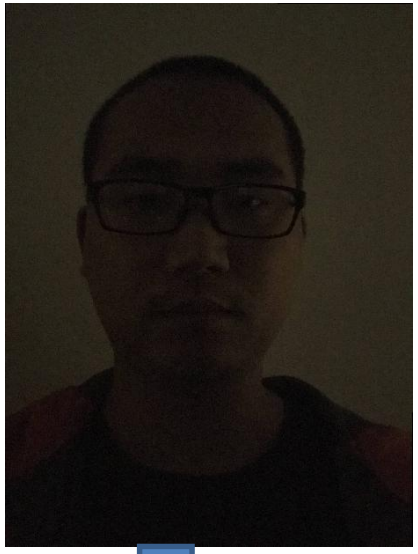
Low Light Face Image Enhancement

Liming Gong
Supervisor: Ha Le
Professor: Ioannis Kakadiaris

What does this project do?



Relighting

?

Face & Landmark detection

Face & Landmark detection

?

Compare?

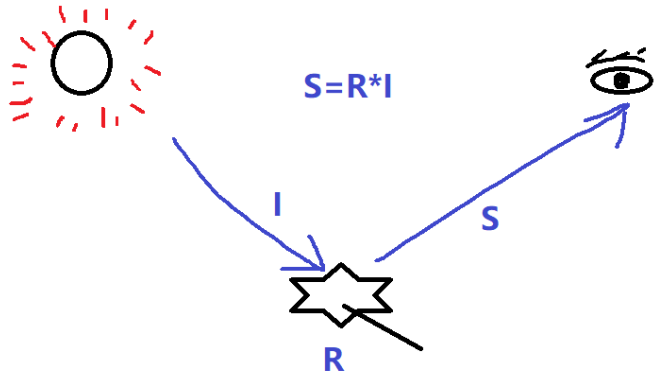
?

Project overview

- Reimplement a state-of-the-art low light image enhancement method: **Deep Retinex Net**
- Apply this network on 100 low light selfies
- Apply one pretrained face and landmark detection method to check improvement on enhanced selfies
- Dig deeper if something weird happens……

Lighting and Reflectance

S= R* I (element-wise multiplication)

