Dark Arts :
Low Image Enhancement and Interpretation.

Group 2

G.C.Jayathilaka (E/14/158)
S.Suren (E/14/339)
W.G.H.S.Weligampola (E/14/379)

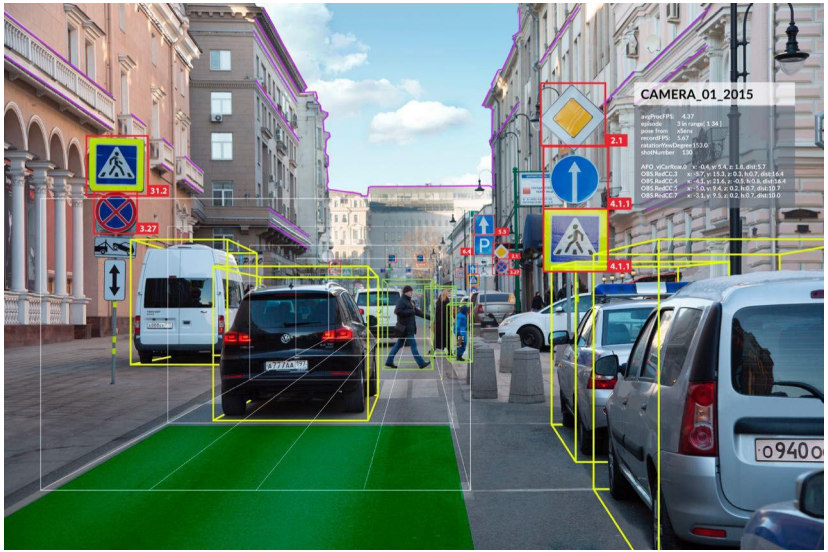
Prof. R.G.Ragel
Dr. R.Godaliyadda
Dr. V.Herath
Dr. P.Ekanayake

Introduction

Presented by E/14/158

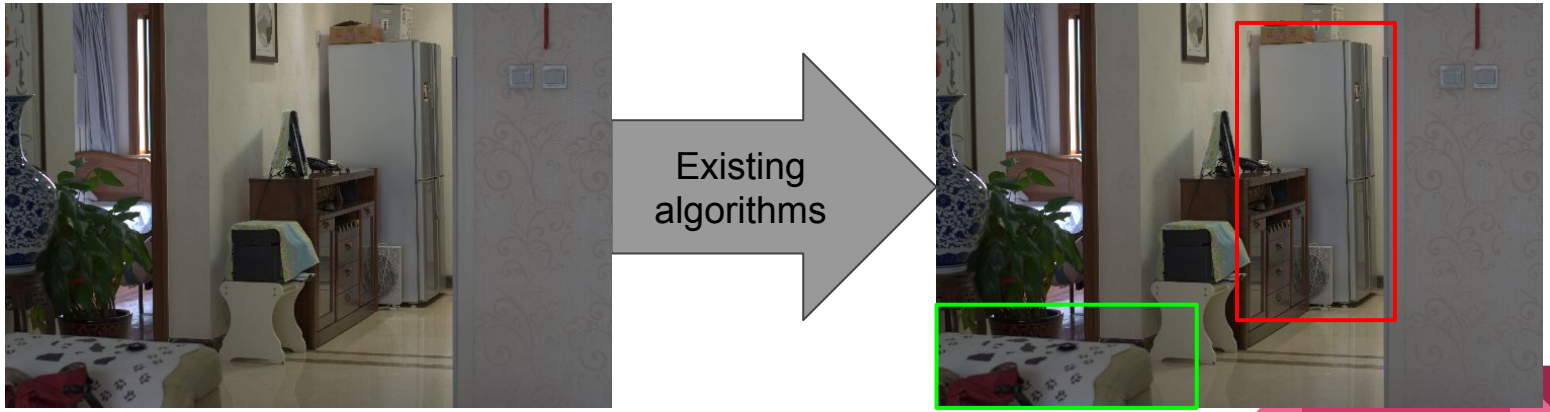
Computer vision

- more than **<u>2000 research papers</u>** are being published on computer vision annually.
  - These papers discuss how to interpret visual input for **object detection, scene interpretation, colour adjustments**, etc.
- There are **<u>many vision based products</u>** based on these researches.



