질의 향상이 대부분 '아래'라 불리는 '풀HD' TV에 셋톱박스 크기의 장치만 설치하면, FHD 방송 콘텐츠도 쉽게 UHD 영상으로 즐길 수 있다고 밝혔다. 이 기술은 360도 가상현실(VR) 콘텐츠에도 적용 가능하다. 연구진은 FHD(1920×1080) 영상을 입력하면, 가로세로 픽셀 수가 각각 2배씩 늘어난 UHD(3840×2160) 영상으로 출력되는 인공지능 기반 딥러닝(딥컨볼루션신경망(DCNN))을 설계했다. FHD 영상을 그냥 늘리면 화면이 깨지겠지만, 이 인공지능은 화면을 구성하는 각 픽셀을 늘려 고화질 영상을 만들어낸다. 이런 일이 가능한 이유는 연구진이 영상 처리를 위해 인공지능에게 수백만 번 반복학습을 시켰기 때문이다. 같은 영상을 저화질

(128×128)과 고화질(256×256)로 만들어 각각 입력 값과 출력 값(정답)으로 한 뒤, 목표 값과 출력 값을 비교한다. 학습되면 출력하는 값이 주어지지 않아도 FHD 영상을 입력하면 UHD 영상을 생성해 낸다. 이때 인공지능에게 저화질과 고화질의 배율을 높여 학습시킨 UHD보다 더 화질이 좋은 초고화질 영상을 만들 수도 있다. 연구진은 인공지능 기반의 영상 효율을 높이고 이를 작은 크기의 하드웨어로 구현하는 데도 성공했다. 김 교수는 "고효율의 합성된 장면 없이도 인공지능을 적용시킬 수 있기 때문에 상용화 가능성이 높다"고 말했다. 이 시스템의 저장용량은 10KB, 소비전력은 12W 정도다. 이날 김 교수는 해당 기술을 국내 특허로 출원했으며 향후 미국과 일본, 중국 등 해외 시장도 출원할 계획이다. 출처: 동아일보 기사 kyun@donga.com

UHD 업스케일링 하드웨어에 FPGA 기반의 CNN 기반 4K-UHD 업스케일링 기술을 시연하는 모습.

기존의 풀 HD 영상은 업스케일링 과정만 거친다. 이번에는 주로 DCNN이 활용된다. DNN의 특징은, 인공지능 기반 학습을 통해 입력된 데이터가 주어지지 않아도 FHD 영상을 입력하면 UHD 영상을 생성해 낸다. 이때 인공지능에게 저화질과 고화질의 배율을 높여 학습시킨 UHD보다 더 화질이 좋은 초고화질 영상을 만들 수도 있다. 연구진은 인공지능 기반의 영상 효율을 높이고 이를 작은 크기의 하드웨어로 구현하는 데도 성공했다. 김 교수는 "고효율의 합성된 장면 없이도 인공지능을 적용시킬 수 있기 때문에 상용화 가능성이 높다"고 말했다. 이 시스템의 저장용량은 10KB, 소비전력은 12W 정도다. 이날 김 교수는 해당 기술을 국내 특허로 출원했으며 향후 미국과 일본, 중국 등 해외 시장도 출원할 계획이다.

2. DCNN based Reverse Tone Mapping



2. DCNN based Reverse Tone Mapping

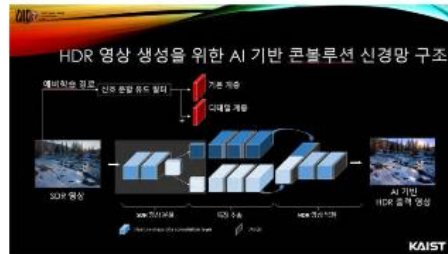
KAIST, 딥러닝 기반 HDR 영상 변환 기술 개발

김영준 입력 2018.04.09 15:00 **닷컴 0개**

국내 연구진이 딥러닝 기술로 일반 영상을 고품질 하이다이내믹레인지(HDR) 영상으로 변환하는 기술을 개발, 초고화질(UHD) 콘텐츠 및 고품질 영상 시장을 확대하는 기반을 마련했다.

KAIST(총장 신성철)는 김문철 전기 및 전자공학부 교수팀이 보통 화질의 스탠더드다이내믹레인지(SDR) 영상을 HDR 영상화하는 기술 개발에 성공했다고 8일 밝혔다.

디스플레이로 표현되는 사물은 실제 사물과 많은 차이를 보인다. 기존 SDR 영상의 표현 가능 밝기는 보통 0.1~100니트(1니트는 1㎡ 면적에 쬐볼 한 겨울 든 밝기) 수준이다. 아주 밝거나 어두운 영상의 디테일을 표현할 수 없고, 색상 재현 능력도 떨어진다.

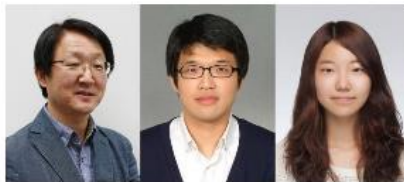


HDR 영상 생성 CNN 구조

대안으로 나온 것이 UHD 방송에 쓰이는 HDR 기술이다. HDR 영상의 밝기는 0.05~1000니트다. 재현 색상도 기존 16만7000가지에서 107만여개로 늘어난다. 더 실제에 가까운 영상을 제공한다.

문제는 기존에 만들어진 SDR 콘텐츠의 HDR화 성능이 떨어진다는 것이다. HDR 기능을 갖춘 TV가 기존 SDR 콘텐츠를 HDR로 변환·재생하는 기능을 갖췄지만 세밀함이 부족하다. 밝거나 색을 일몰로 더하는 식이어서 주변보다 밝은 곳에 일부 그림자가 저 있어도 모두 밝게 표현한다. 아직 대부분의 영상 콘텐츠가 SDR 규격이어서 고성능 변환 기술 개발이 필수다.

연구팀은 최신신경망(CNN) 딥러닝 학습으로 SDR 영상을 고품질 HDR화하는 기술을 개발했다. 같은 대상을 촬영한 SDR-HDR 영상을 반복 학습시켜 이후 어떤 SDR 영상을 입력해도 고품질 HDR 영상으로 변환할 수 있도록 했다. 영상의 기본·세부 정보를 분리해서 HDR화하는 방법으로 표현에 세밀함을 더했다. 사물의 텍스처(표면, 예지(테두리)) 정보를 따로 뽑아내 각 특징 영역을 세분, 처리했다.



HDR 영상 기술 개발에 KAIST 연구진, 왼쪽부터 김문철 전기 및 전자공학부 교수, 김대훈 박사과정, 김수재 석사과정

연구팀은 앞으로 이 기술을 하드웨어(HW)로 구현, 상용 TV에 적용할 방침이다. 방송사·스트리밍 사업자용 SDR-HDR 변환 플랫폼 구축도 염두에 두고 있다.

김문철 교수는 "이 기술이 상용화되면 기존에 만들어진 영상을 더욱 실제에 가깝게 변환해서 보는 것이 가능해진다"면서 "HDR 콘텐츠 부족 문제를 해결할 수 있다"고 말했다.

대건=김영준기자 kjy85@etnews.com

일반 영상도 고화질로...HDR 변환 기술 개발

(하이다이내믹레인지)



김문철 kyj85@ 과학 기자

김문철 KAIST 교수팀
최신신경망 딥러닝 학습 활용
SDR 영상→고품질 HDR 영상화
향후 HW 구현·상용 TV 적용 계획
SDR-HDR 변환 플랫폼 구축도

국내 연구진이 딥러닝 기술로 일반 영상을 고품질 하이다이내믹레인지(HDR) 영상으로 변환하는 기술을 개발, 초고화질(UHD) 콘텐츠 및 고품질 영상 시장을 확대하는 기반을 마련했다.

KAIST(총장 신성철)는 김문철 전기 및 전자공학부 교수팀이 보통 화질의 스탠더드다이내믹레인지(SDR) 영상을 HDR 영상화하는 기술 개발에 성공했다고 8일 밝혔다.

디스플레이로 표현되는 사물은 실제 사물과 많은 차이를 보인다. 기존 SDR 영상의 표현 가

능 밝기는 보통 0.1~100니트(1니트는 1㎡ 면적에 쬐볼 한 겨울 든 밝기) 수준이다. 아주 밝거나 어두운 영상의 디테일을 표현할 수 없고, 색상 재현 능력도 떨어진다.

대안으로 나온 것이 UHD 방송에 쓰이는 HD

R 기술이다. HDR 영상의 밝기는 0.05~1000니트다. 재현 색상도 기존 16만7000가지에서 107만여개로 늘어난다. 더 실제에 가까운 영상을 제공한다.

문제는 기존에 만들어진 SDR 콘텐츠의 HDR화 성능이 떨어진다는 것이다. HDR 기능을 갖춘 TV가 기존 SDR 콘텐츠를 HDR로 변환·재생하는 기능을 갖췄지만 세밀함이 부족하다. 밝거나 색을 일몰로 더하는 식이어서 주변보다 밝은 곳에 일부 그림자가 저 있어도 모두 밝게 표현한다. 아직 대부분의 영상 콘텐츠가 SDR 규격이어서 고성능 변환 기술 개발이 필수다.

연구팀은 최신신경망(CNN) 딥러닝 학습으로 SDR 영상을 고품질 HDR화하는 기술을 개

발했다. 같은 대상을 촬영한 SDR-HDR 영상을 반복 학습시켜 이후 어떤 SDR 영상을 입력해도 고품질 HDR 영상으로 변환할 수 있도록 했다. 영상의 기본·세부 정보를 분리해서 HDR화하는 방법으로 표현에 세밀함을 더했다. 사물의 텍스처(표면, 예지(테두리)) 정보를 따로 뽑아내 각 특징 영역을 세분, 처리했다.

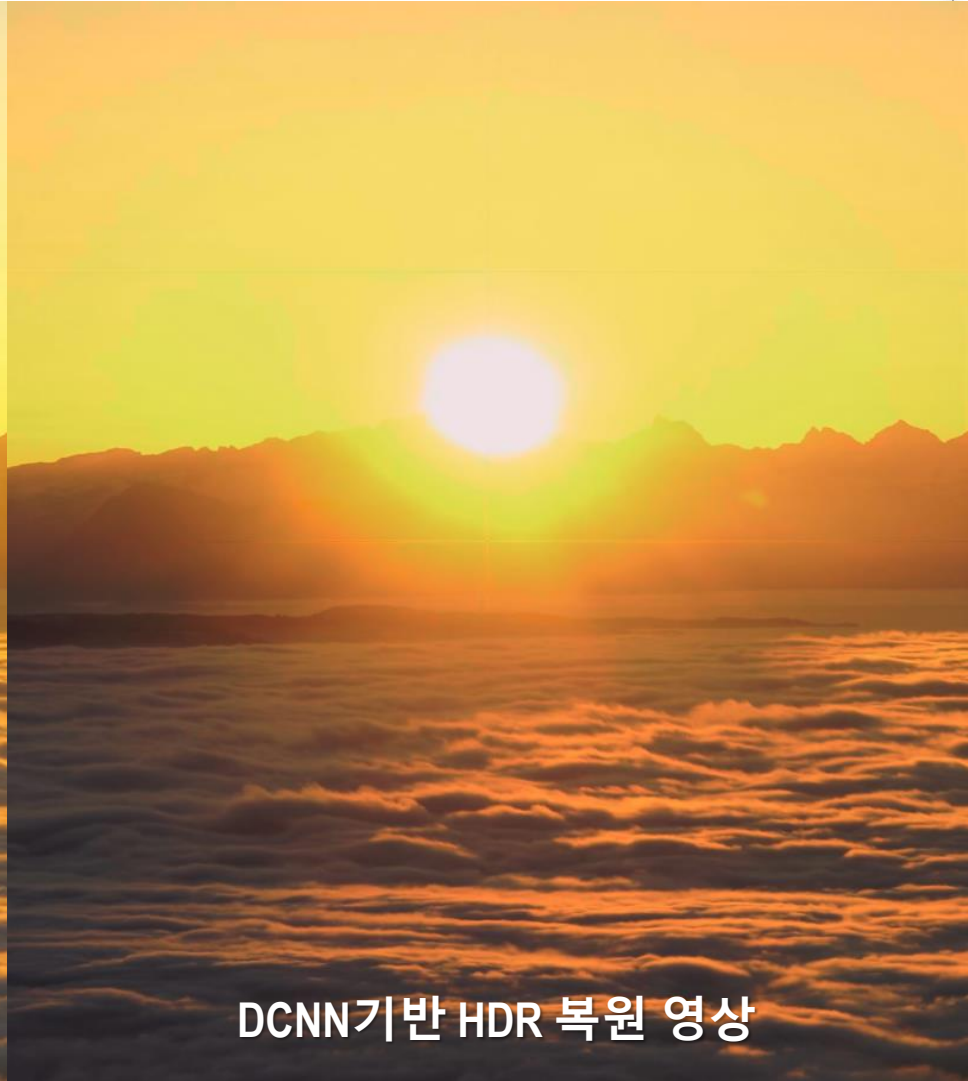
연구팀은 앞으로 이 기술을 하드웨어(HW)로 구현, 상용 TV에 적용할 방침이다. 방송사·스트리밍 사업자용 SDR-HDR 변환 플랫폼 구축도 염두에 두고 있다.

