

# **NTIRE 2021 Challenge on High Dynamic Range Imaging: Dataset, Methods and Results**

**박예인**

*Vision & Display Systems Lab.*

*Dept. of Electronic Engineering, Sogang University*

# Outline

- NTIRE 2021 Challenge 소개
- Background
  - HDR 개념
  - HDR 복원 연구의 필요성
  - HDR 복원 연구
- NTIRE 2021 Challenge on High Dynamic Range Imaging: Dataset, Methods and Results
  - Dataset
  - Results
  - 참고) “Attention-guided Network for Ghost-free High Dynamic Range Imaging”, CVPR, 2019
  - Ranking 1 of Track 2 : MegHDR
  - Ranking 2 of Track 2 : SuperArtifacts
- Conclusion

# NTIRE 2021 Challenge

- **N**ew **T**rends in **I**mage **R**estoration and **E**nhancement workshop and challenges on image and video processing in conjunction with CVPR 2021
- Computer vision task (image restoration, enhancement, manipulation)
  - 자동차 산업, 전자, 원격 감지, 의료 이미지 분석 등과 같은 분야에서 응용 범위 증가
  - 모바일 및 웨어러블 장치 사용
- 이 workshop은 해당 분야의 새로운 trend와 발전에 대한 개요를 제공하는 것을 목표로 함
  - Image/video inpainting, deblurring, denoising, upsampling and super-resolution
  - Image/video de-hazing, de-raining, de-snowing
  - Image/video compression, enhancement
  - Style transfer
  - Image/video restoration, enhancement, manipulation
  - Image/video semantic segmentation, depth estimation

# Background

- HDR 개념

- Dynamic range

- 대비의 최대값과 최소값의 비율

- ※ 가장 밝은 부분과 가장 어두운 부분 사이의 밝기 비율

- LDR (low dynamic range) : small dynamic range

- HDR (high dynamic range) : large dynamic range

- LDR 이미지에 비해 어두운 부분과 밝은 부분의 디테일이 살아있어 화면에 사실감을 더해 줌



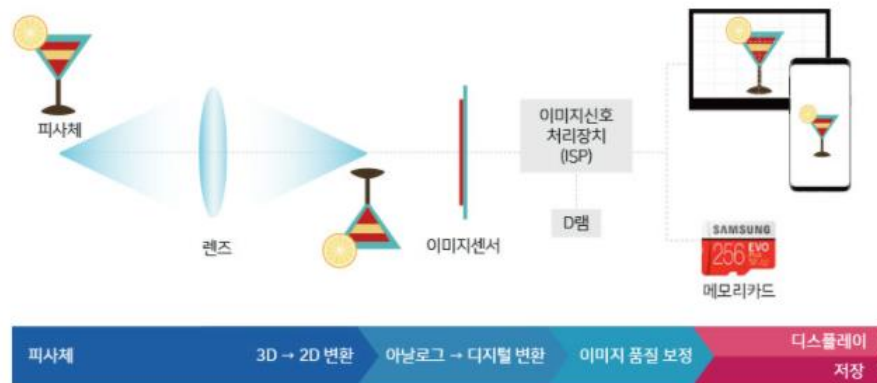
LDR



HDR

# Background

- HDR 복원 연구의 필요성
  - 일반 카메라의 이미징 센서의 고유한 한계
    - High-irradiance 영역의 포화와 low-light 영역의 판독 값의 불확실성
    - 단일 노출 샷으로 다양한 조명으로 장면을 획득하는데 있어 어려움
  - 고성능 카메라의 단점
    - 비싸고 부피가 크며 별도의 많은 장비를 필요로 함



출처 : 삼성전자

# Background

- HDR 복원 연구

- Single-image HDR reconstruction

- 다양한 노출 값을 가진 이미지들이 없이 구현 가능함

- Ghost artifact가 발생되지 않음

- Single 이미지는 multiple frame 이미지에 비해 불리한 점이 명확함

- ※ 상대적으로 적은 정보가 제공되므로 detail과 color 측면에서 복원력이 떨어짐

- ※ Multiple frame을 사용할 수 없거나 시간 제약으로 인해 multiple frame을 캡처할 수 없는 경우에 좋은 대안이 됨

- Multi-image HDR reconstruction

- Single 이미지 기반 보다 우수한 성능을 가짐

- ※ 다양한 노출 값을 가진 이미지들이 입력되기 때문에 상대적으로 많은 정보가 주어지고 이로 인해 detail과 color 측면에서 복원력이 높음