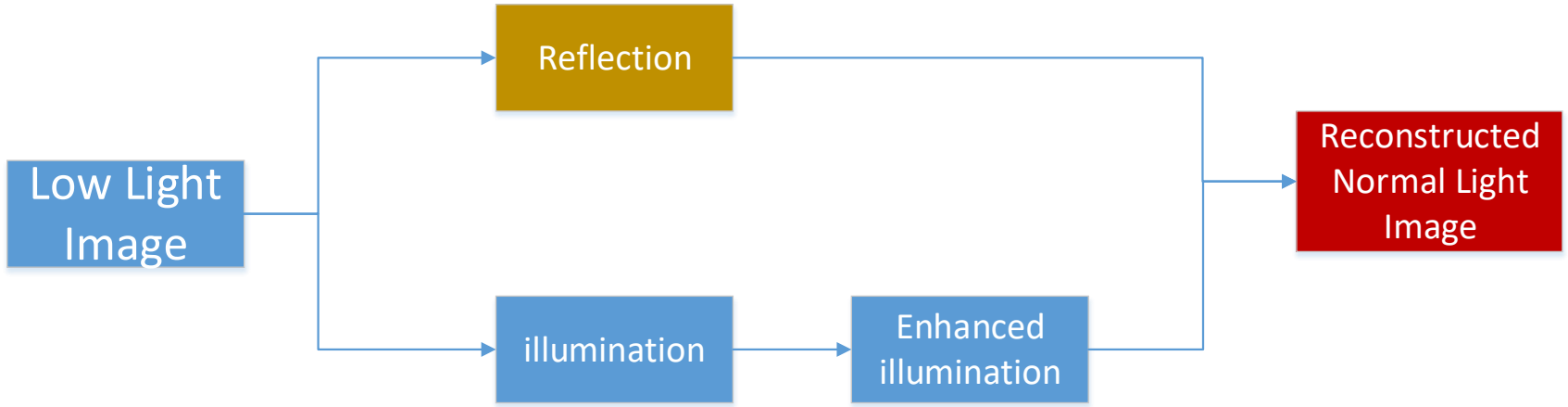


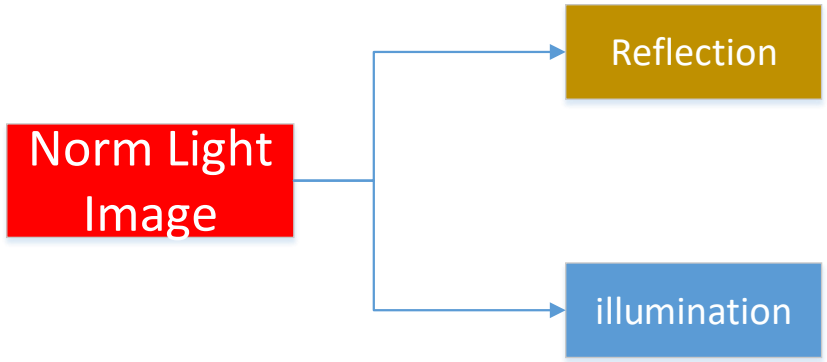
S: the scene you see

R: reflectance map, describes the intrinsic property of captured objects, which should not change for the same object no matter the light condition

I: illumination map, which reflects the lighting

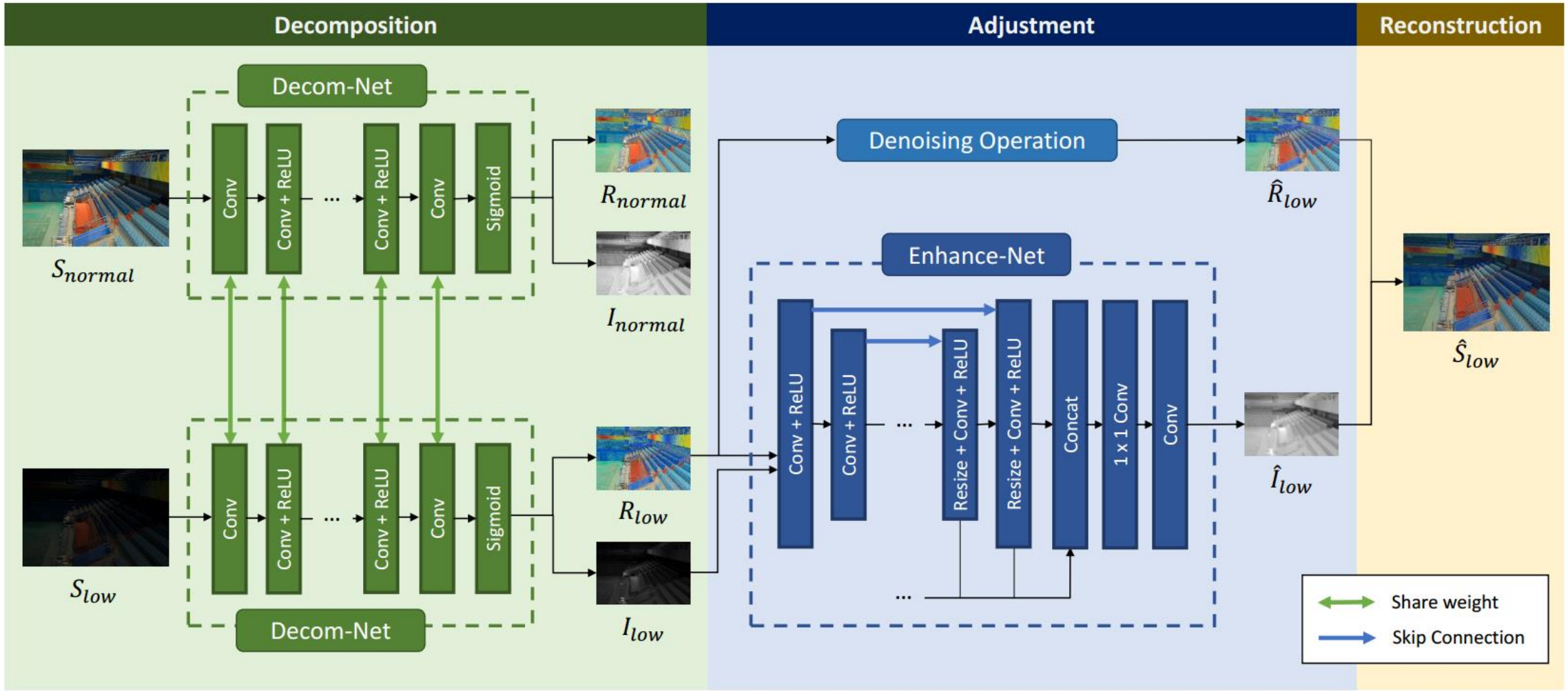
Giving low-normal pair





- the reflectance should be the same
- the decomposed I and R could reconstruct original image
- the final reconstructed image should be as close as to normal light image
- the illumination should be smooth (overall gradient should be small)