김문철 교수는 "이 기술이 상용화되면 기존에 만들어진 영상을 더욱 실제에 가깝게 변환해서 보는 것이 가능해진다"면서 "HDR 콘텐츠 부족 문제를 해결할 수 있다"고 말했다.

전자신문 (2018.04.09)



기존 복원 영상



DCNN기반 HDR 복원 영상



기존 복원 영상



DCNN기반 HDR 복원 영상



기존 복원 영상



DCNN기반 HDR 복원 영상



기존 복원 영상



DCNN기반 HDR 복원 영상

3. Motion Deblurring (움직임 흐림 제거)

❖ Motion blur(움직임 흐림)

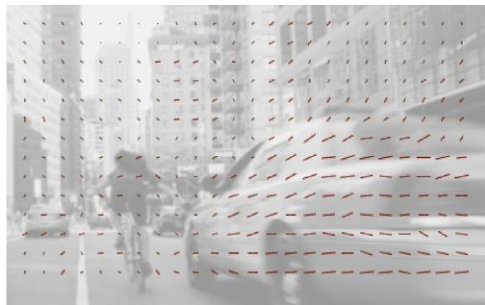
- 영상 획득 과정에서 카메라의 빠른 움직임 또는 장면내의 객체 움직임이 빠르거나 카메라 셔터막의 노출 시간이 긴 경우 발생
- 영상이 일반적으로 흐려져 주관적 화질을 떨어뜨리는 요인
- 하나의 영상 내에 균일 움직임 흐림과 불균일 움직임 흐림이 존재 가능



Sharp Image



Uniform Motion Blur



Blur Kernel

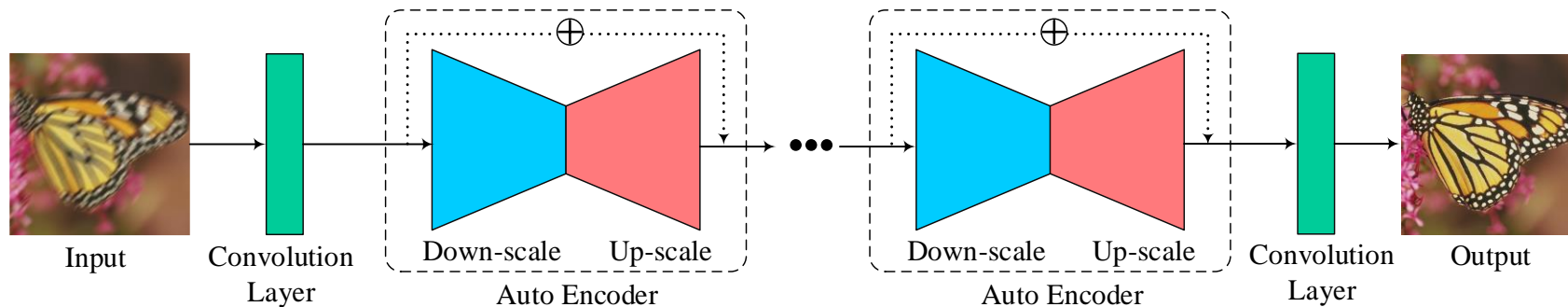


Non-uniform Motion Blur

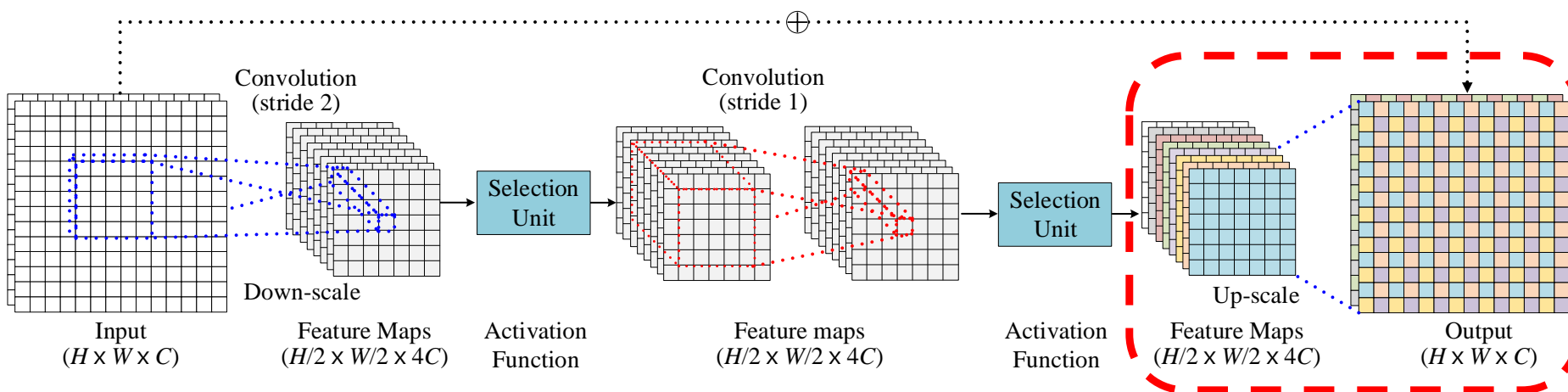
3.1 CNN 기반 움직임 흐림 제거

❖ CNN-based Motion Deblur Network with Our ERC Unit

■ ERC – Extended Receptive-Field Convolution Unit



Proposed CNN Structure with Enlarged Receptive Convolution (ERC) Units



3.2 Experimental Results of GOPRO dataset

❖ Input Blurry Image



3.2 Experimental Results of GOPRO dataset

❖ Result of Our ERC32 net



3.2 Experimental Results of GOPRO dataset

❖ Input Blurry Image



3.2 Experimental Results of GOPRO dataset

❖ Result of Our ERC32 net



3.2 Experimental Results of GOPRO dataset

❖ Input Blurry Image



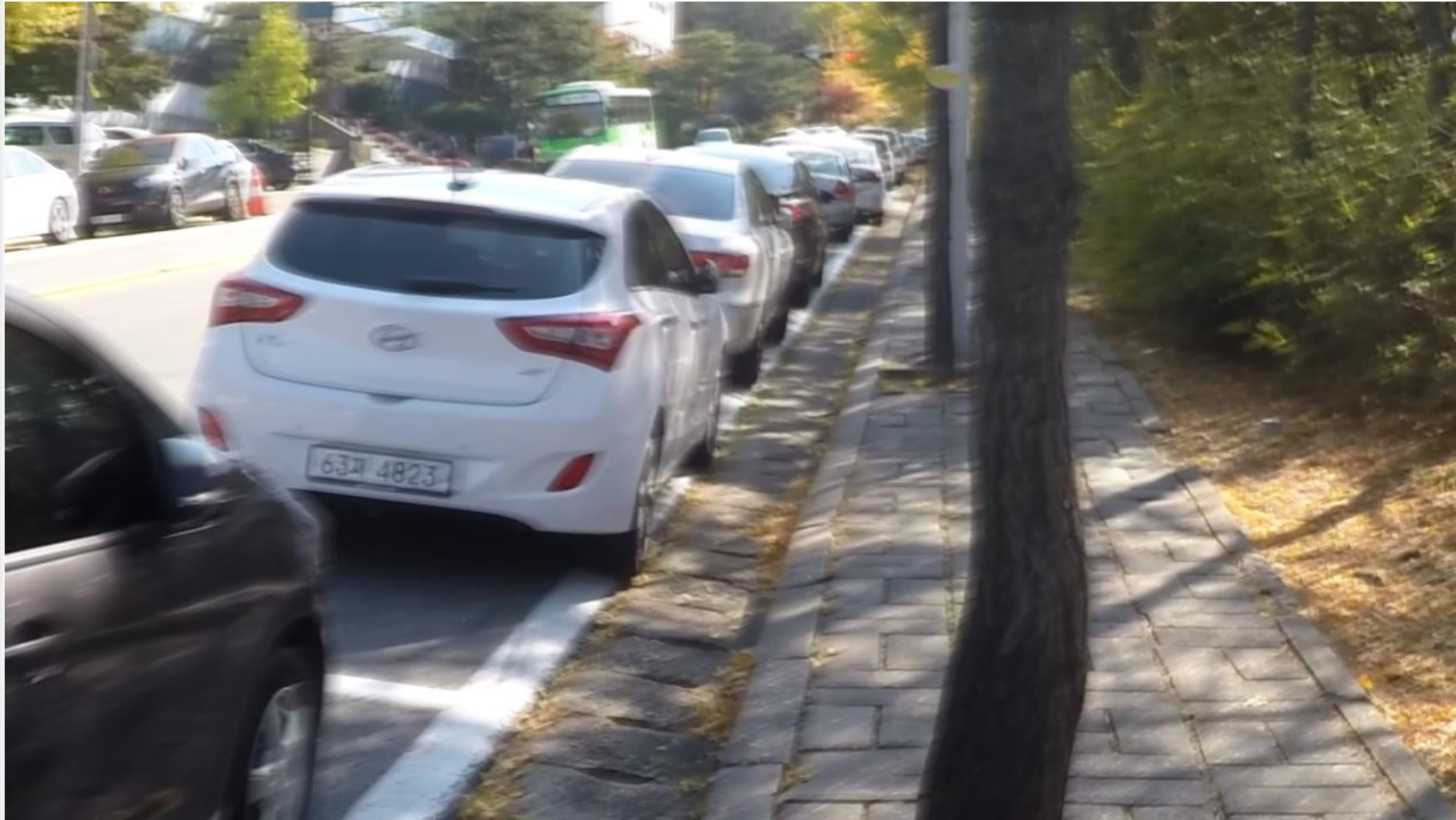
3.2 Experimental Results of GOPRO dataset