- Ghost artifact 존재

- ※ 화질 저하를 일으킴

---

# NTIRE 2021 Challenge on High Dynamic Range Imaging: Dataset, Methods and Results

# NTIRE 2021 Challenge on High Dynamic Range Imaging

- Challenge

- HDR image enhancement task를 다루는 첫번째 에디션
- HDR imaging에 있어 최신 기술을 측정하고 발전시키는 것을 목표로 함
- 카메라, 장면, 광원 측면에서 complex motion이 발생하는 challenging 시나리오에 중점을 둠

- Dataset

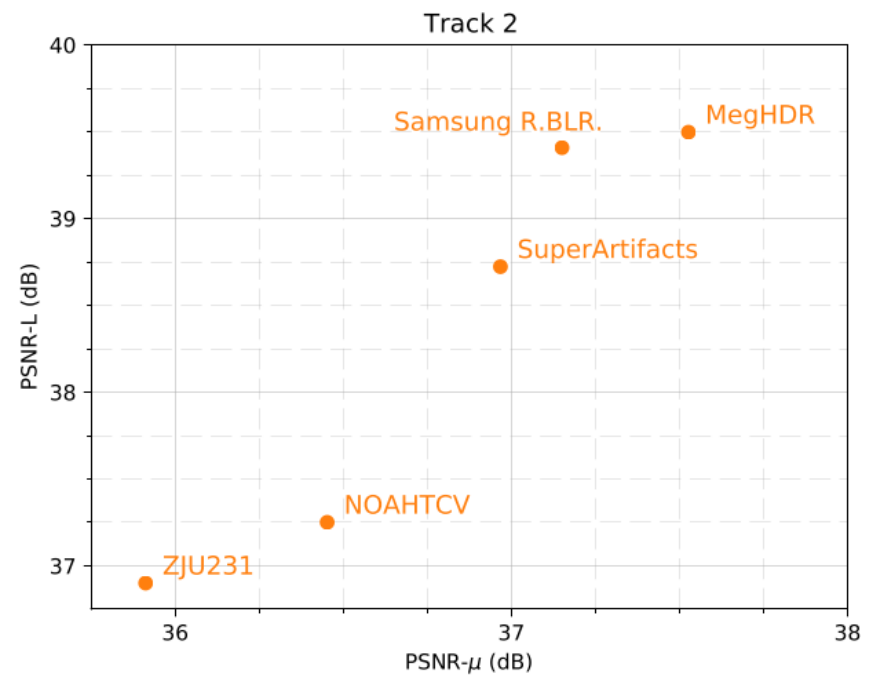
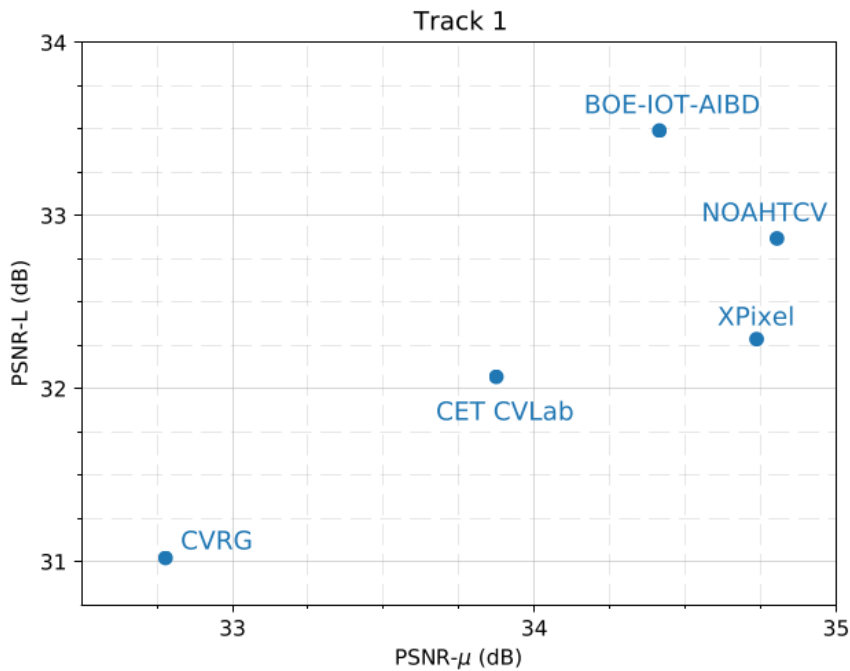
- 1500 training, 60 validation, 201 test 데이터로 구성
- Short, medium, long 노출 값을 가진 3개의 LDR 이미지들과 medium 노출 이미지에 맞게 aligned된 ground truth HDR image로 구성
- Challenging한 장면들이 포함됨
  - Moving light sources, brightness changes over time, high contrast skin tones, specular highlights and bright, saturated colors 등