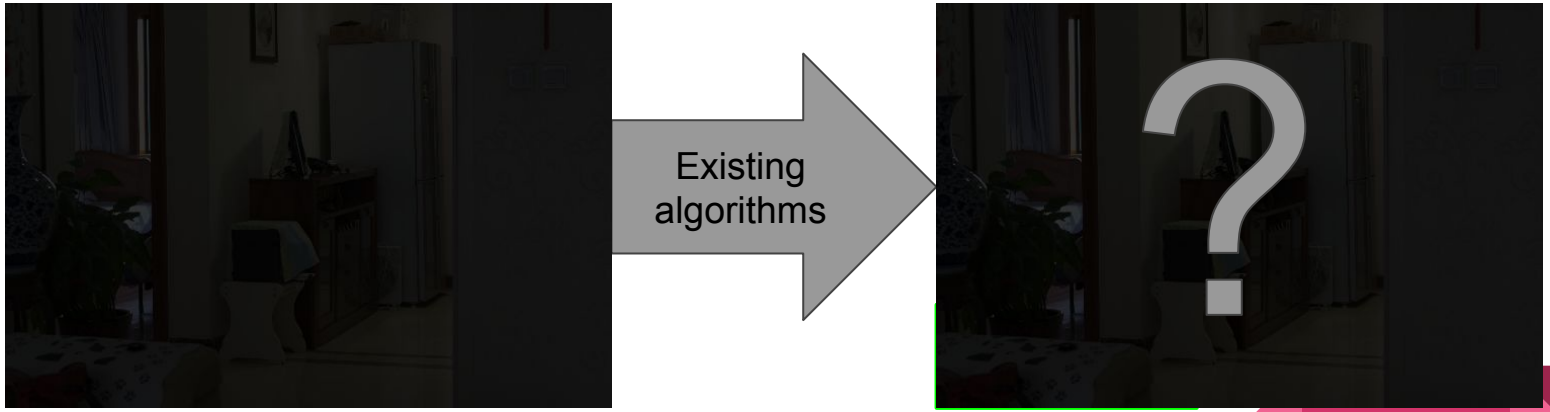
Computer vision: The problem

99% of the existing work in computer vision applies for good lighting conditions which restricts its application.



Computer vision: The problem

99% of the existing work in computer vision applies for good lighting conditions which restricts its application.



Existing solutions (Non Algorithmic)

- Artificial lighting
  - Consumes energy
  - Disturbs natural ecosystems.
- Sophisticated camera hardware
  - The night mode in cameras is enabled through expensive hardware.
- High-Dynamic-Range (HDR) Imaging
  - Movement of dynamic objects cause “ghosting effect”.

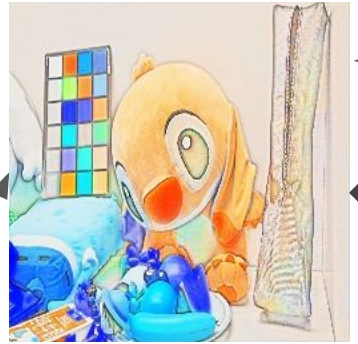
Related concepts

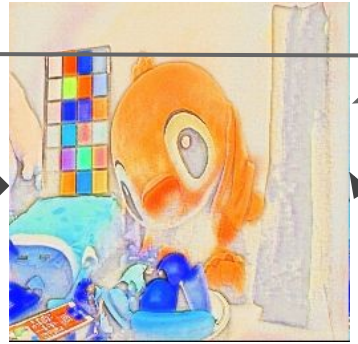
These terms will be used throughout the presentation.

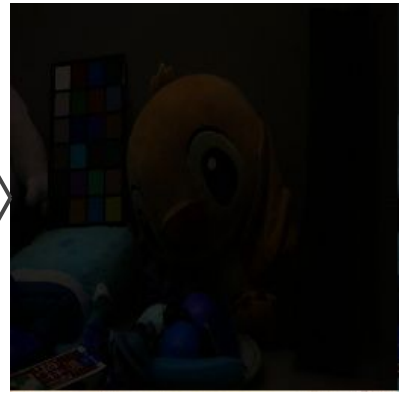
- Retinex theory
  - Reflectance and Illumination.
- Datasets
  - Enhancement problems
    - Paired datasets
    - Unpaired datasets
  - Image classification problems
    - One class datasets