❖ Result of Our ERC32 net



3.2 Experimental Results of GOPRO dataset

❖ Input Blurry Image



3.2 Experimental Results of GOPRO dataset

❖ Result of Our ERC32 net



3.3 Experimental Results of DBN dataset

❖ Input Blurry Image



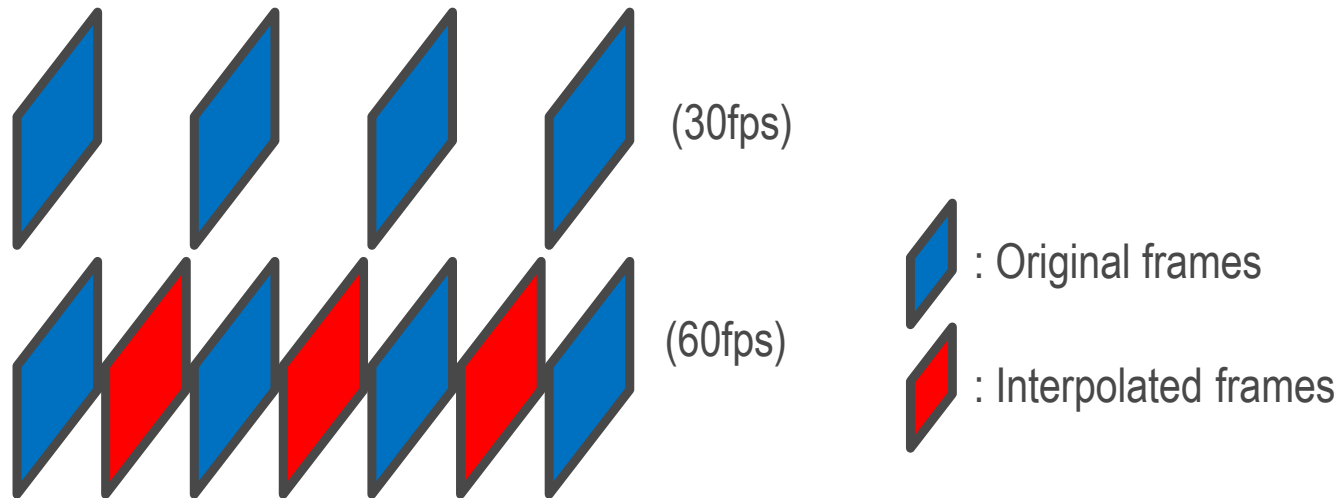
3.3 Experimental Results of DBN dataset

❖ Result of Our ERC32 net

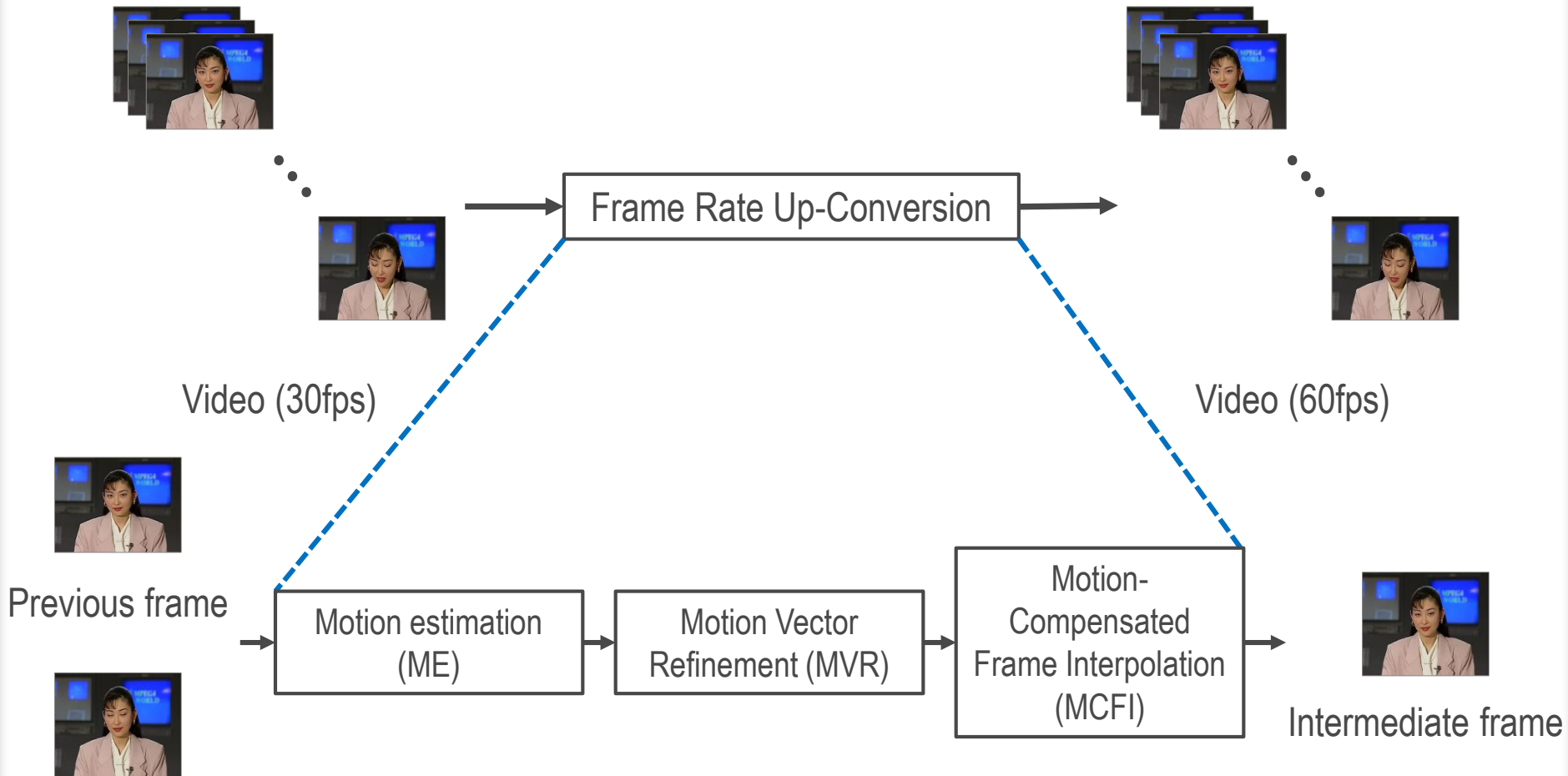


4. Frame Rate Up-Conversion (FRUC)

- ❖ Motion judder is jerky movement on screen.
- ❖ FRUC is used to increase the number of frames for improving visual quality.
 - A video plays more smoothly because increasing the number of frames alleviates discontinuous motion which is called *motion judder*.