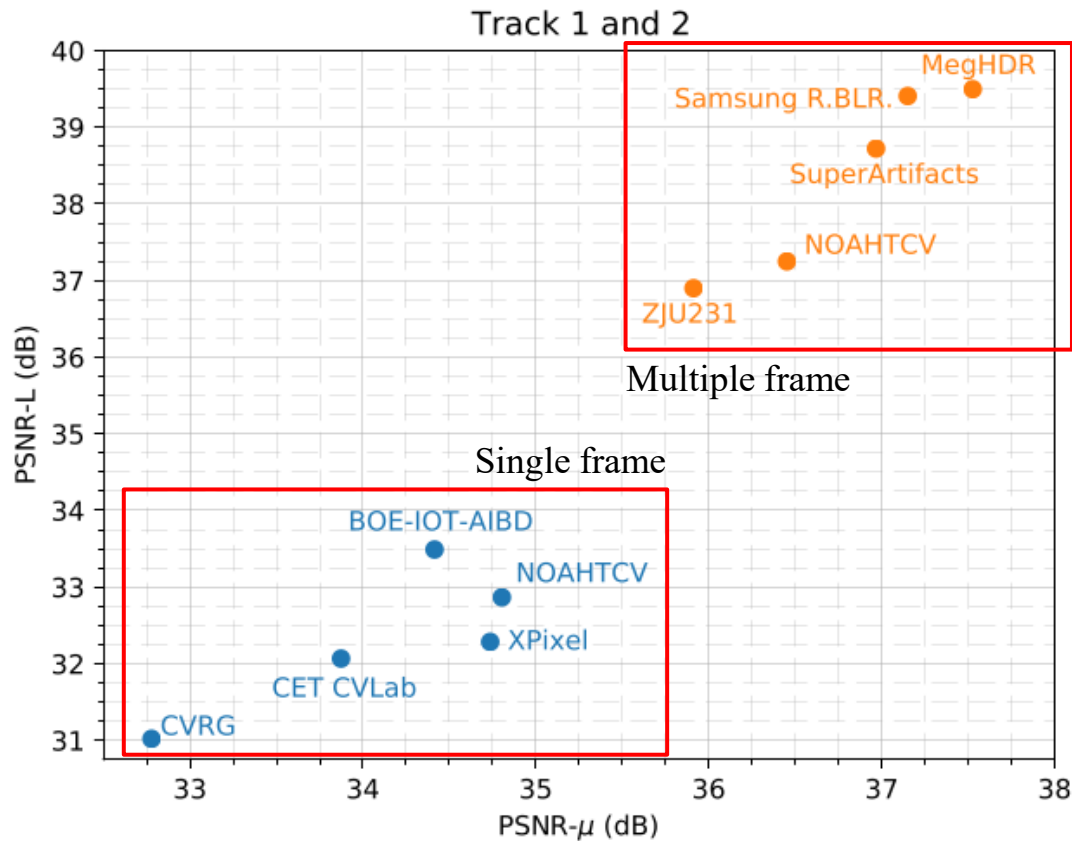
# NTIRE 2021 Challenge on High Dynamic Range Imaging

- Results
  - Track 1 (single frame) Track 2 (multiple frame)



# NTIRE 2021 Challenge on High Dynamic Range Imaging

- Results



# NTIRE 2021 Challenge on High Dynamic Range Imaging

- Results

Team	Username	PSNR- $\mu$	PSNR	Runtime (s)	GPU	Ensemble
NOAHTCV	noahcv	34.804 (1)	32.867 (2)	61.52 (5)	Tesla P100	flips, transpose
XPixel	Xy_Chen	34.736 (2)	32.285 (3)	0.53 (2)	RTX 2080 Ti	-
BOE-IOT-AIBD	chenguannan1981	34.414 (3)	33.490 (1)	5.00 (4)	Tesla V100	-
CET CVLab	akhilkashok	33.874 (4)	32.068 (4)	0.20 (1)	Tesla P100	flips, rotation
CVRG	sharif_apu	32.778 (5)	31.021 (5)	1.10 (3)	GTX 1060	-
<i>no processing</i>	-	25.266 (6)	27.408 (6)	-	-	-