Their approach



Input and output

- Training

Input: Image pair of low light image and corresponding normal light image, patch size: 48*48

Output: Reconstructed image patch with enhanced illumination

- Testing

Input: full size low light image

Output: Reconstructed image with enhanced illumination

Loss function

$$\mathcal{L}_{recon} = \sum_{i=low,normal} \sum_{j=low,normal} \lambda_{ij} ||R_i \circ I_j - S_j||_1 .$$

$$\mathcal{L}_{ir} = ||R_{low} - R_{normal}||_1 .$$

$$\mathcal{L}_{is} = \sum_{i=low,normal} ||\nabla I_i \circ exp(-\lambda_g \nabla R_i)||$$

S: input image
R: reflectance map
I: illumination map

- Decom Net

$$\mathcal{L} = \mathcal{L}_{recon} + \lambda_{ir}\mathcal{L}_{ir} + \lambda_{is}\mathcal{L}_{is}$$

- Enhance Net

$$\mathcal{L} = \mathcal{L}_{recon} + \lambda_{is}\mathcal{L}_{is}$$

Train and test steps

Train:

1, Random crop image pair patches and apply data augmentation.

2, Train Decom-Net with input image pairs.

3, Fix Decom-Net, train Enhance-Net with input image pairs.

Test:

1, Feed low light images.

Contributions

- The first dataset with paired low/normal-light images captured in real scenes.

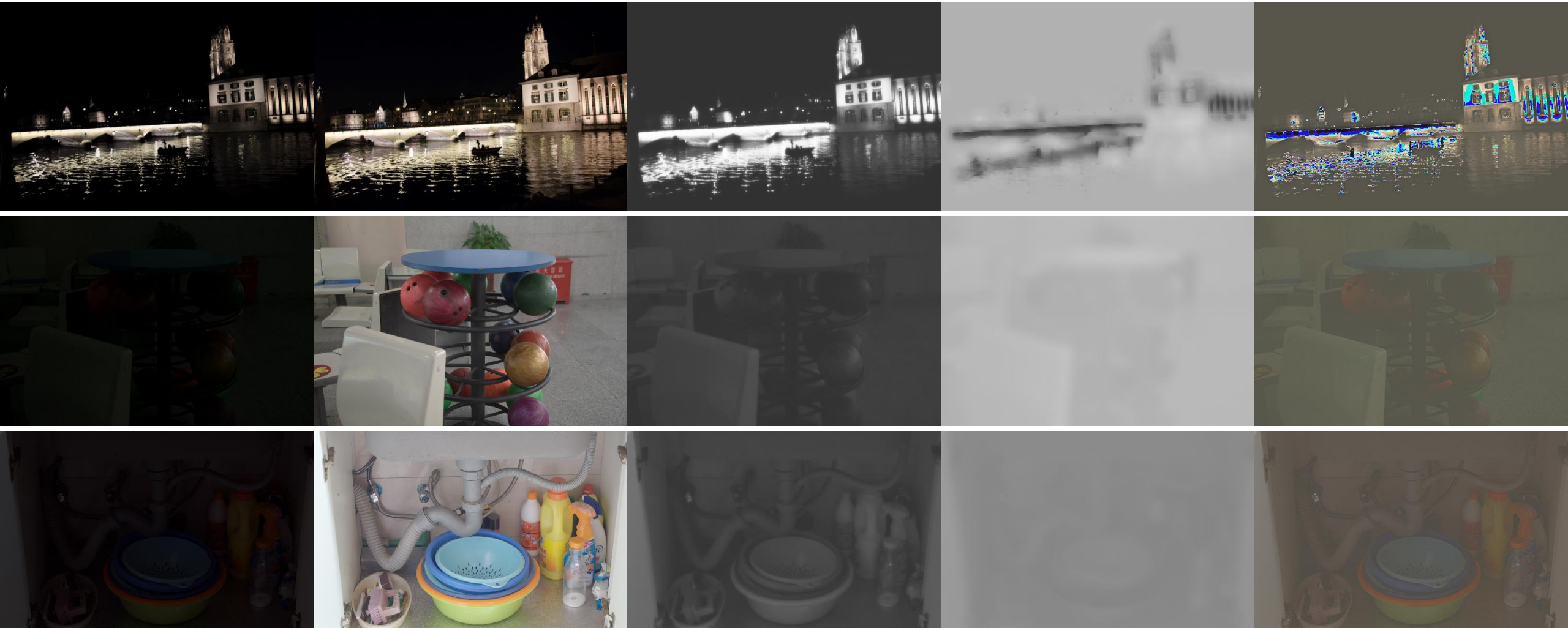
- Light weighted, end to end trainable network: Deep Retinex Net.