4.1 Workflow of conventional ME & MC based FRUC

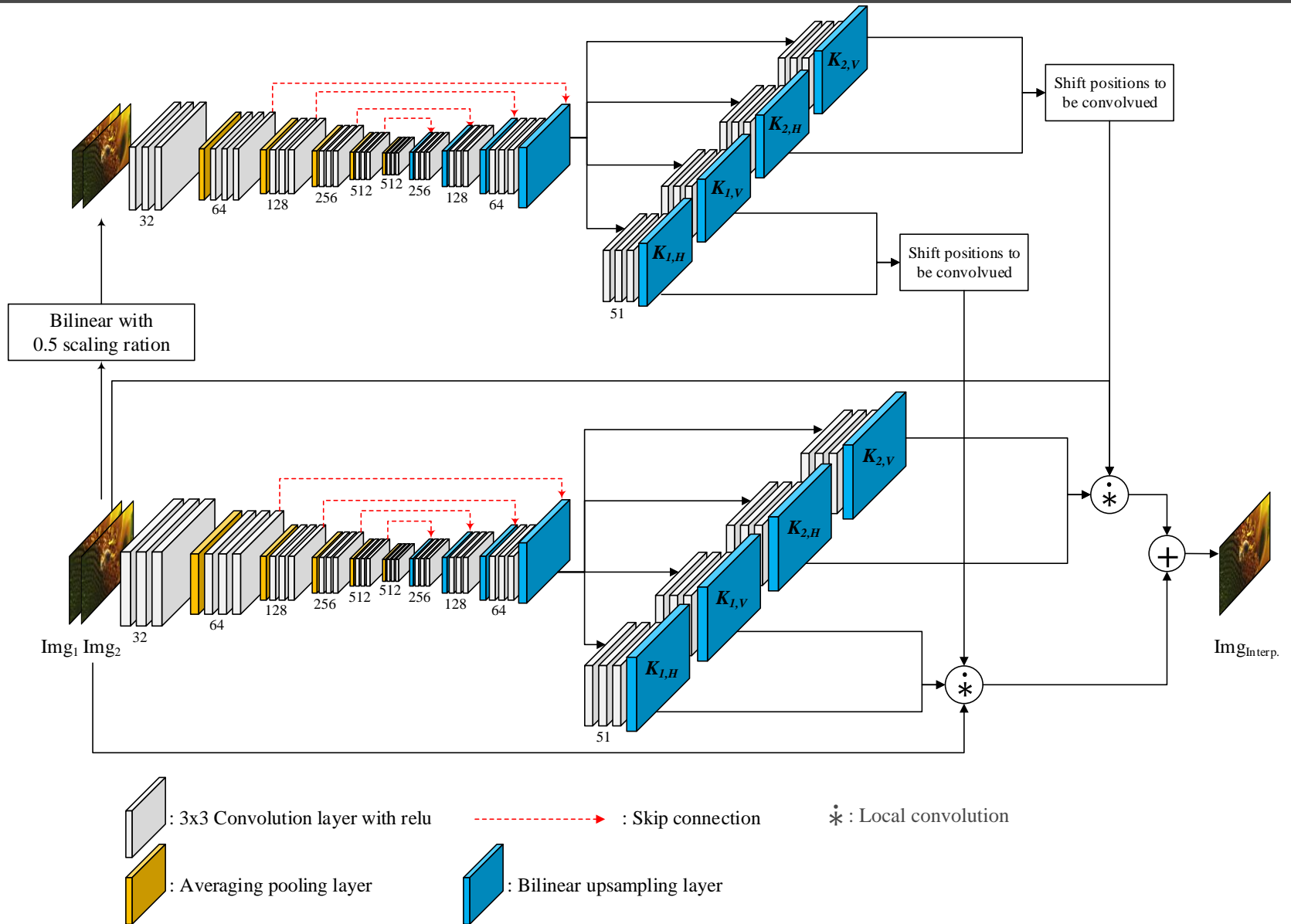


Previous frame

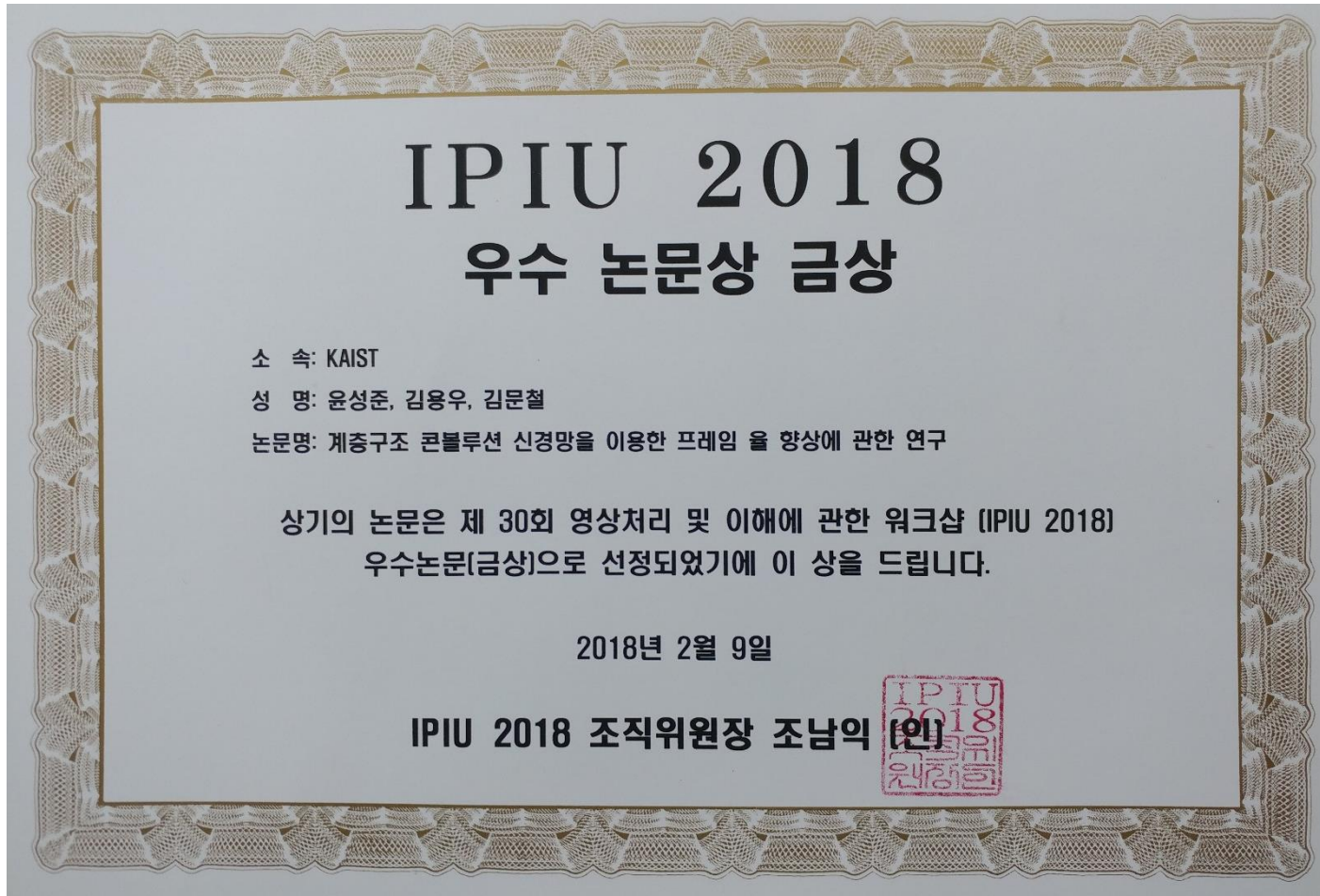
Current frame

ME*: Motion estimation
MC*: Motion compensation

4.2 Hierarchical CNN-based FRUC Network

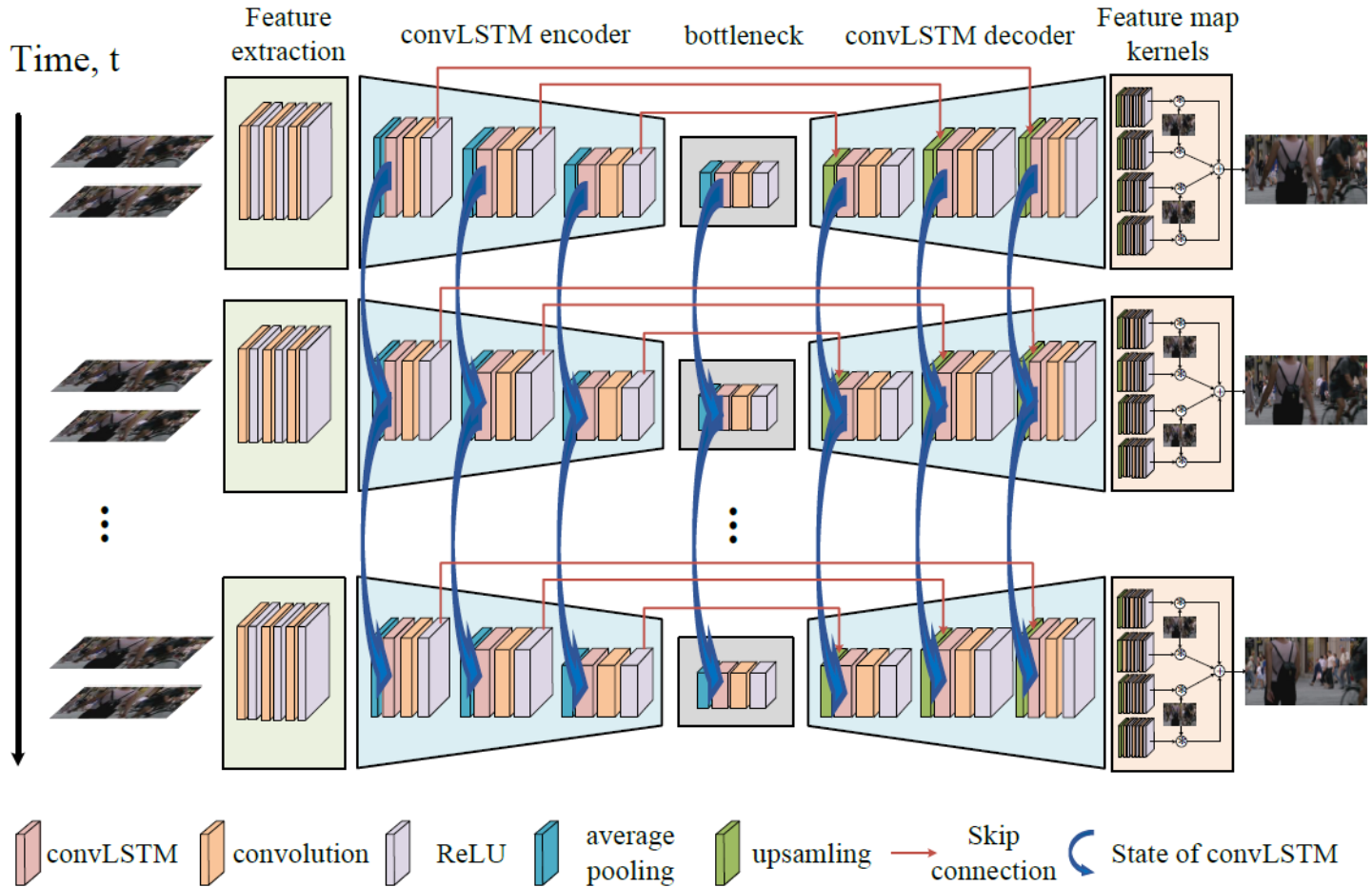


4.2 Hierarchical CNN-based FRUC Network



4.3 convLSTM 기반 FRUC

❖ convLSTM Network for FRUC



Architecture of the proposed convLSTM FRUC

4.3 convLSTM 기반 FRUC

❖ convLSTM Network for FRUC

AI 활용...영상 프레임 초당 30→60장으로 늘린다



김영준 kyj85@ 과학 기자

KAIST(총장 신성철)가 인공지능(AI)을 활용해 비디오 영상 프레임을 초당 30장에서 60장으로 늘려주는 기술을 개발했다.

김문철 KAIST 전기 및 전자공학부 교수팀은 기존 영상 고프레임화 기술을 개선해 'AI 딥러닝 기반 고프레임율 비디오 생성 기술'을 개발했다고 30일 밝혔다.

비디오 영상은 다수의 정지 화면을 연속해 보는 것이다. 프레임율이 높을 수록 움직임이 끊김없이 부드럽다. 저해상도 콘텐츠는 초당 30프레임 영상이 주를 이루지만, 4K UHD 비디오는 초당 60프레임이 표준이다. 해상도가 증가할수록 사람이 체감하는 프레임 변화가 커진다.

고프레임화 기술로는 일본 영상의 움직임을

김문철 KAIST 교수팀
기존 영상 고프레임화 기술 개선
AI 딥러닝 기반 움직임 변화 학습

TV·빔프로젝터 등에 적용 가능
UHD 방송 활성화에도 도움

예측해 새로운 프레임을 생성하는 '보간' 기술이 나왔다. 하지만 아직은 기술 성숙도가 낮아 화질이 떨어진다. '컨볼루션신경망(CNN)'을 이용했는데, 예측 정확도가 낮다.

김 교수팀은 기존 CNN 모델에 AI 기술을 더해 시간 변화에 따른 영상 속 움직임 변화를 정확하게 학습하도록 하는 '순환신경망(RNN)' 구조를 결합한 'convLSTM-CNN' 모델을 개



김문철 KAIST 교수(오른쪽)와 연구진이 기존 영상과 개발 기술로 프레임율을 높인 영상을 비교하고 있다.

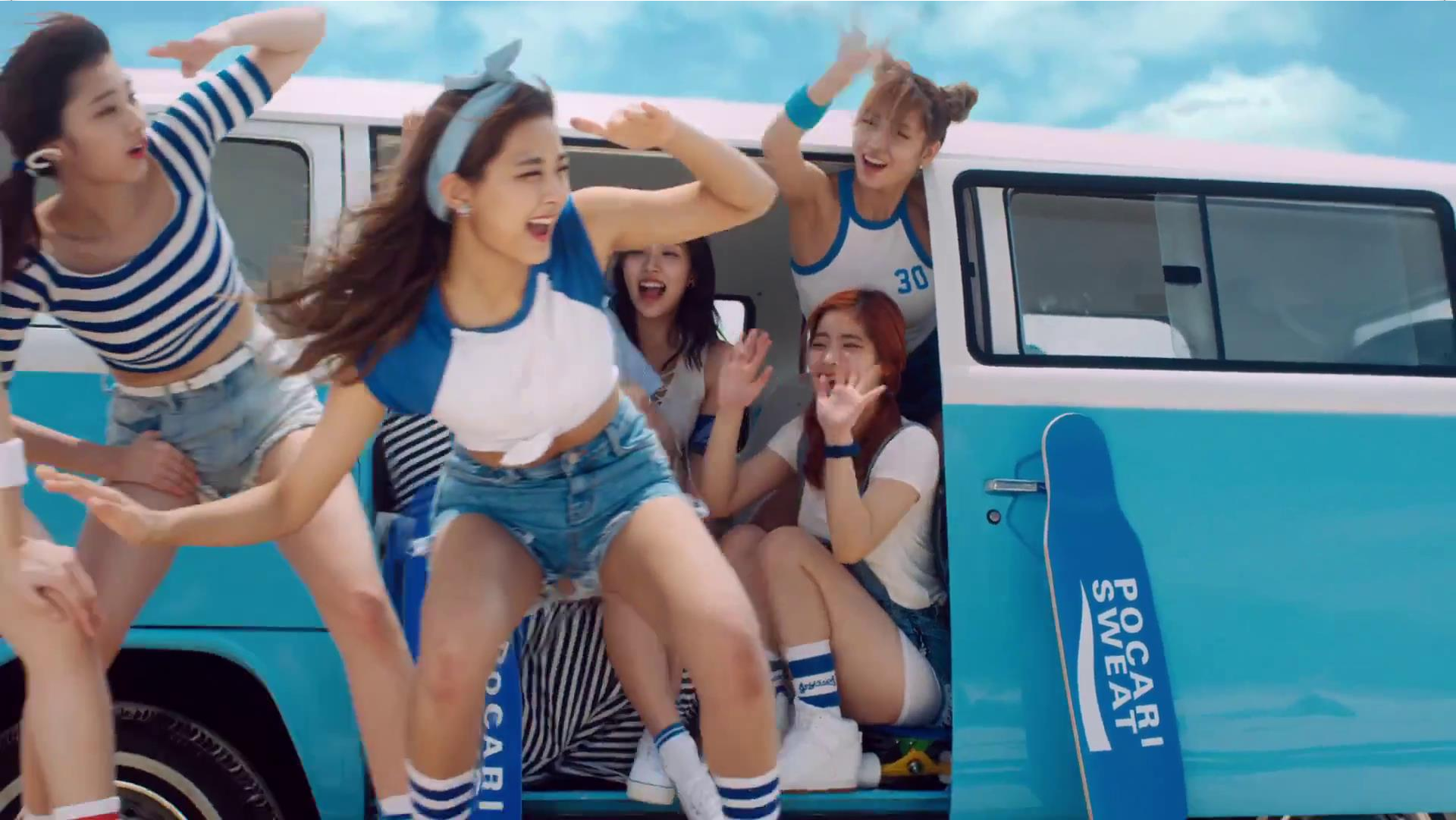
발해 문제를 해결했다.

UHD 방송·IPTV·스트리밍 사업자 등이 이 기술을 활용하면 고화질 콘텐츠를 대거 확보할 수 있다. TV 및 빔프로젝터, 헤드마운티드디스플레이(HMD)에도 적용할 수 있다.

김문철 교수는 "UHD 방송 활성화를 위해서는 공간 해상도뿐만 아니라 프레임율을 높이는 기술도 필요하다"며 "KAIST의 AI 기반 영상 품질 향상 기술이 관련 산업 분야에 넓게 쓰일 전기를 마련했다"고 말했다.

전자신문 (2018.05.31)