< Results and rankings of methods submitted to the Track 1: Single frame HDR >

Team	Username	PSNR- $\mu$	PSNR	Runtime (s)	GPU	Ensemble
MegHDR	liuzhen	37.527 (1)	39.497 (1)	1.35 (3)	RTX 2080 Ti	flips, transpose
SuperArtifacts	evelynchee	36.968 (2)	38.723 (2)	3.80 (4)	RTX 2080 Ti	-
NOAHTCV	noahcv	36.452 (3)	37.250 (3)	1.26 (2)	Tesla V100	-
ZJU231	ZJU231	35.912 (4)	36.900 (4)	0.37 (1)	RTX 2080 Ti	flips, rotation, $\times 4$ models
Samsung Research Bangalore*	AnointedKnight	37.151	39.408	15.77	Tesla P40	flips, transpose
<i>no processing</i>	-	25.266 (5)	27.408 (5)	-	-	-

< Results and rankings of methods submitted to the Track 2: Multiple frames HDR >

# NTIRE 2021 Challenge on High Dynamic Range Imaging

- 요약

- 같은 evaluation metric (PSNR)과 같은 ground truth 이미지를 공유하는 두개의 다른 track으로 구성됨

- **Track 1: Single Frame**

- 단일 LDR 프레임만 사용하여 HDR 복원 정도를 평가
- 서로 다른 노출의 인접 프레임에서 정보를 활용할 수 없기 때문에 저/과 노출 영역을 복원할 때 더 어려움
- 해당 track에 참여한 팀들은 대부분 HDR to LDR pipeline을 reverse시키는 것\*을 목표로 하는 여러 sub network를 구성함

- **Track 2: Multiple Frame**

- 3개의 다른 노출 값을 가진 LDR 이미지들을 사용하여 HDR 복원 정도를 평가
- 해당 track에 참여한 팀들은 대부분 attention 메커니즘 기반의 AHDRNet\*\*을 사용함

# 참고) Attention-guided Network for Ghost-free High Dynamic Range Imaging\*

## • Overview of the AHDRNet Architecture

- **Attention network**와 Attention network에서 나온 features를 결합하는 **Merging network**로 구성

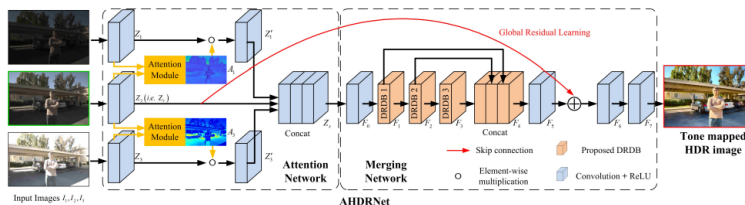
- **Attention network**

- 각각의 LDR 영상에서 특징 추출
- Reference 영상과 non-reference 영상 사이에서 attention map을 추출
- Attention map을 non-reference 영상과 pixel wise 곱을 통해 non-reference 영상의 필요 없는 영역을 제거

∴ 정제된 non-reference feature와 reference feature를 결합

- **Merging network**

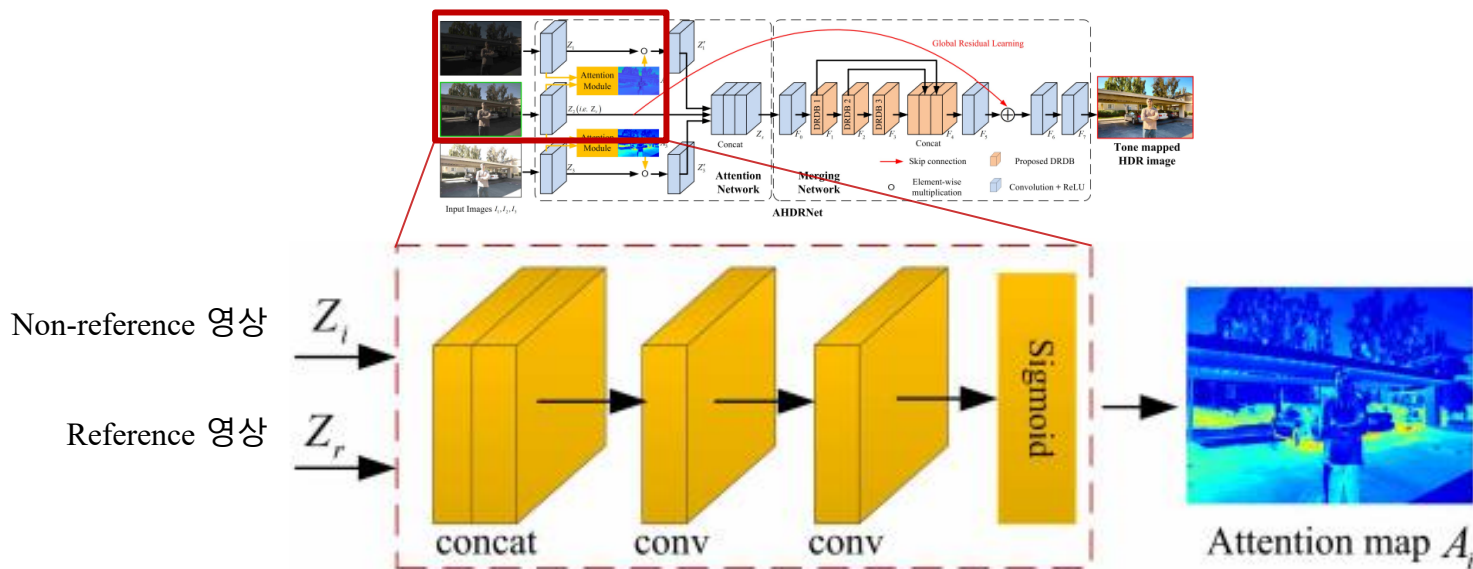
- 결합된 feature를 이용하여 HDR 수행
- 제안된 block인 Dilated residual dense block(DRDB)를 이용하여 복원된 HDR 영상의 디테일 성능을 향상