Retinex based model

Image : S



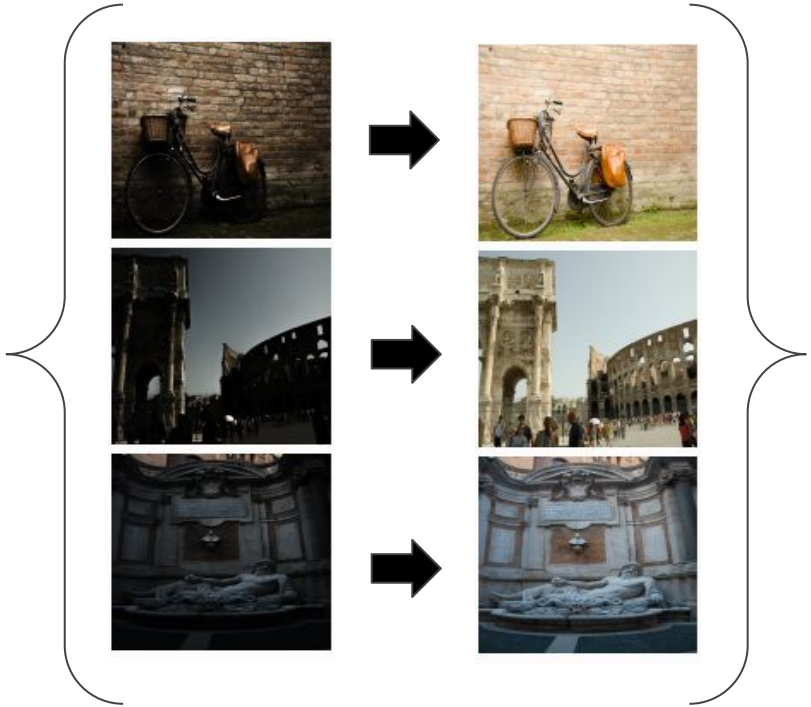




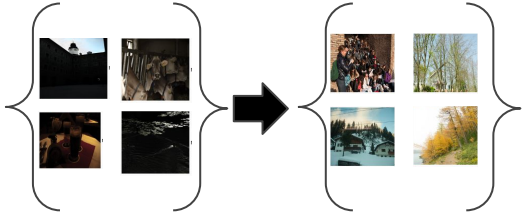
Reflectance : R
(colour information)
Invariant property

Illumination : I
(Lighting information)
Light dependant property

Dataset: Types (1/2)

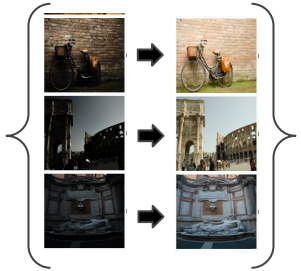


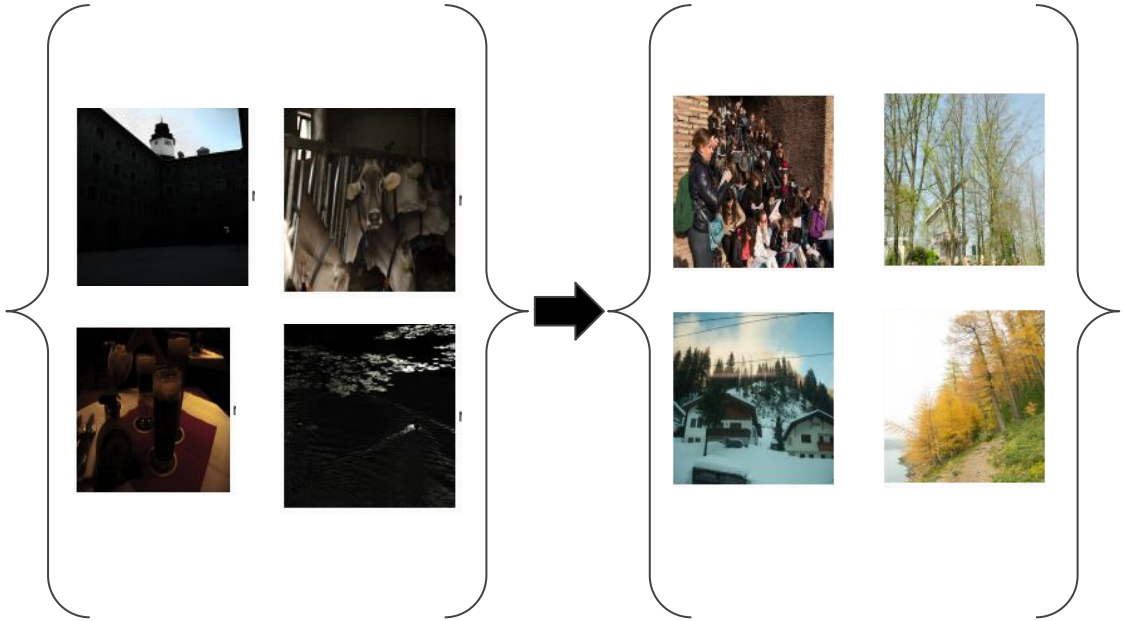
- **Paired dataset**: Every dark image has it's well light counterpart.
  - Difficult to collect.
  - More information.