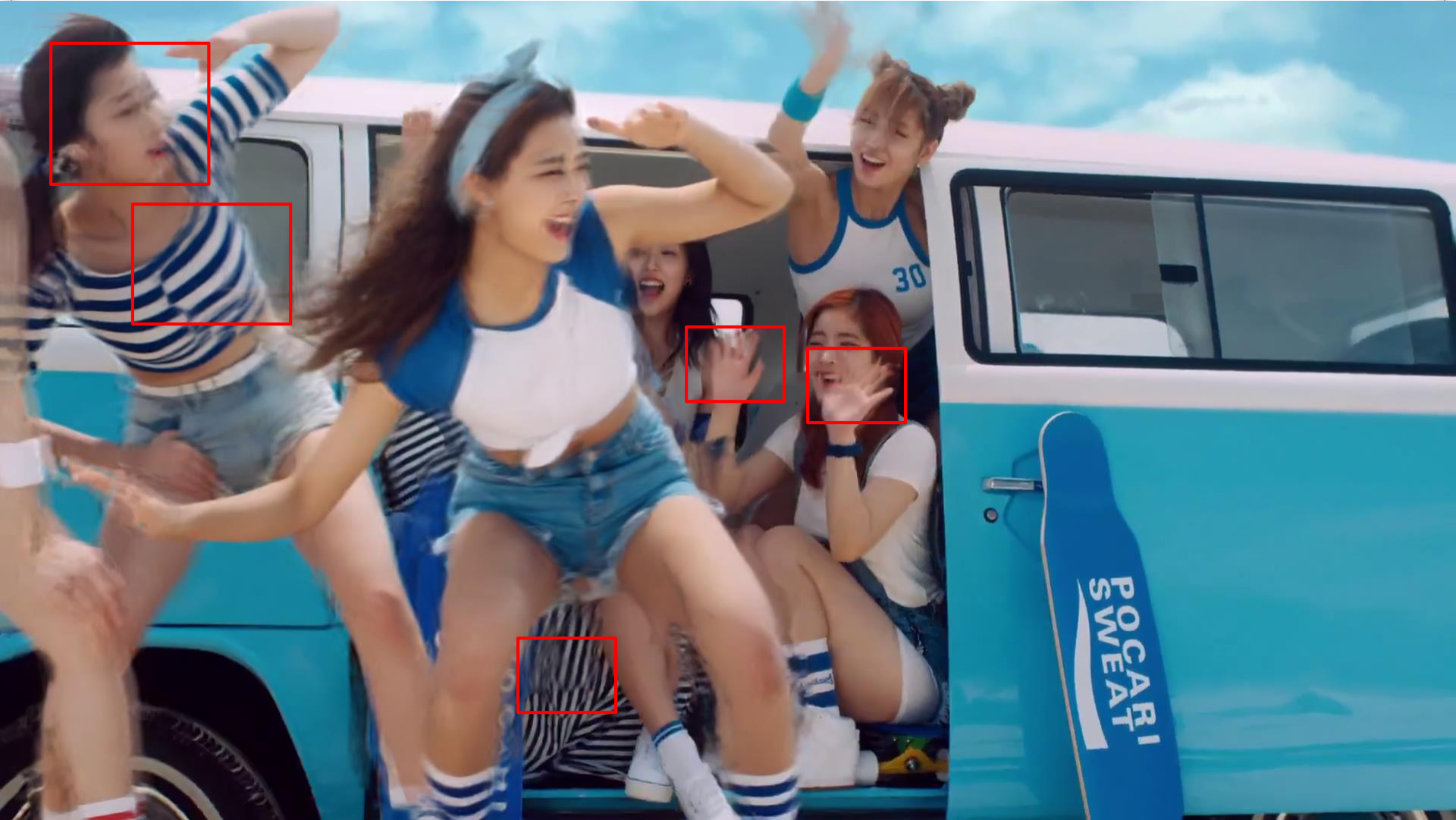
Pocari area 71st frame



Pocari area 72nd frame

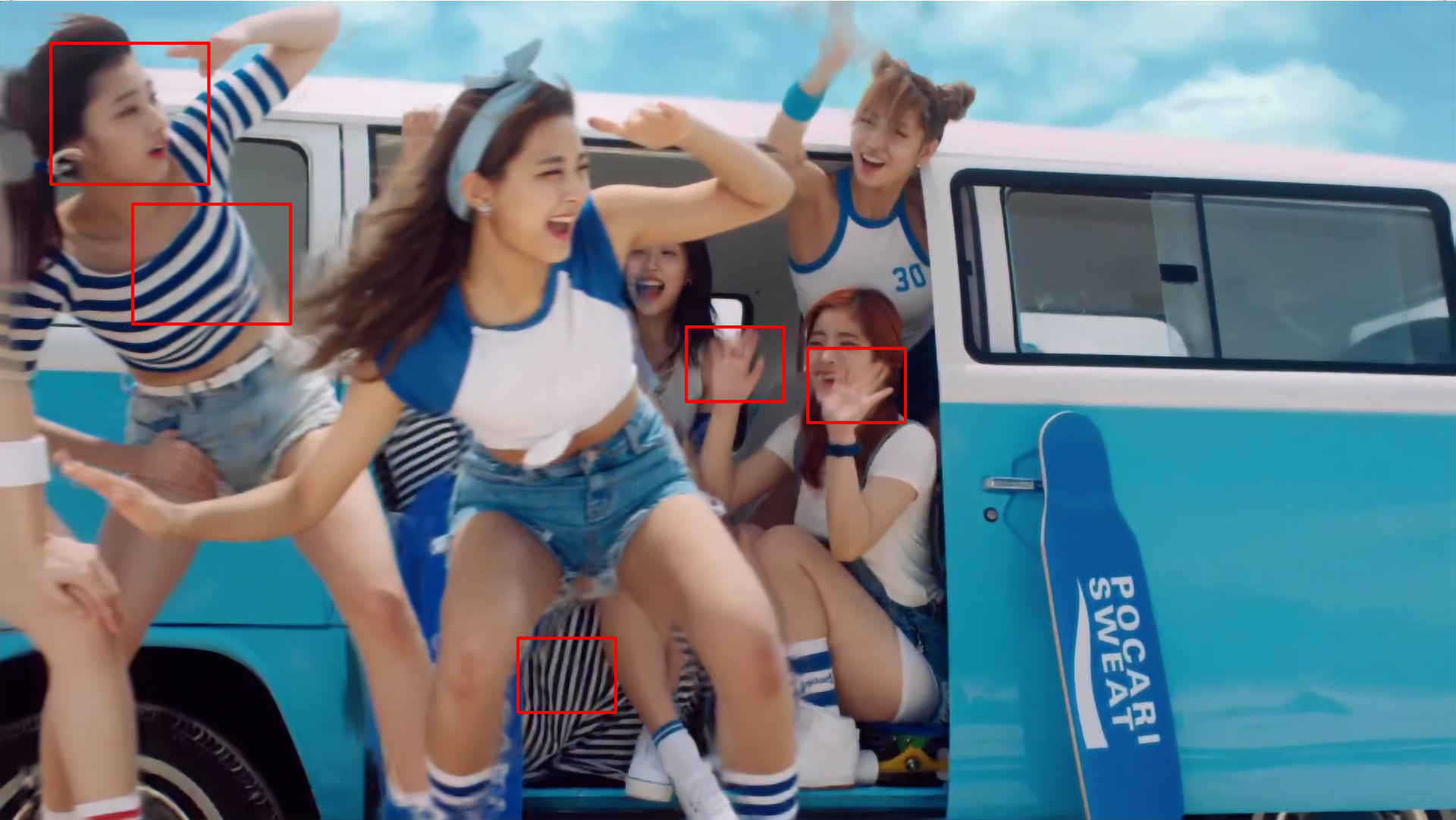


Niklaus [2]



[2] S. Niklaus, L. Mai, and F. Liu, "Video Frame Interpolation via Adaptive Separable Convolution," Proc. IEEE Int. Conf. Computer Vis. (ICCV), Venice, Italy, Oct. 2017.

Proposed Hierarchical CNN-based VFI



Short track 743th frame



2016-17 ISU 쇼트트랙 월드컵 1차
남자 500m(1) 준준결승 **용경완 출전**

SBS Sports
LIVE

2016 WBSU-23 세계선수권
니카라과 vs 대한민국
오늘 아침 **8:50**



WR 39.937
4.3

Short track 744th frame



2016-17 ISU 쇼트트랙 월드컵 1차
남자 500m(1) 준준결승 **용경완 출전**

SBS Sports
LIVE

2016 WBSU-23 여자 월드컵
니카라과 vs 대한민국
오늘 아침 **8:50**



WR 39.937
4.4

Niklaus [2]



[2] S. Niklaus, L. Mai, and F. Liu, "Video Frame Interpolation via Adaptive Separable Convolution," Proc. IEEE Int. Conf. Computer Vis. (ICCV), Venice, Italy, Oct. 2017.

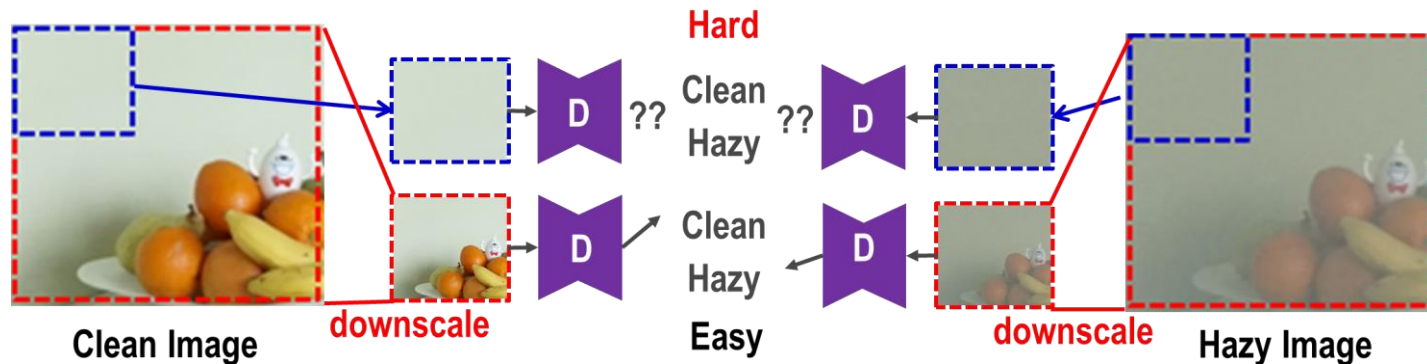
Proposed Hierarchical CNN-based VFI



0743 029776

5. Image Dehazing

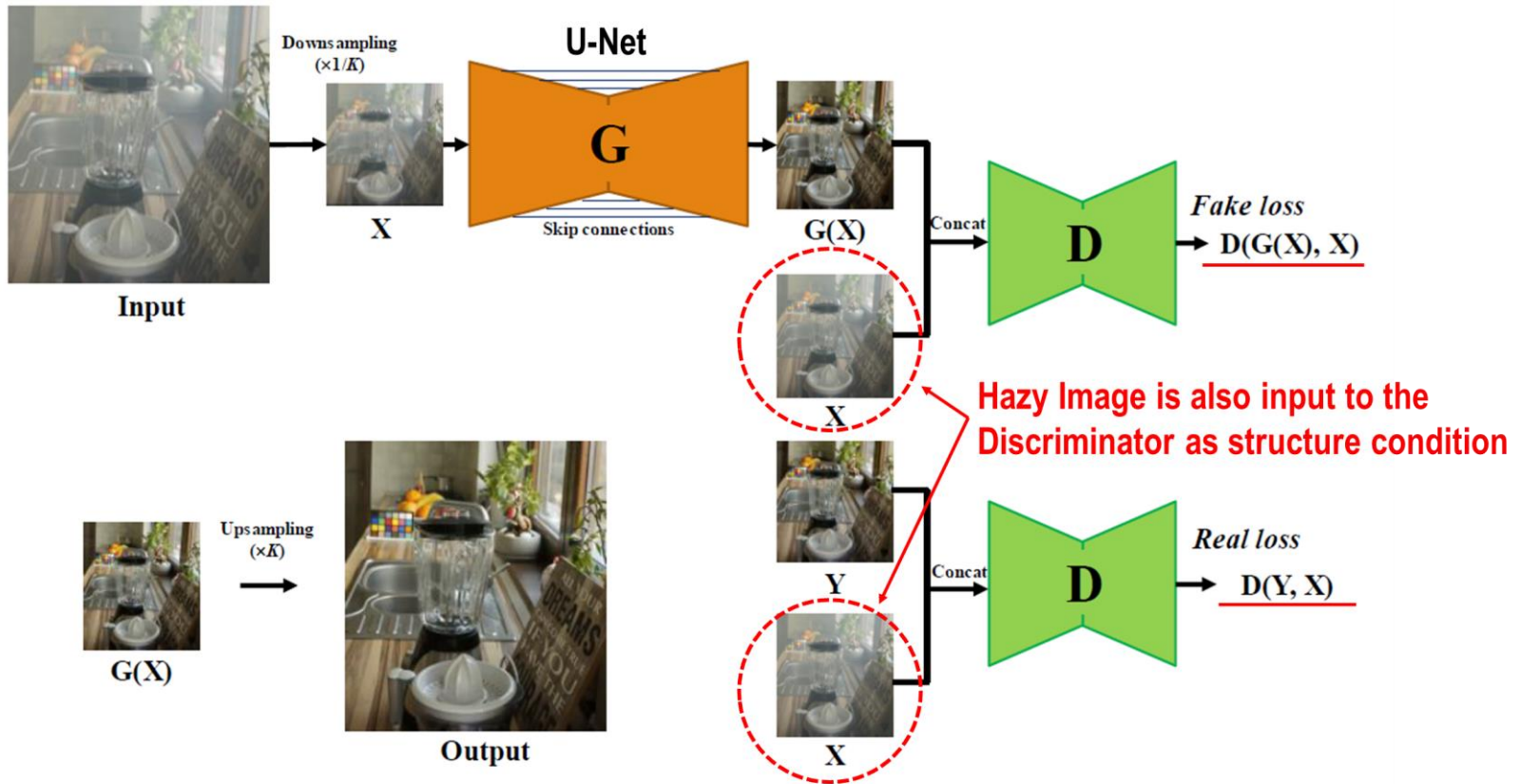
- ❖ Image haze is of the sources that not only lower image aesthetical values, but also cause significant performance degradation for object recognition.
- ❖ Thus, image dehazing is an essential preprocessing to both aesthetic photography and computer vision applications.
- ❖ **Our Approach**
 - Downscale of input hazy images for large receptive field sizes



- Reducing the input image sizes by using a simple bicubic interpolation method, and then extending the generated dehazed images by the same interpolation method.
 - Homogeneous region-dominant image (Indoor dataset) → Large downscale ratio($K\uparrow$)
 - Complex texture-dominant image (Outdoor dataset) → Small downscale ratio($K\downarrow$)

5. Image Dehazing

❖ Our Fully End-to-End learning based Conditional BEGAN for Image Dehazing



$$L_G = \lambda_{GAN} \underbrace{err D_{fake}}_{GAN \text{ loss}} + \lambda_{L1} \underbrace{|G(x) - y|}_{L1 \text{ loss}} + \lambda_{VGG} \underbrace{L_{VGG}}_{Perceptual \text{ loss}}$$

5.1 NTIRE Challenge on Image Dehazing

NTIRE 2018 Challenge on Image Dehazing: Methods and Results

Cosmin Ancuti, Codruta O. Ancuti, Radu Timofte, Luc Van Gool, Lei Zhang, Ming-Hsuan Yang, Vishal M. Patel, He Zhang, Vishwanath A. Sindagi, Ruhao Zhao, Xiaoping Ma, Yong Qin, Limin Jia, Klaus Friedel, Hyeonjun Sim, Jae-Seok Choi, Sooye Kim, Soomin Seo, Saehun Kim, Munchurl Kim, Ranjan Mondal, Sanchayan Santra, Bhabatosh Chanda, Jinlin Liu, Kangfu Mei, Juncheng Li, Luyao Faming, Fang Awen, Jiang Xiaochao, Qu Ting, Liu Pengfei, Wang Biao, Sun Jiangfan, Deng Yuhang, Zhao Ming, Hong Jingying, Huang Yzhi, Chen Ern, Chen Xiaoli, Yu Tingting, Wu Anli, Genc Deniz Engin, Hazim Kemal Ekenel, Wenzhe Liu, Tong Tong, Gen Li, Qi Quan, Gao Zhan, Li Daofa, Tang Yuling, Chen Ziyong, Huo Aitor, Alvarez-Gila, Adrian Gaidran, Alessandro Bria, Javier Vazquez-Corral, Marcelo Bertalmio, H. Seckin Demir, Omer Faruk Adil, Huynh Xuan, Phung Xin, Jin Jiale, Chen Chaowei, ShanZhibo Chen

Context

- In the last decade a significant amount of literature focused on single image dehazing research.
- Despite this growing interest, the field lacks standardized benchmarks to allow for evaluating objectively and quantitatively the performance of the existing dehazing techniques.
- The NTIRE 2018 challenge is a step forward in benchmarking single image dehazing. It uses two novel dataset (I-HAZE and O-HAZE) that can be employed as a more representative benchmark to assess dehazing algorithms in outdoor scenes, based on ground truth images.