# 참고) Attention-guided Network for Ghost-free High Dynamic Range Imaging\*

- Attention Network for Feature Extraction

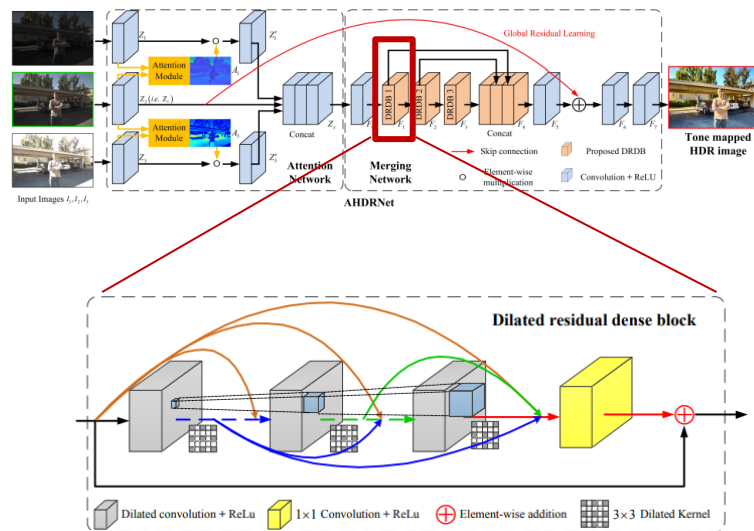
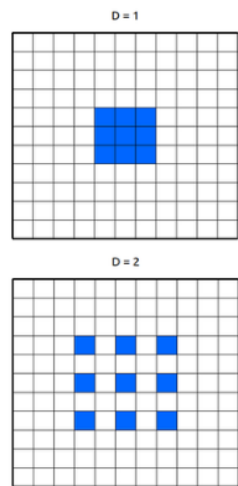
- Reference 영상과 non-reference 영상의 feature를 입력으로 받아 attention map을 형성시키는 network
- 2개의 feature를 concat한 입력을 convolution을 통해 1 채널의 attention map으로 구성
  - 출력을 sigmoid로 하여 어떤 픽셀이 중요한지 0~1사이의 값으로 출력



# 참고) Attention-guided Network for Ghost-free High Dynamic Range Imaging\*

## • Merging Network for HDR Image Estimation

- Dilated convolution을 이용한 block인 DRDB를 통해 영상의 디테일을 명확히 생성
  - Dilated rate가 2인 3\*3 커널은 9개의 파라미터를 사용하면서 5\*5 커널과 동일한 시야를 갖게 됨
  - 필터 내부에 zero padding을 추가하여 receptive fields를 늘리는 방법
- Reference feature를 가지고 오는 Global residual learning을 통해 reference의 대한 정보를 더 명확히 사용할 수 있도록 함



# Ranking 1 of Track 2 : MegHDR

- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”
  - Noise-free, ghost-free, hallucinating more accurate contents in over saturated areas
  - How to handle saturation and noise
    - Spatial attention 모듈 채택
      - ※ LDR 이미지들에서 가장 적절한 영역을 적응적으로 선택
  - How to tackle misalignment caused by object motion or camera jittering
    - 감마 보정된 이미지들을 feature level에서 Pyramid, Cascading and Deformable (PCD) alignment 모듈로 정렬
      - ※ LDR 이미지들과 감마 보정된 이미지들을 dual branch로 각각 처리
  - 즉, Spatial attention 모듈로 LDR 이미지들을 처리하고 PCD alignment 모듈로 감마 보정된 이미지들을 처리