Dataset: Types (2/2)

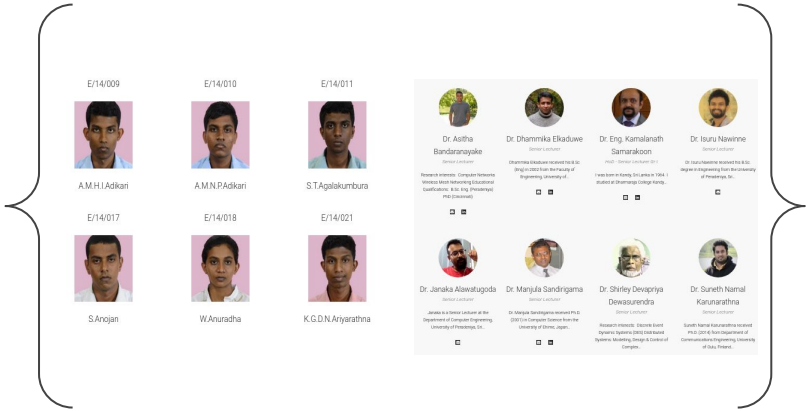
- **Unpaired dataset**: There are unrelated sets of well lit and dark images.
  - Easy to obtain.





One Class Datasets

These datasets are for anomaly detection. We have examples for the “usual” data points but none for the anomalies.



Class (can obtain)

Out of class (cannot obtain)

Case studies

The research is divided into 3 seperate case studies for convenience

1. **Low-light image enhancement**
2. **Multi-level controlled image enhancement**
3. **Low light object detection**

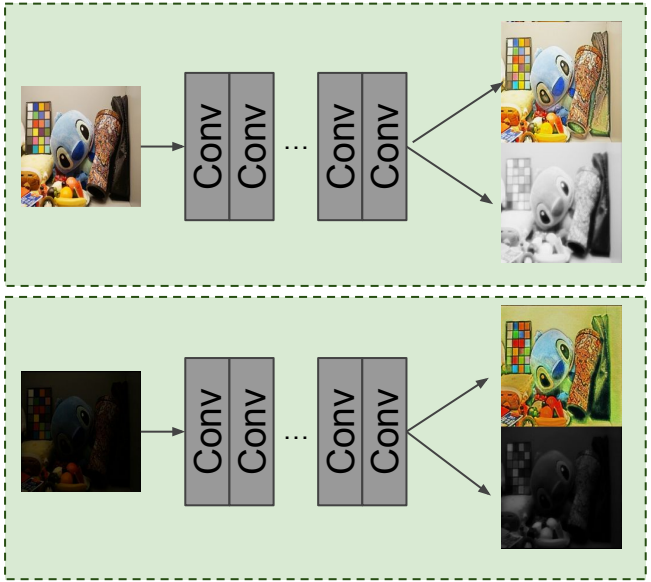
Low-light image enhancement

Presented by E/14/379

Evolution of low-light enhancement algorithms

- Classical algorithms (*unpaired dataset*)
  - Intensity based (Histogram Equalization) / Gradient based (Grad-Enhance)
- Retinex-theory (*paired/unpaired dataset*)
- Deep Convolutional Neural Network (*paired/unpaired dataset*)
  - LLNet, LLCNN, **RetinexNet**
- Adversarial learning (*paired/unpaired dataset*)
  - Retinex-GAN, **Enlighten-GAN**