Challenge phases

- Development (training) phase:** the participants got train data (hazy and haze-free images) (25 sets for the indoor track and 35 for the outdoor track);
- Validation phase:** the participants received 5 additional sets of images for both tracks and had the opportunity to test their solutions on the hazy validation images and to receive immediate feedback by uploading their results to the server. A validation leaderboard is available;
- Final evaluation (test) phase:** the participants got the hazy test images (5 sets for each track) and had to submit both their dehazed image and a description of their methods before the challenge deadline. One week later the final results were made available to the

Results and final ranking

Team	user	Track 1: Indoor		Track 2: Outdoor	
		PSNR	SSIM	PSNR	SSIM
Scarlet Knights[36]	rutgersprinter	24.973 ₁	0.881	24.029 ₁	0.775
RJTU	Team T-brain	22.966 ₁	0.857	24.538 ₁	0.777
FKS	fs	20.354 ₇	0.829	24.232 ₂	0.687
KAIST-VICLAB[27]	hjsim	20.911 ₂	0.751	23.180 ₆	0.705
Ranjanisil[2]	ranjanisil	22.421 ₄	0.852	22.705 ₉	0.707
KAIST-VICLAB[17]	Team KAIST-VICLAB	19.860 ₈	0.747	22.997 ₄	0.701
Ranjanisil[2]	ccsuplab	17.622 ₁₂	0.817	23.207 ₅	0.770
Dg-hifriends	liuj09	20.549 ₆	0.803	20.230 ₁₈	0.722
CLFStudio	Team CLFStudio	18.543 ₉	0.808	23.124 ₂	0.755
MLPhoenix	MLLab	16.653 ₁₃	0.787	22.080 ₁₆	0.731
Xdhm	hly_2	18.305 ₁₀	0.809	19.628 ₁₄	0.674
MLPhoenix	qxc	18.254 ₁₁	0.800		
SIMT-Lab[10]	Team DeAn			21.750 ₁₁	0.717
Xdhm	Tia			20.291 ₁₂	0.683
IVLab	Team IVLab			17.547 ₁₇	0.652
CLEAR	lizhangray	16.544 ₁₄	0.787	18.123 ₁₆	0.675
Dehazing-by-retinex[12]	shuffle	15.147 ₁₇	0.751	18.494 ₁₅	0.669
AESLSAN	DemirAdil	15.292 ₁₆	0.746		
AHappyFacel	Team HCLiLab	14.787 ₁₈	0.748		
Xdhm	hly			16.527 ₁₈	0.616
Xdhm	IMCL			15.784	0.634
hazy images	baseline	12.876	0.704		

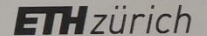
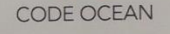
Table 1. NTIRE 2018 Challenge dehazing results and final rankings on indoor (I-HAZE) and outdoor (O-HAZE) test data.

Reported runtimes and platforms

Team	Runtime [s]		Platform	CPU/GPU (at runtime)	Architecture
	Track1: Indoor	Track2: Outdoor			
Scarlet Knights	10.0	10.0	pytorch	Titan X	Multi-scale Dense-residual Net
RJTU	6.4	9.9	Python 2.7, Pytorch	Titan X	DenseNet
FKS	11.5	10.3	Keras with tensorflow	Nvidia GTX1080Ti	
KAIST-VICLAB	3.9	4.8	Matlab, Python 3.6, Pytorch	Titan X	CBEGAN
Ranjanisil	5.7	4.8	Keras with tensorflow backend	Nvidia GeForce GTX 745	Fully Convolutional Net (FCN)
KAIST-VICLAB	9.6	5.2	Python 2.7, Tensorflow	CUDA v8.0, CuDNN 6	pix2pixGAN
CLFStudio	10.0	10.0	pytorch	Titan X1	
Dg-hifriends	6.0	8.0	Tensorflow	Nvidia GTX1080Ti	
MLPhoenix	19.7	17.4	pytorch	Nvidia GTX1080Ti	
Xdhm	1.7		pytorch		pix2pix and IRCNN
SIMT-Lab	8.0	8.0	Tensorflow	Titan X	Cycle-GAN
IVLab			Tensorflow	Titan X	
CLEAR	207.4	382.0	Matlab	(CPU)	
Dehazing-by-retinex	37.0	37.0	C++	(CPU)	
AESLSAN	30.0	30.0	Matlab	(CPU)	
AHappyFacel	21.8	18.7	Tensorflow, caffe2 with python	Titan X	
IMCL		4.5	Tensorflow	Nvidia GTX1080Ti	DDC-GAN

Table 2. Reported runtimes per image on I-HAZE and O-HAZE test data and details from the factsheets

Sponsors



New Trends in Image Restoration and Enhancement Workshop
Conference on Computer Vision and Pattern Recognition, Salt Lake City, June 18th 2018

presents
4th Place Award

in
NTIRE 2018 Challenge on Image Dehazing

to the Authors

**Hyeonjun Sim, Sehwan Ki, Jae-Seok Choi, Sooye Kim,
Soomin Seo, Saehun Kim, Munchurl Kim**

Korea Advanced Institute of Science and Technology (KAIST), Korea



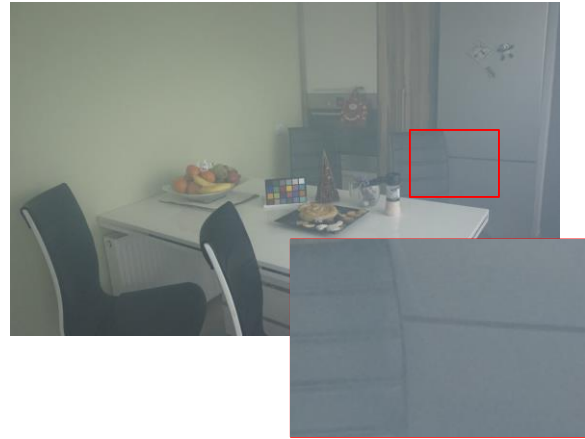
considered the Peak
ed in decibels (dB)
(SSIM) computed
truth.

participants on average
the final phase and
s, and factsheets.
Its of the challenge
the major details for

5.2 Experimental Results



(a) Clean



(b) Hazy



(c) pix2pix



(d) BEGAN



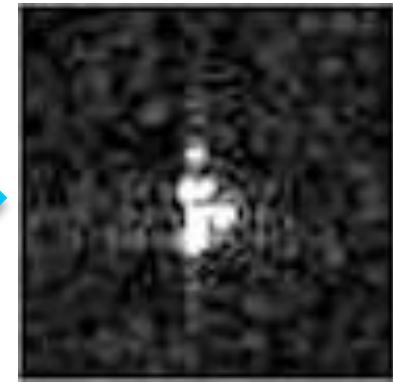
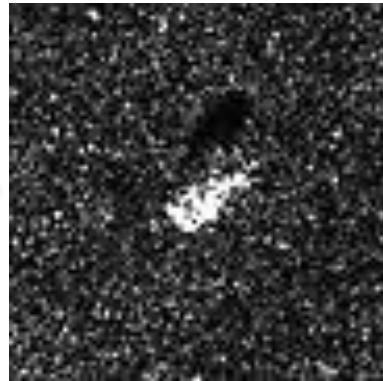
(e) CBEGAN

Dehazing Methods	PSNR (dB)	SSIM
DehazeNet [TIP 16]	21.14	0.8472
MSCNN [ECCV 16]	17.57	0.8102
AOD-net [ICCV 16]	19.06	0.8504
Proposed (e)	22.07	0.9187

6. Automatic Target Detection/Recognition

❖ Synthetic aperture radar(SAR)

- Can operate in all weather conditions, day and night
- Back-scattering of electromagnetic-wave
- Difficult to interpret intuitively



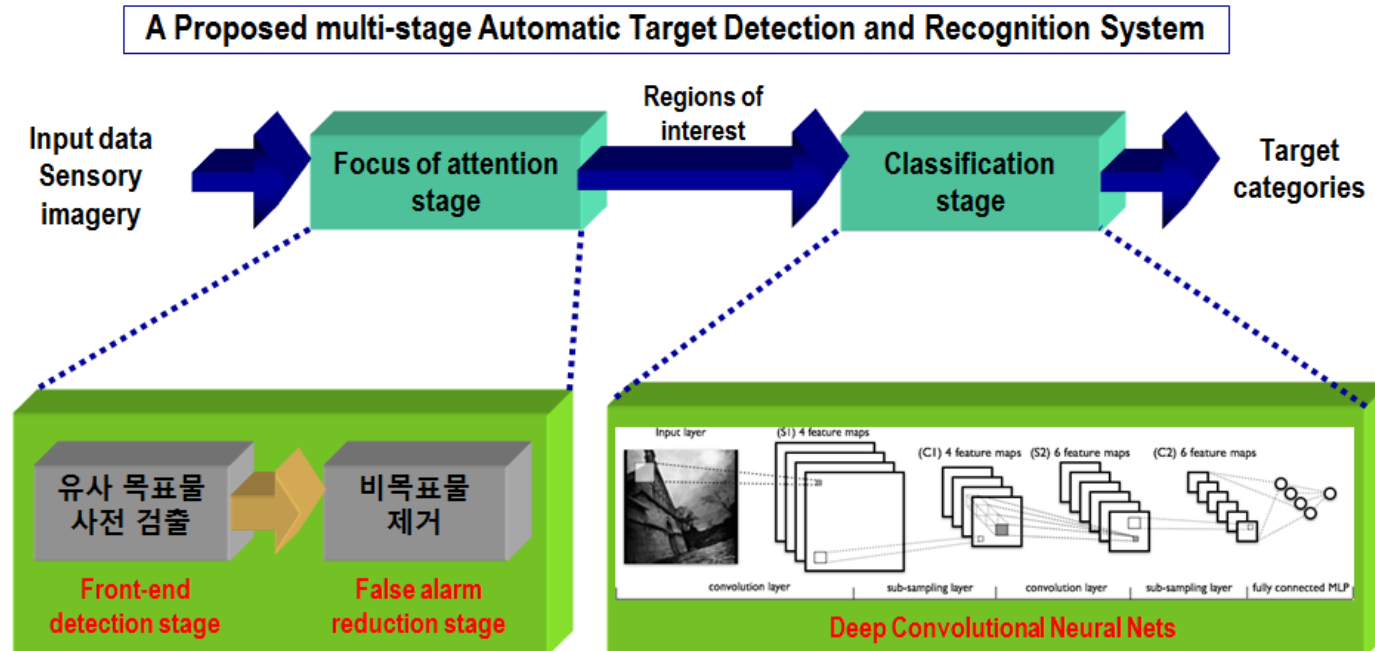
❖ Decision based on expert group

- Huge expense and time

❖ Needs of recognize and analysis algorithm for SAR

6. Automatic Target Detection/Recognition

❖ Deep CNN based Approach with Multi-Stage ATD/R



- Target Candidate Detection
- Discrimination of False Targets
 - Low-Complexity-CNN based Discriminator
- Target Classification and Recognition
 - Deep CNN based Classifier

6. Automatic Target Detection/Recognition

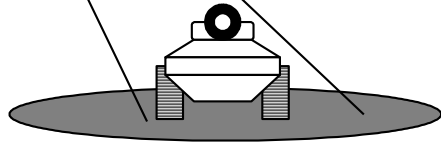
❖ Deep CNN based Approach with Multi-Stage ATD/R