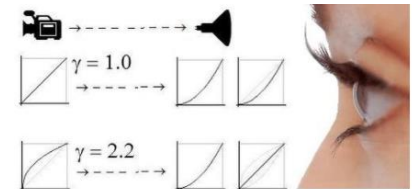


# Ranking 1 of Track 2 : MegHDR

- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”

- Gamma correction 적용

- 1) Non-linear한 LDR 이미지들을 linear space로 변환
- 2) Gamma correction을 통해 입력 LDR 이미지들을 HDR domain에 맵핑



∴ Gamma correction은 이미지를 우리가 눈으로 인식하는 것과 더 가까운 domain으로 맵핑시켜 줌

$$\check{I}_i = \frac{(I_i)^\gamma}{t_i}, \quad i = 1, 2, 3$$

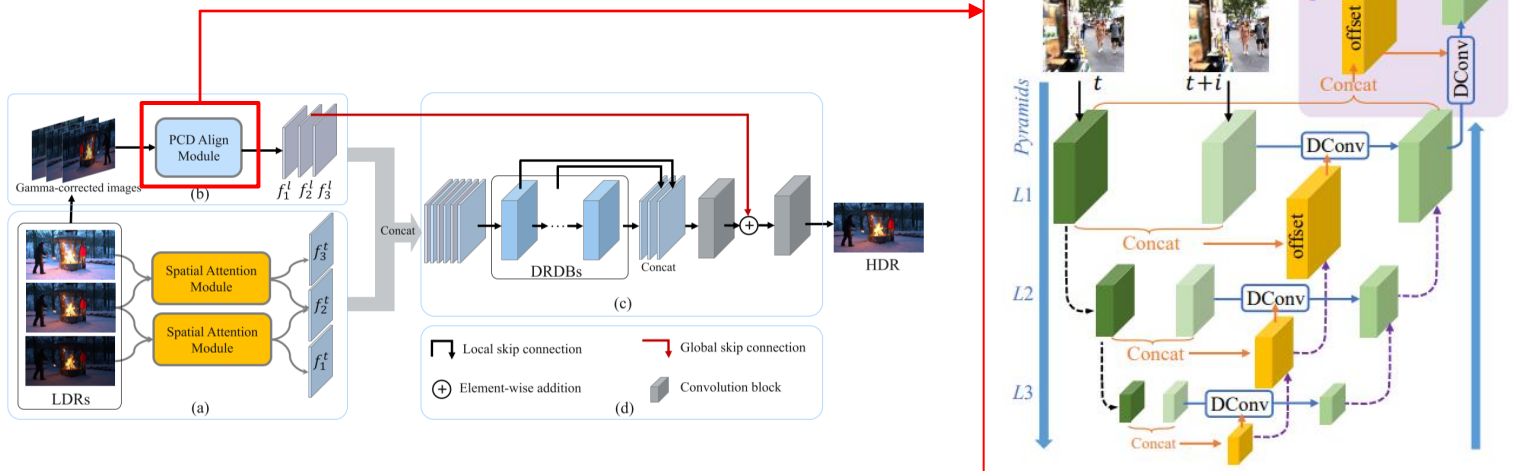
- $I_i$  : 3 LDR images ( $i = 1, 2, 3$ ),  $t_i$  : exposure time,  $\gamma$  : gamma correction parameter (2.2),  $\check{I}_i$  : gamma corrected images

$$I^H = f(\mathcal{A}(I_i), \mathcal{P}(\check{I}_i); \theta)$$

- Spatial attention module A에서는 LDR 이미지들을, PCD alignment module P에서는 감마 보정된 이미지들을 각각 처리하여 재구성된 HDR 이미지를 생성

# Ranking 1 of Track 2 : MegHDR

- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”
  - Pyramid, Cascading and Deformable (PCD) alignment 모듈
    - 감마 보정된 이미지들을 정렬시키기 위함
    - Feature level에서 정렬 수행
    - Convolution을 사용하여 pyramid feature를 추출하고, 각 feature의 scale에 대해 reference feature와 deformable alignment를 수행



# Ranking 1 of Track 2 : MegHDR

- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”