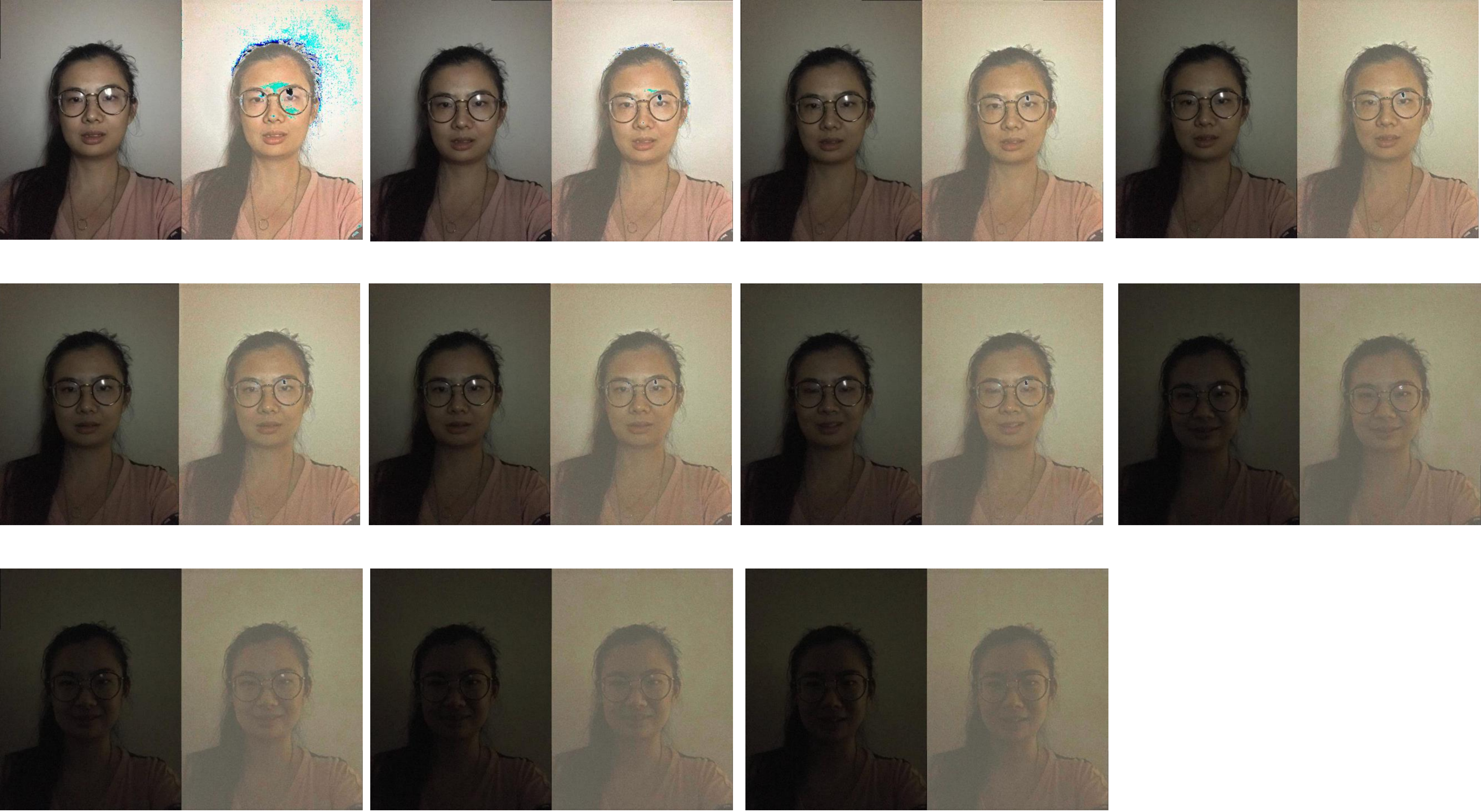
My result for Decom Net

low Light | normal light | reflectance | illumination

My result for Enhance Net

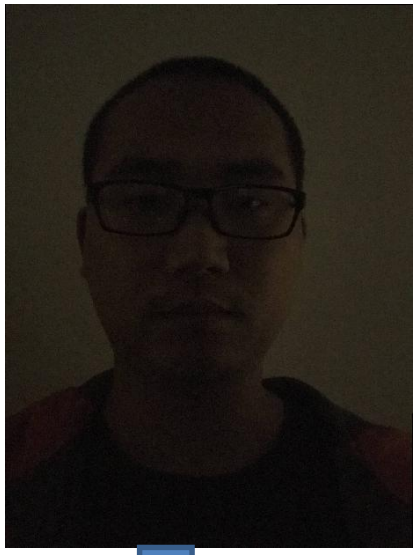
low light | normal light | illumination | enhanced I | reconstructed





Author’s vs Mine

Does this really help?



Relighting

Face & Landmark detection

Face & Landmark detection

?

Compare?

?

The pretrained model