■ Data Augmentation of ISAR Target Data

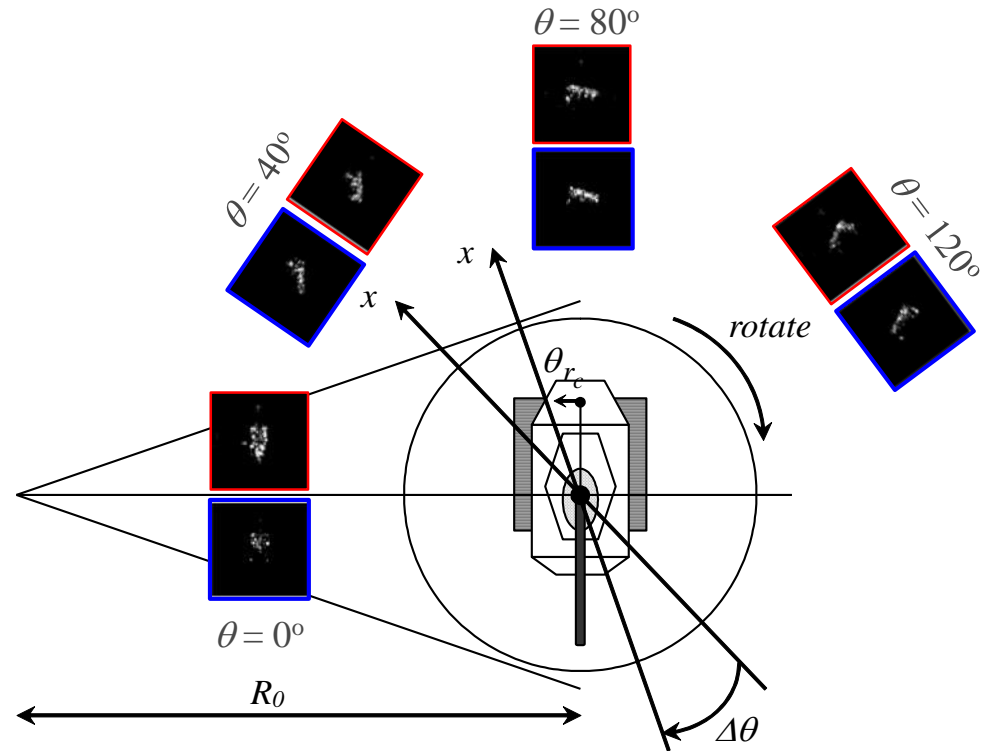
- Translation
- Speckling Noise
- Pose (azimuth angle) Synthesis



Stationary radar
DA: depression angles



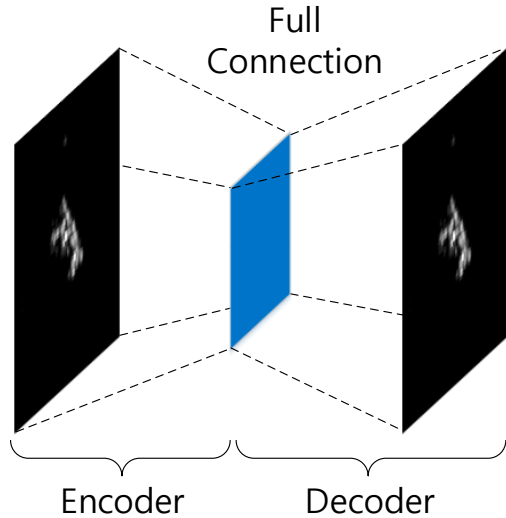
Stationary radar



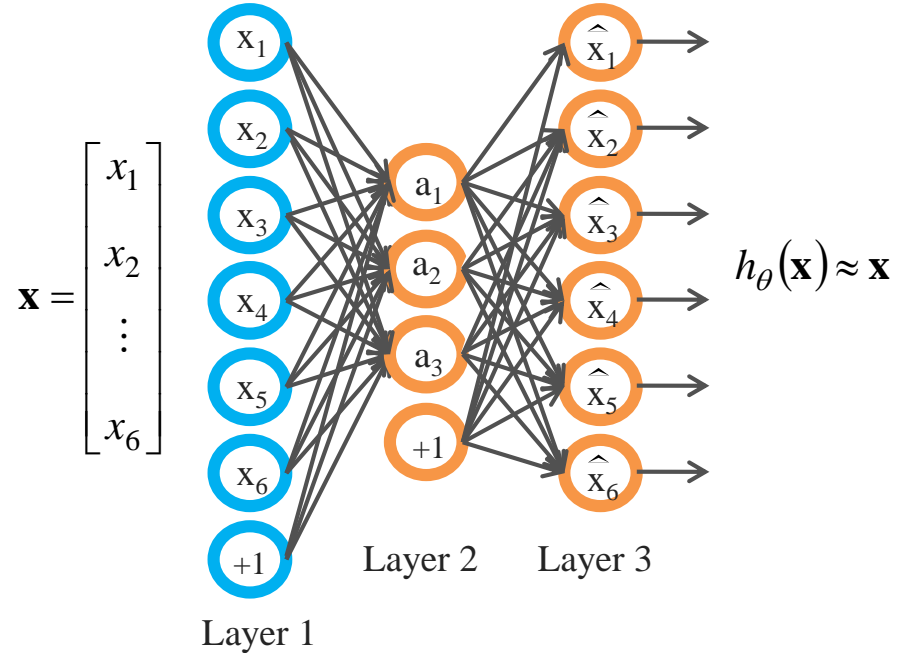
6.1 Automatic Target Detection/Recognition

❖ Deep CNN based Approach with Multi-Stage ATD/R

- Feature Learning of ISAR Target Data
 - Sparse Autoencoder for low-complexity and fast processing



$$\min_{\theta} \underbrace{\|h_{\theta}(\mathbf{x}) - \mathbf{x}\|^2}_{\text{reconstruction error term}} + \lambda \cdot \underbrace{\sum_i |a_i|}_{L_1 \text{ sparsity term}}$$

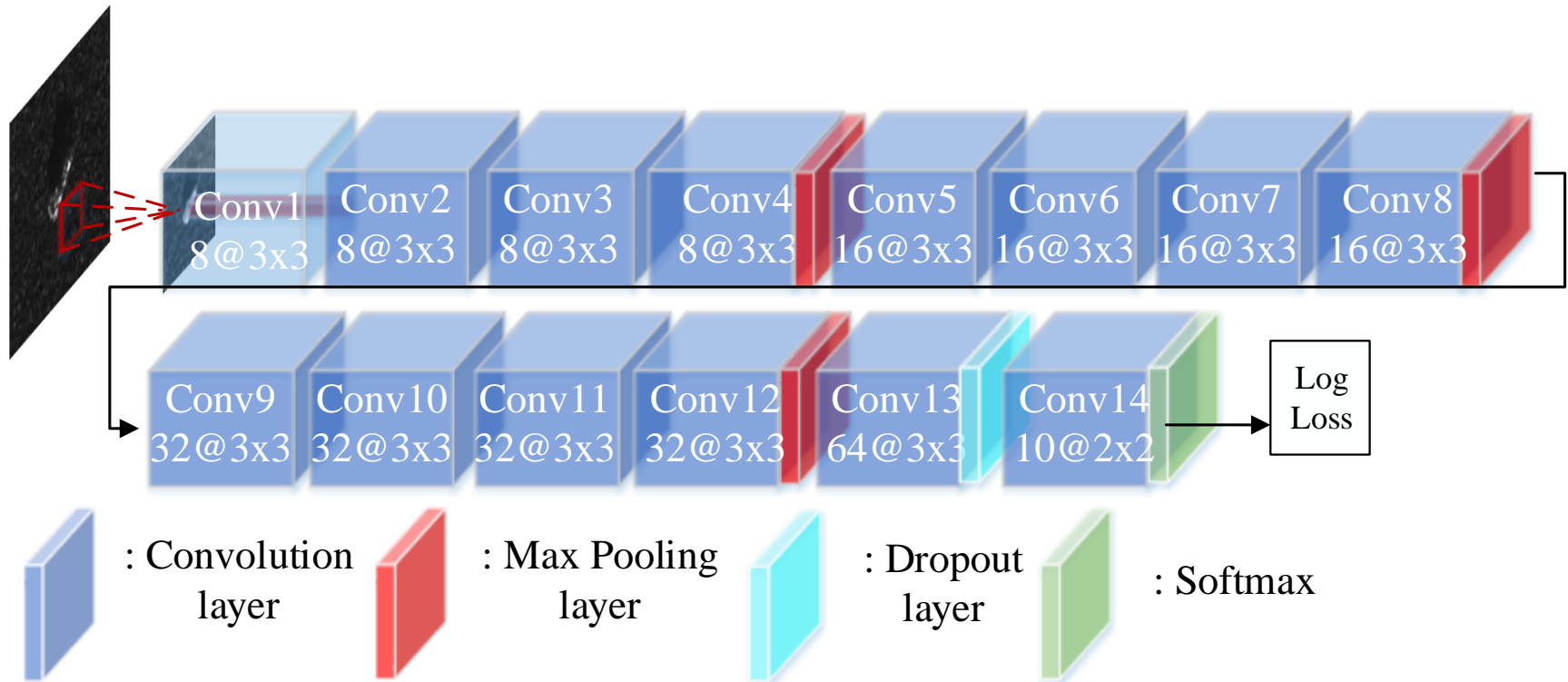


6.2 Automatic Target Detection/Recognition

❖ Deep CNN based Approach with Multi-Stage ATD/R

- Deep CNN based Classifier for SAR Target Recognition

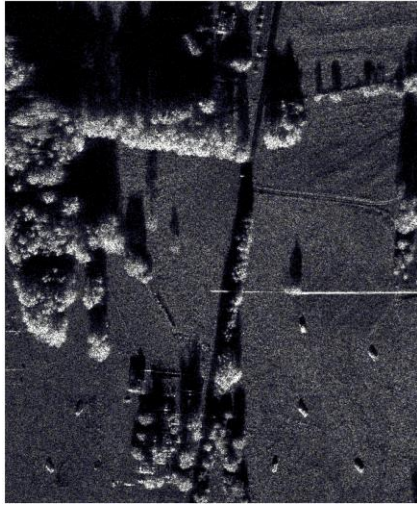
- 99.42% accuracy for 10 class classification (15 layers)