- Pyramid, Cascading and Deformable (PCD) alignment 모듈

- Deformable convolution

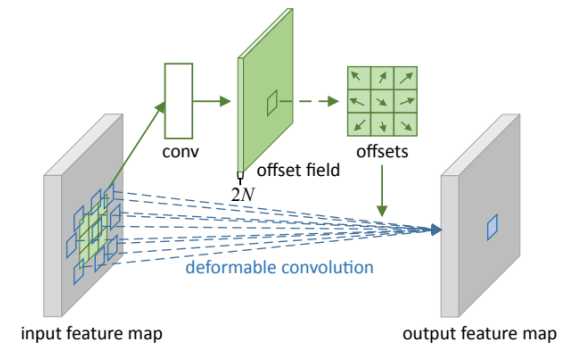
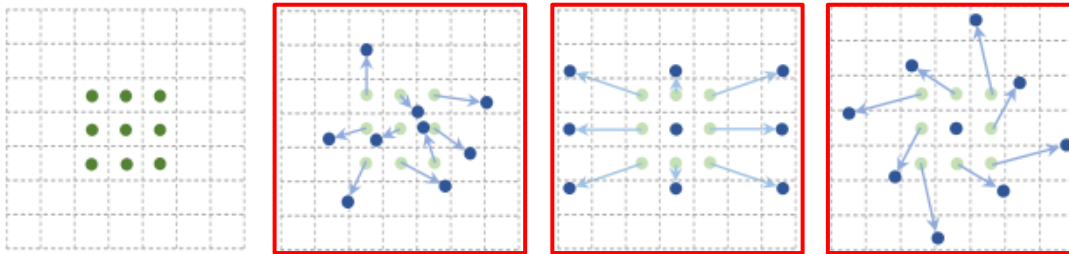
- ※ 기존 CNN은 고정된 커널 영역에서만 feature를 추출

- ✓ 3\*3 conv filter를 사용하면 3\*3 receptive field에서만 feature 추출

- ※ 고정된 receptive field에서만 feature를 추출하는 것이 아니라, 좀 더 flexible한 영역에서 feature를 추출하는 방법

- ✓ Offset을 계산하는 conv layer

- ✓ Offset 정보를 받아 conv 연산을 수행해서 output feature map을 생성



# Ranking 1 of Track 2 : MegHDR

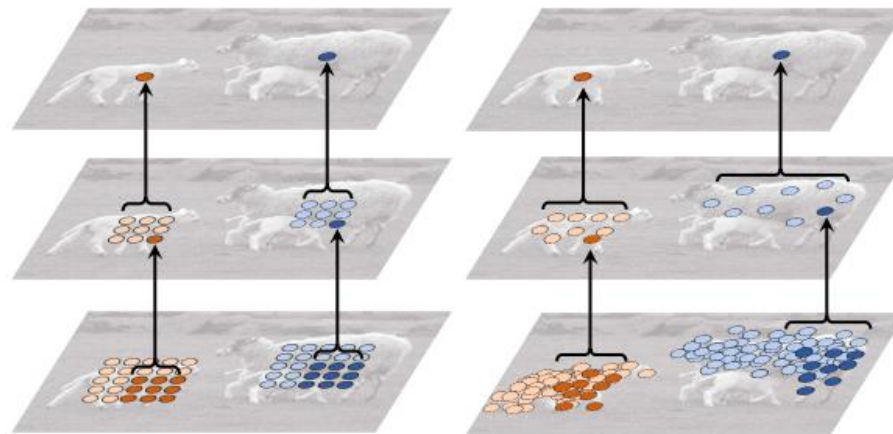
- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”