- **How far are we from solving the 2D & 3D Face Alignment problem? (and a dataset of 230,000 3D facial landmarks) (ICCV 2017) (SFD face detector)**

Detect 2D facial landmarks in pictures

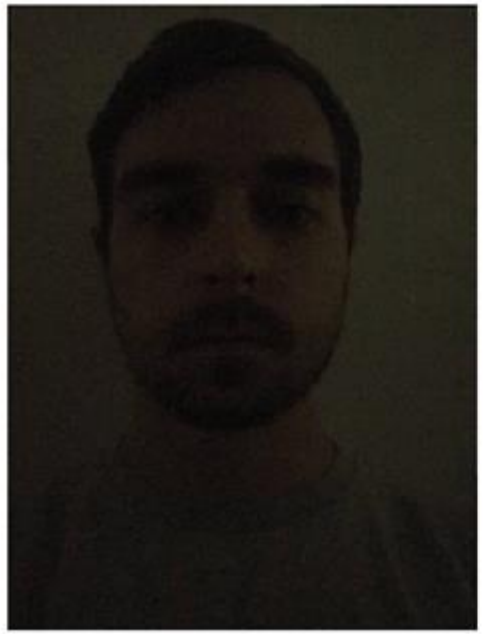
Face detection rate

- The original face images: 118/124 = 95.16%
- My relighted faces: 105/124 = 84.68%
- Author’s relighted faces: 79/124 = 63.71%

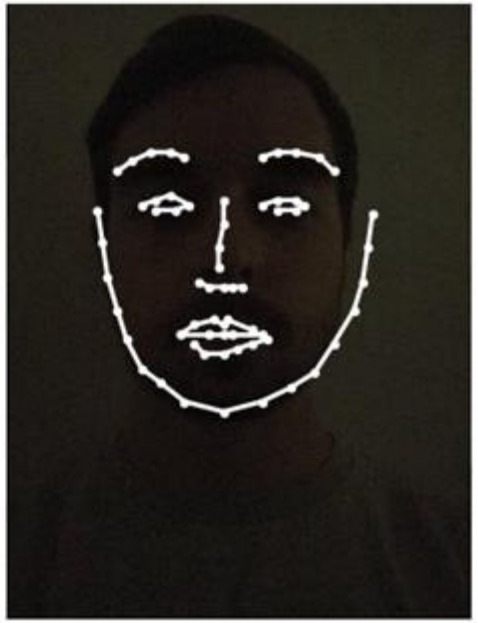
Original > My enhancement > Author’s enhancement