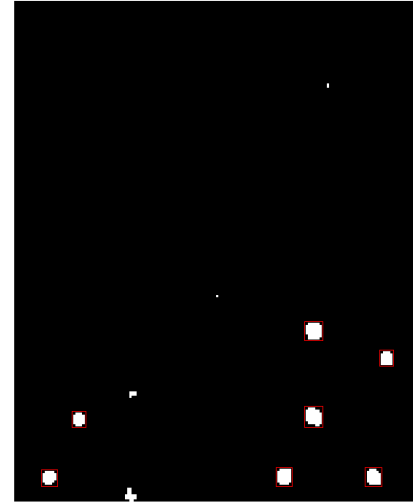
6.3 실험 결과 - ATD/R

❖ Deep CNN based Approach with Multi-Stage ATD/R

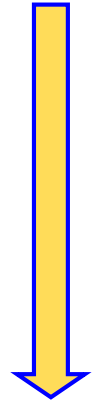
Input
SAR data



Target Detection
Network



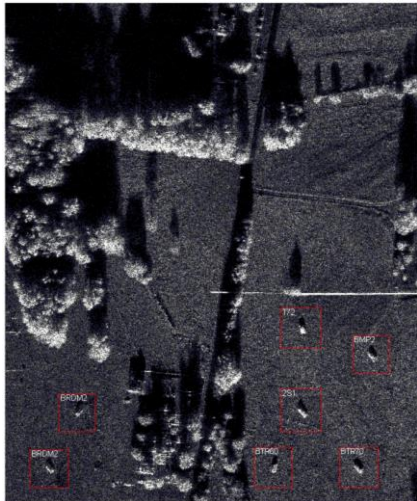
<Result of target
detection network >



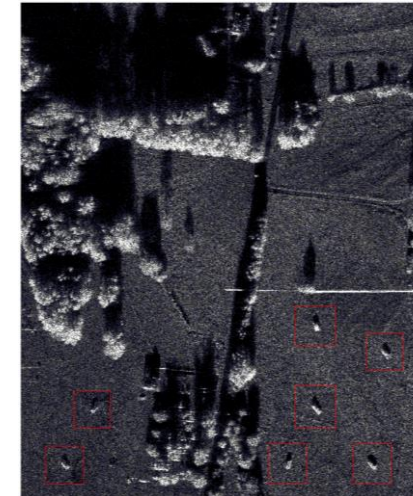
Receptive fields

Target
후보 영역 검출

SAR ATD-ATR
Output

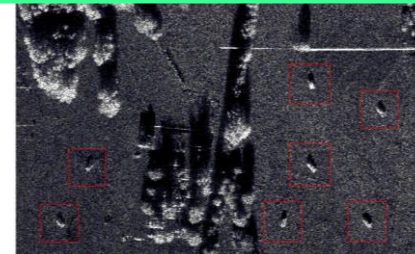
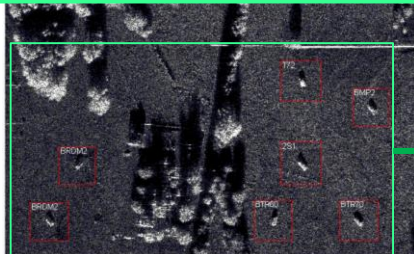
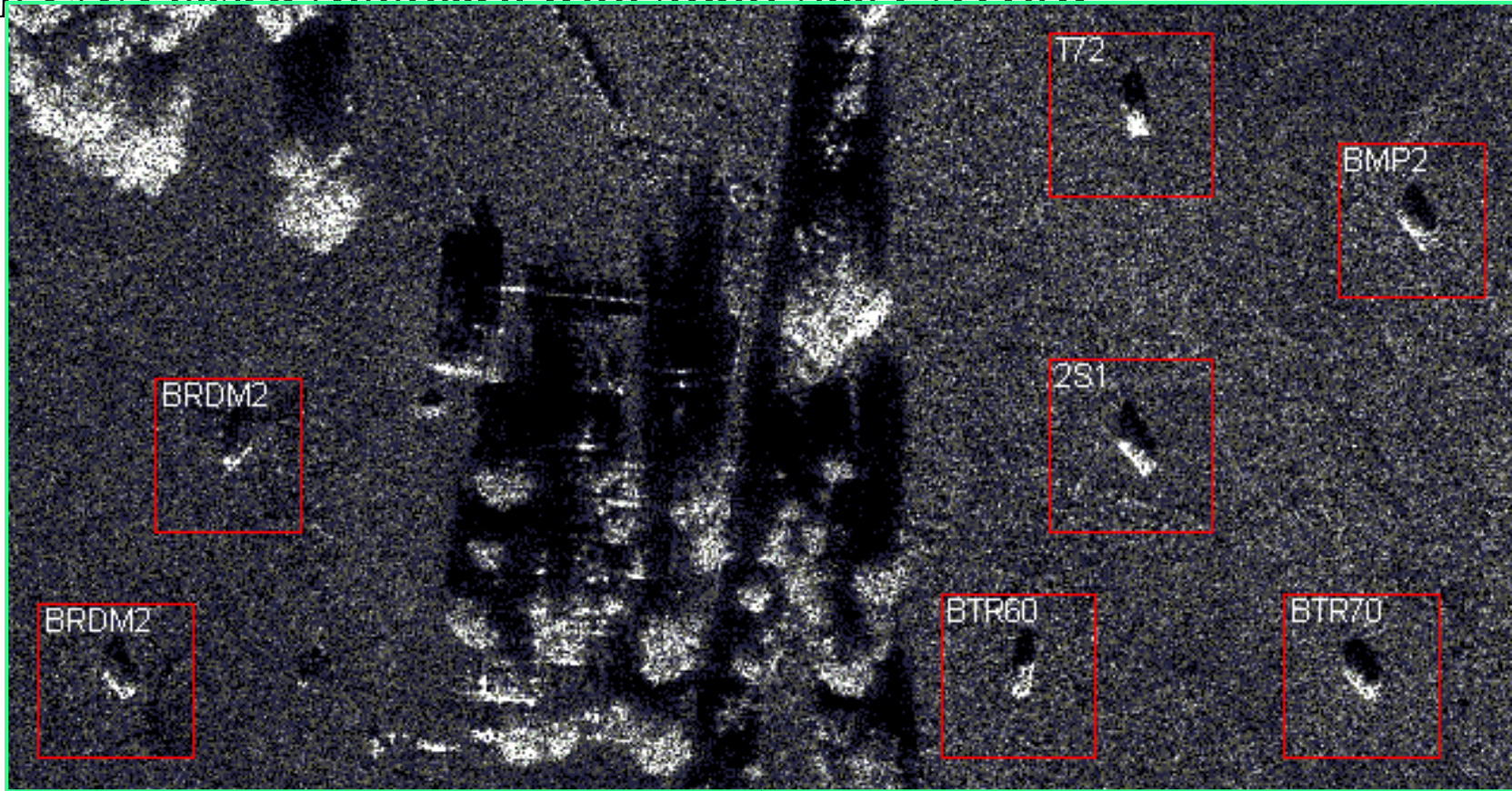


Target Recognition
Network



6.3 실험 결과 - ATD/R

❖ Deep CNN based Approach with Multi-Stage ATD/R



<SAR ATD-ATR 결과>

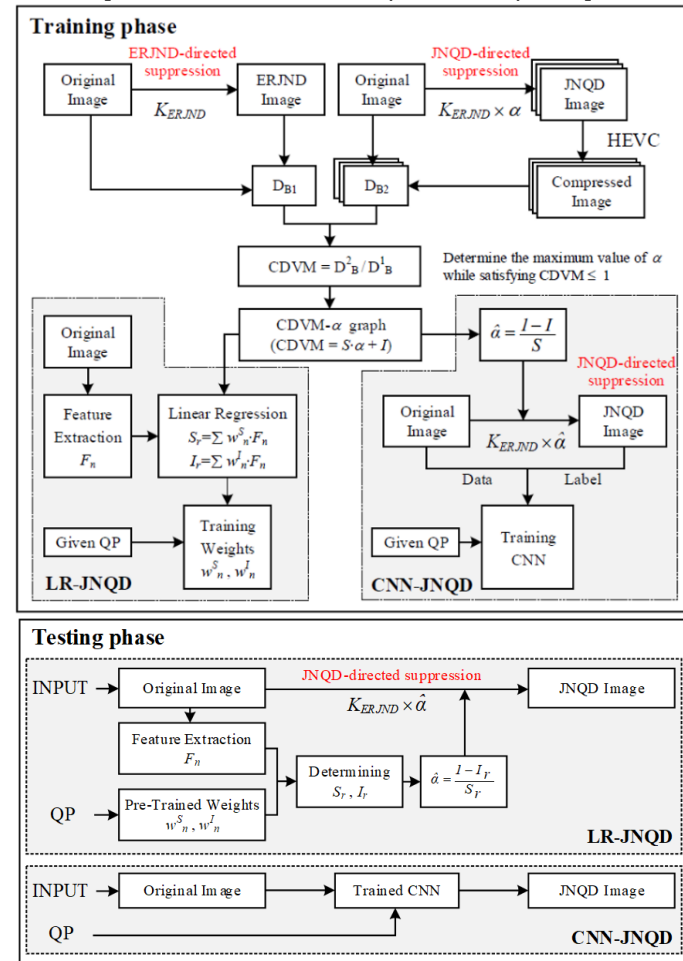
<Target 후보 영역 검출>

7. 영상 압축을 위한 최소 왜곡 인지 경계 모델링

● JNQD(Just-Noticeable-Quantization-Distortion)

- ▶ 영상 복잡도와 양자화 파라미터에 적응적인 JNQD 모델 기반 영상 압축 전처리 기법 개발
 - 영상 압축과 관련된 최소 왜곡 인지 경계 실험을 통해 얻은 왜곡 경계 값을 영상 압축의 전처리 과정에 양자화 파라미터 값에 적응적으로 적용되어 동일 주관적 화질 대비 높은 압축 효과를 보임.

[JNQD 모델링 알고리즘 (학습, 적용) 구조]



JNQD 모델 기반 전처리 방법

- ▶ 영상 압축에서 발생하는 왜곡을 고려한 실험을 통해 에너지 감소 왜곡 기반 최소 인지 왜곡 (ERJND: Energy Reduced JND) 값을 모델링.
- ▶ 영상의 DCT Block 별로 영상 압축으로 인해 발생하는 왜곡을 추정하여, 압축 후 복원 영상의 주관적 화질이 유지되는 인지 왜곡 경계 값 (JNQD)을 결정.
 - JNQD를 결정할 때 **선형 회귀 방법 (LR-JNQD)**과 **컨볼루션 신경망 방법 (CNN-JNQD)**을 사용.
- ▶ 입력영상의 주파수 성분을 Block 별로 결정된 JNQD 값만큼 줄여서 기존의 압축 과정에 입력.

기존 최소 인지 왜곡 기반 주관적 영상 압축 대비 우수성

- ▶ **높은 양자화 파라미터에서도 높은 압축률 향상**
 - 큰 양자화 파라미터인 경우, 잔차 신호가 이미 양자화에 의해 제거되어 기존 방법으로는 추가적으로 압축 효율을 높일 수 없는 반면 제안 방법의 경우 추가적인 압축률 향상을 얻을 수 있음.
- ▶ 표준 압축의 **추가적인 수정 없이** 기존의 표준 압축에 적용 가능
 - 전처리 과정에서 적용하여 인지 민감도에 따라 정보를 제거하는 방식으로 압축 코덱 내부의 추가적인 수정 없이 표준 코덱에 적용이 가능.
- ▶ 양자화 파라미터에 **적응적으로 전처리 강도 조절 가능**
 - 기존의 전처리 방법의 경우, 영상 압축에서 발생하는 왜곡을 추가로 고려하지 않은 것과 다르게 제안 방법은 양자화 파라미터의 값에 따라 다르게 발생하는 압축 왜곡을 고려하여 전처리 강도를 조절 가능함.

6.1 실험 결과 - 최소 왜곡 인지 경계 모델링

● JNQD 모델을 HEVC의 전처리에 적용한 예

- 원본 압축 영상 (HM 11.0, All Intra) 대비 주관적 압축 영상 화질 및 압축률 결과 비교
- $R = 100 * (1 - ((\text{원본 압축 영상의 비트량}) - (\text{주관적 압축 영상의 비트량})) / (\text{원본 압축 영상의 비트량}))$
- $\text{DMOS} = (\text{원본 영상의 MOS}) - (\text{주관적 압축 영상 MOS})$, MOS : mean opinion score

Sequence (sizes)	QP	R (%)			DMOS (std)		
		Bae's JND	LR-JNQD	CNN-JNQD	Bae's JND	LR-JNQD	CNN-JNQD
BQTerrace (1920x1080)	22	39.96	37.35	42.25	-0.7(0.9)	0.1(0.7)	0.0(0.8)
	27	16.43	22.45	28.93	-0.6(1.0)	-0.5(0.4)	-0.4(0.7)
	32	0.02	12.00	19.88	-0.4(1.0)	0.0(1.0)	0.2(0.6)
	37	0.58	2.21	10.84	-0.8(0.7)	0.0(0.4)	0.3(0.6)
Cactus (1920x1080)	22	27.94	23.96	27.81	-0.8(1.1)	-0.2(0.9)	-1.0(1.2)
	27	8.33	13.16	19.52	-0.7(0.8)	-0.2(0.9)	-0.6(0.9)
	32	1.52	5.72	13.48	-0.5(1.7)	-0.4(1.1)	-0.5(1.1)
	37	0.39	1.20	6.40	-0.2(0.9)	-0.1(1.2)	0.4(0.9)
ParkScene (1920x1080)	22	24.28	21.16	21.50	-0.7(1.0)	-0.1(0.7)	-0.4(0.9)
	27	7.95	14.26	19.64	-0.3(0.7)	0.0(0.4)	-0.1(0.6)
	32	1.97	5.66	14.82	-0.4(0.7)	0.1(0.7)	-0.7(0.8)
	37	0.55	0.27	7.18	0.1(0.7)	-0.2(0.9)	-0.1(2.4)
BQMall (832x480)	22	23.31	19.29	22.07	-0.7(0.9)	-0.6(1.4)	-0.5(0.7)
	27	9.75	15.17	20.44	-1.0(1.1)	-0.2(0.8)	0.1(0.9)
	32	0.55	8.90	15.98	0.1(0.9)	0.0(0.9)	0.0(1.0)
	37	0.32	1.60	8.99	0.3(1.0)	0.4(1.4)	0.7(1.3)
PartyScene (832x480)	22	32.85	22.46	31.24	-0.8(0.9)	-0.4(1.7)	0.0(1.1)
	27	16.88	21.05	32.19	-0.6(0.9)	-0.3(1.4)	-0.5(0.9)
	32	0.91	15.53	29.21	-0.5(1.0)	-0.1(1.1)	-0.4(1.1)
	37	0.40	3.08	19.95	-0.2(1.0)	-0.2(1.3)	-0.7(1.0)
RaceHorses (832x480)	22	26.75	17.19	24.36	-0.2(1.0)	-0.3(1.0)	-0.2(0.9)
	27	15.64	17.25	26.23	-0.7(0.9)	-0.1(1.1)	-0.3(0.9)
	32	0.57	10.62	23.79	-0.2(0.6)	0.1(1.1)	-0.2(0.5)
	37	0.35	1.63	14.42	-0.4(0.8)	-0.3(1.4)	-0.2(0.8)
Avg.		10.76	13.05	20.88	-0.5(0.9)	-0.2(1.0)	-0.2(0.9)



7.1 실험 결과 - 최소 왜곡 인지 경계 모델링

Original HM 11.0 (32.3 Mbps)

QP = 22, Random Access

Proposed PVC (CNN) (15.1 Mbps)



1번 영상 53% 비트율 감소, 2번 영상 33% 비트율 감소

방송·스마트미디어ICT기초연구실

감사합니다.



Q&A

김문철 교수 (mkimee@kaist.ac.kr)
한국과학기술원 전기및전자공학부