- Pyramid, Cascading and Deformable (PCD) alignment 모듈

- Deformable convolution

- ※ 아래 그림을 통해, object scale에 따라 receptive field가 다른 것을 확인할 수 있음

- ※ 반면에, 기존 convolution은 object scale과 관계없이 고정된 영역에서만 feature 추출

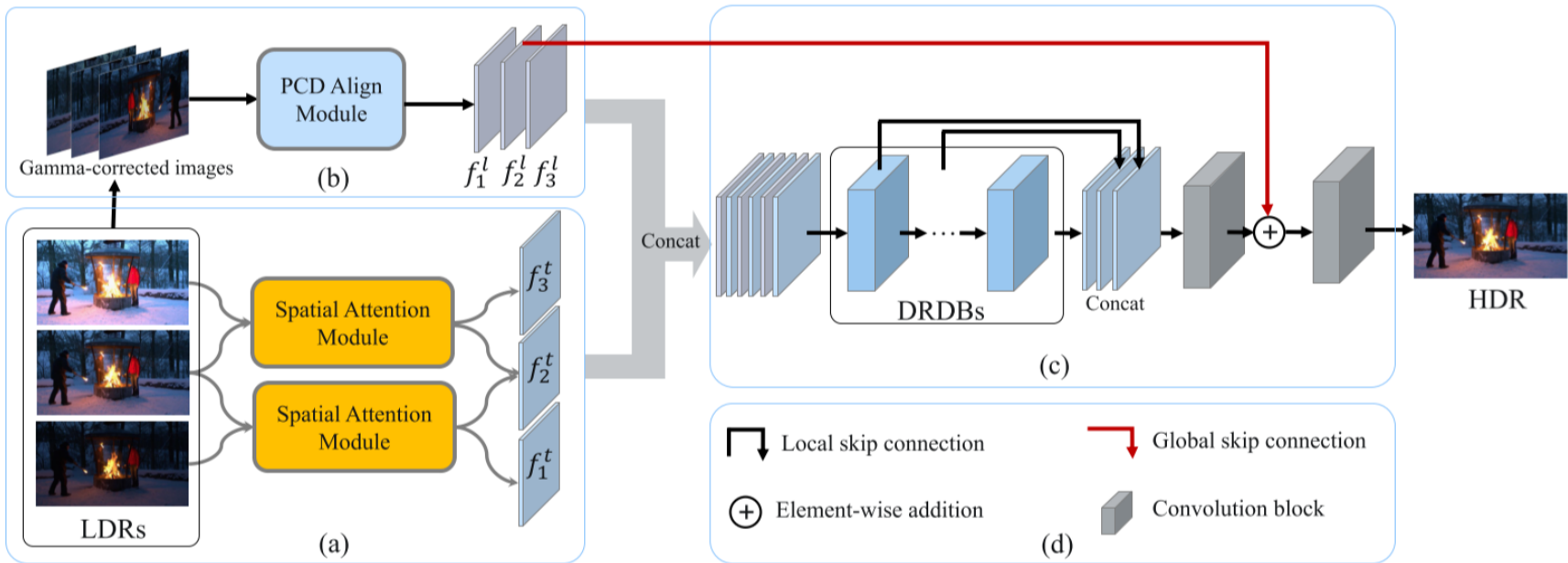


(a) standard convolution

(b) deformable convolution

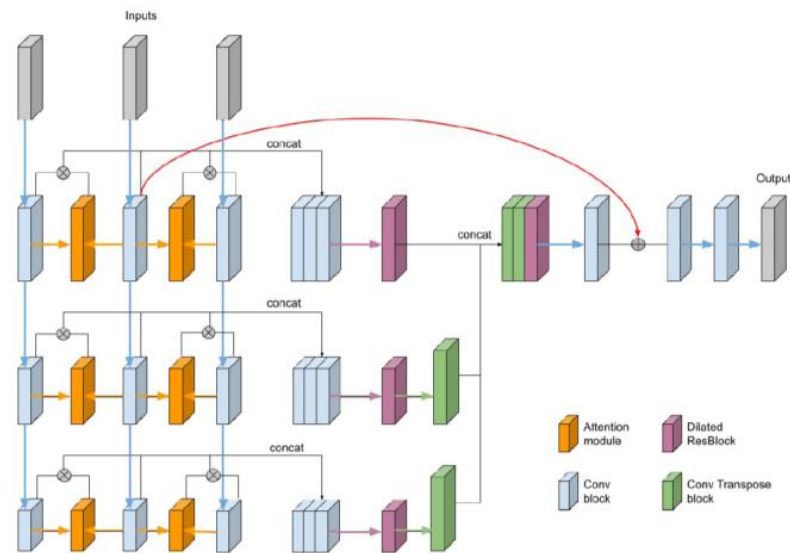
# Ranking 1 of Track 2 : MegHDR

- Zhen Liu et al., “ADNet: Attention-guided Deformable Convolutional Networks for High Dynamic Range Imaging”
  - Spatial attention 모듈, fusion subnet
    - AHDRNet 구조 활용



# Ranking 2 of Track 2 : SuperArtifacts

- Evelyn Yi Lyn Chee et al., “Multi-Level Attention on Multi-Exposure Frames for HDR Reconstruction”
  - 3개의 서로 다른 해상도에서 feature를 처리하고 병합하는 다단계 구조 제안
    - 이전 level의 절반 해상도
    - 큰 전경 움직임을 더 잘 처리하는데 효과적임
  - ※ Object detection 분야에서 작은 물체를 탐지해내기 위해서 사용된 접근 방식
  - AHDRNet 구조 활용



# Conclusion

- Challenge의 경우, 아이디어 보다는 성능에 중점을 두어 기존의 잘 구현된 네트워크를 적극적으로 이용하는 추세를 보임
  - 기존 SOTA 모델들을 기반으로 하여 간단한 아이디어만을 추가함
- 이러한 challenge에 참여할 때 기존 SOTA 모델들을 분석하고 다양한 아이디어를 적용하는 방향으로 가는 것이 중요할 것 같음