Landmark detection

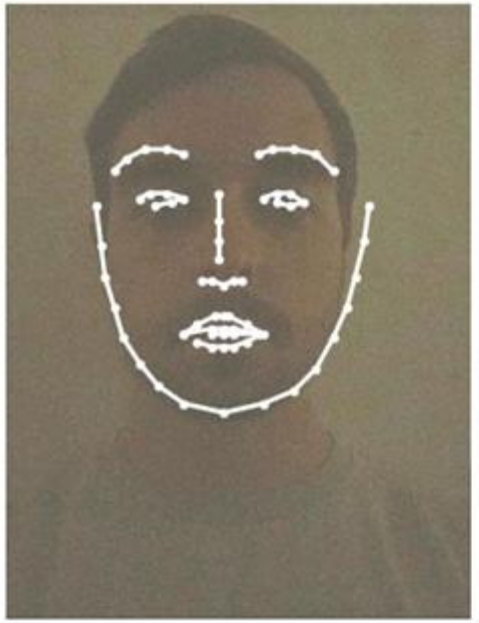
original



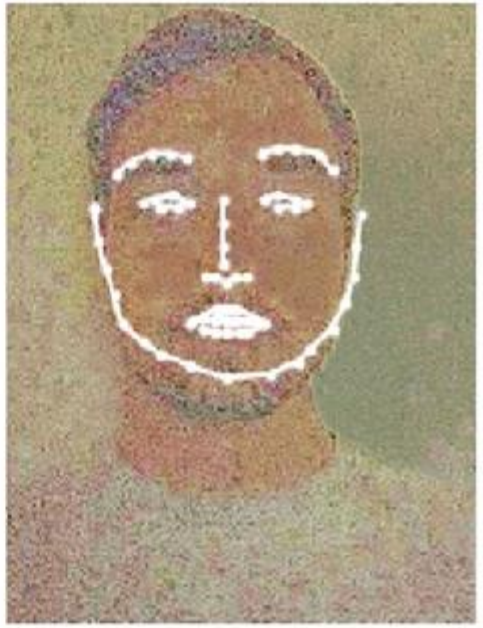
original



mine



author



Dig deeper: add Gamma enhancement

original

gamma

mine

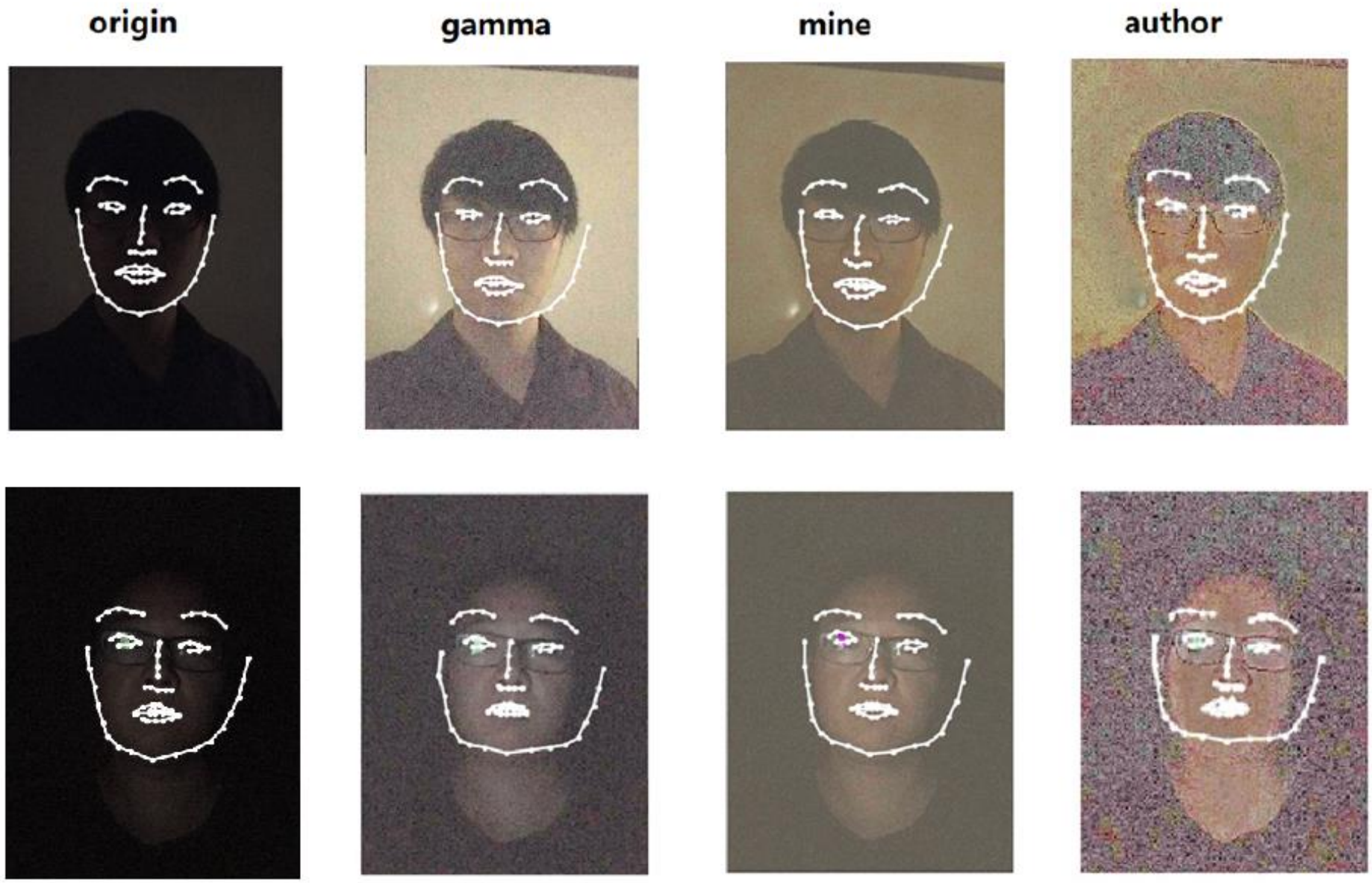
author

Face detection rate

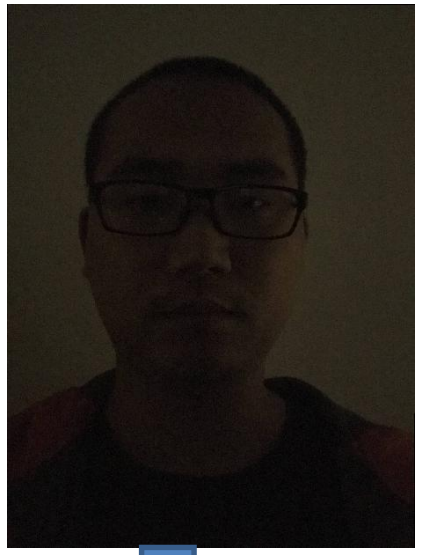
- The original face images: 118/124 = 95.16%
- My relighted faces: 105/124 = 84.68%
- Author's relighted faces: 79/124 = 63.71%
- Gamma enhancement: 116/124 = 93.54%

Original > Gamma > My enhancement > Author's enhancement

Landmark detection



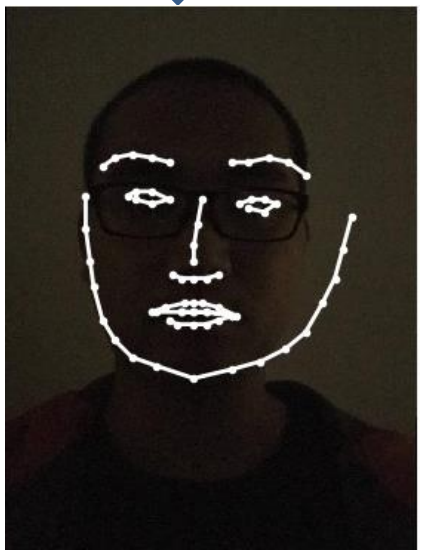
Does this really useful?



Relighting

Landmark detection

Landmark detection



Compare?

Result summarize

- Image enhancement is not helping face detection. It actually makes it worse.

- No sure if it helps landmark detection.

Possible explanation

- The face/landmark detection method is trained on natural images. Any operation on input images may just change the distribution and make them “unnatural”.

What have I learned?

- Instinct may go wrong.
- Solve tasks straight forward. You don’t need enhance images before face detection.
- A machine with the capability to solve “harder” problems may fail on “easy” ones.
- Machine intelligence is really nothing humanlike.

My Notebook Environment

- Inteli5-7300HQ+24G RAM + GTX1060(6G RAM)
- Windows 10 system
- Tensorflow1.10 + Python3.66

- Training set: 1500 images offered by the author
- Testing set: 124 face images acquired by myself
- Training takes less than 1 hour for 100 epochs.

Complementary material:

- Paper website:

https://daooshee.github.io/BMVC2018website/

- My GitHub code repo:

https://github.com/stephenkung/FaceEnhancement

- Face alignment method:

https://adrianbulat.com/face-alignment

- BM3D denoising

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.300.5214&rep=rep1&type=pdf

Thank you!