Retinex based decomposition network



$$\mathcal{L} = \lambda_{IR}\mathcal{L}_{IR} + \mathcal{L}_{recon} + \lambda_{IS}\mathcal{L}_{IS}$$

Invariable reflectance Loss

$$\mathcal{L}_{IR} = \|R_{low} - R_{high}\|$$

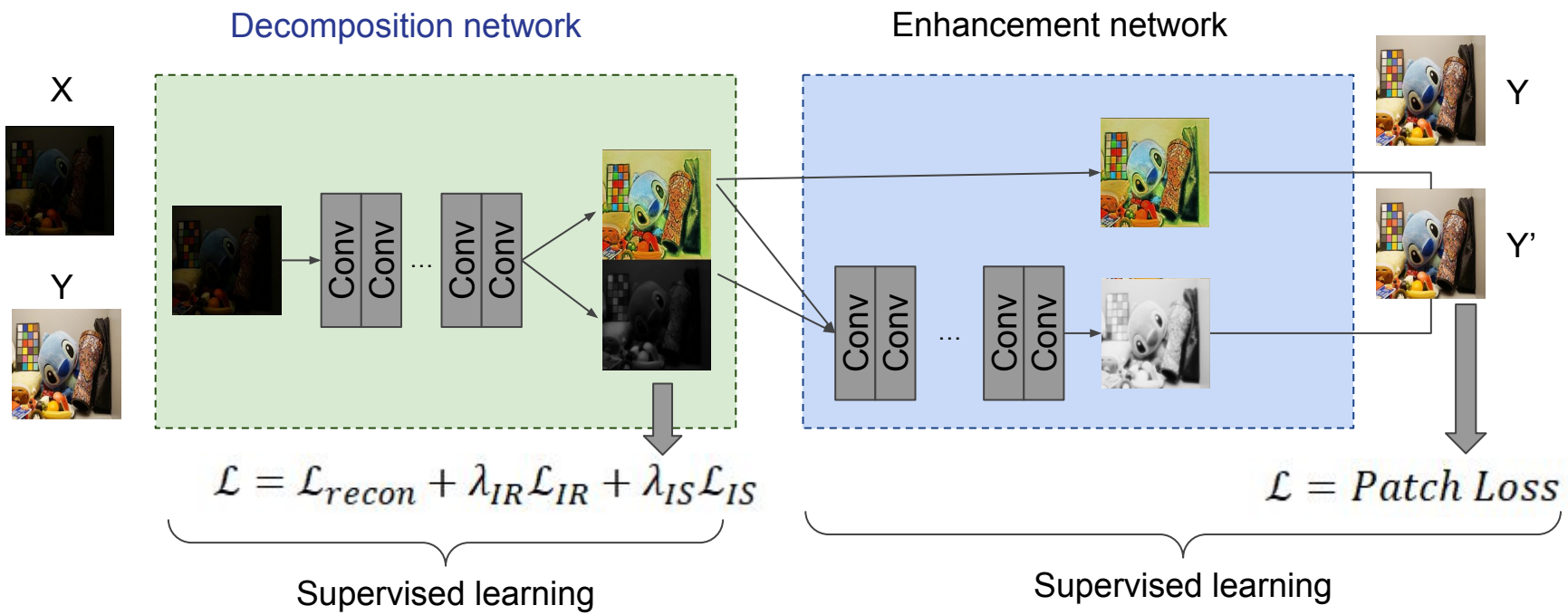
Reconstruction Loss

$$\mathcal{L}_{recon} = \lambda_r \|R_{low} \circ I_{high} - S_{high}\| + \lambda_h \|R_{high} \circ I_{low} - S_{low}\|$$

Illumination Smoothness Loss

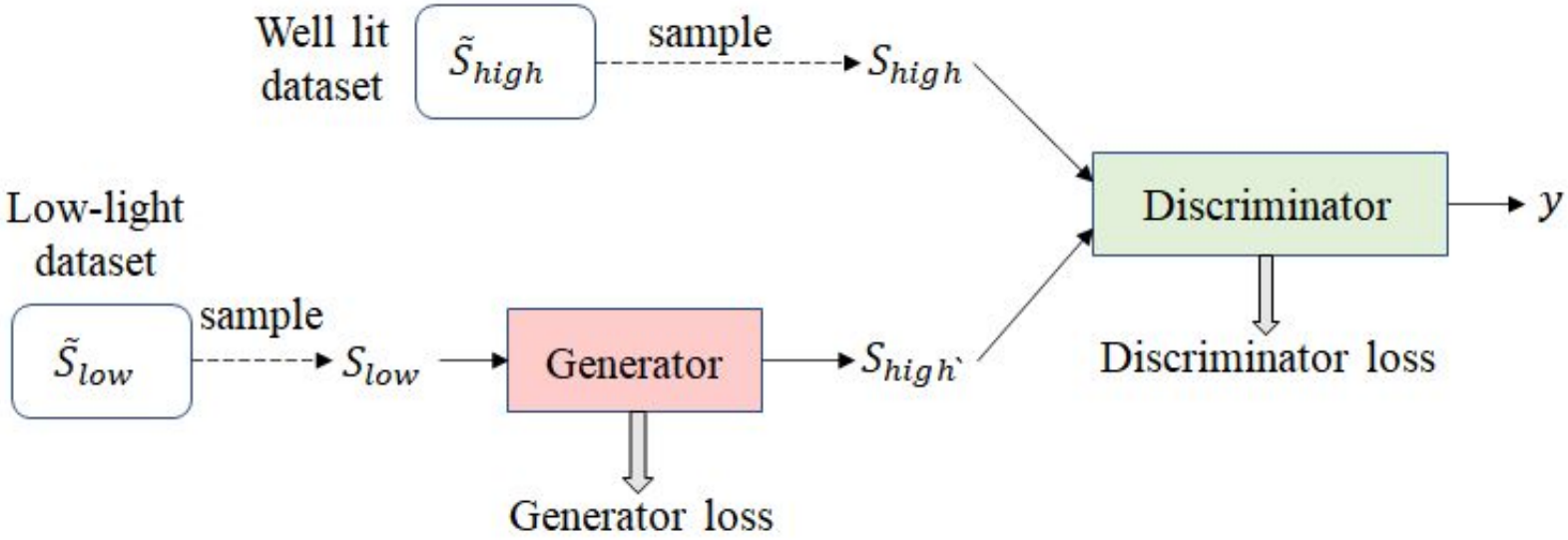
$$\mathcal{L}_{IS} = \|\nabla I \circ \exp(-\lambda_g \nabla R)\|$$

RetinexNet (2018)



[1] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. “Deep retinex decomposition for low-light enhancement”. In BMVC, 2018

GAN & DCGAN*

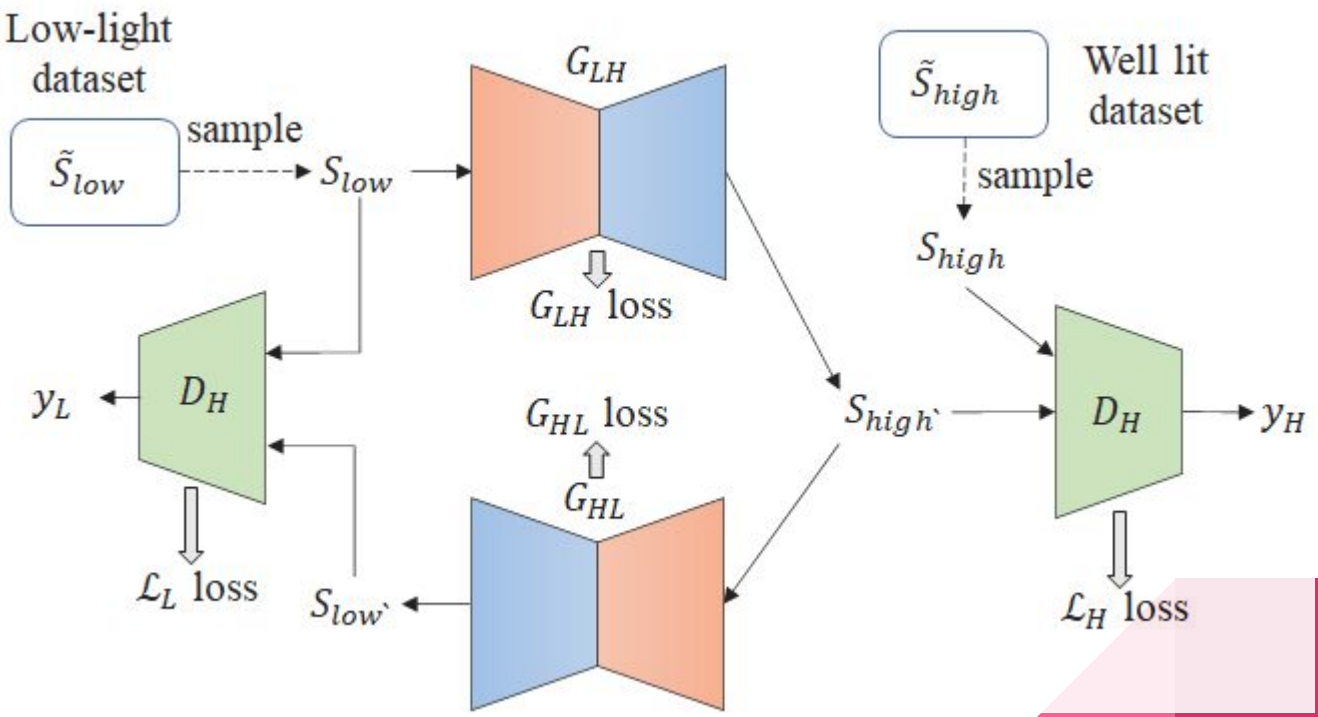


*Deep Convolutional Generative Adversarial Network

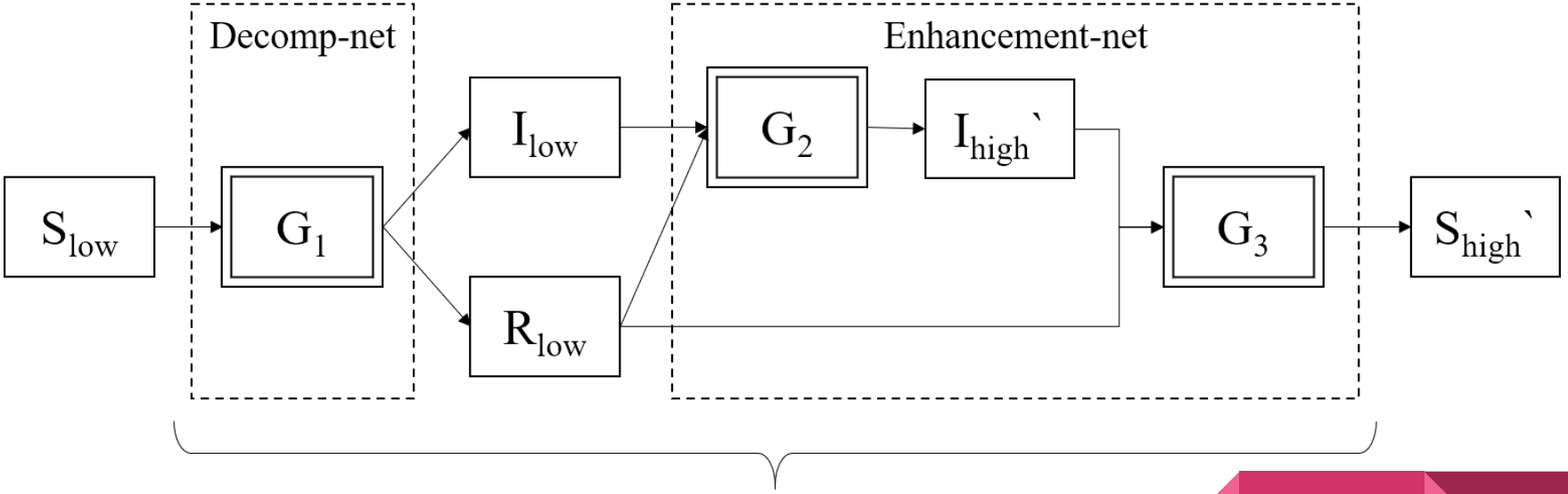
Proposed method: Steps

1. Identification of **illumination level**.
2. Extracting **color information** even in the poorly-light condition.
3. Increase image illumination while **preserving and enhancing the color** information.
4. **Handle the noise** and deformations introduced to the image during the enhancement process.

CycleGAN

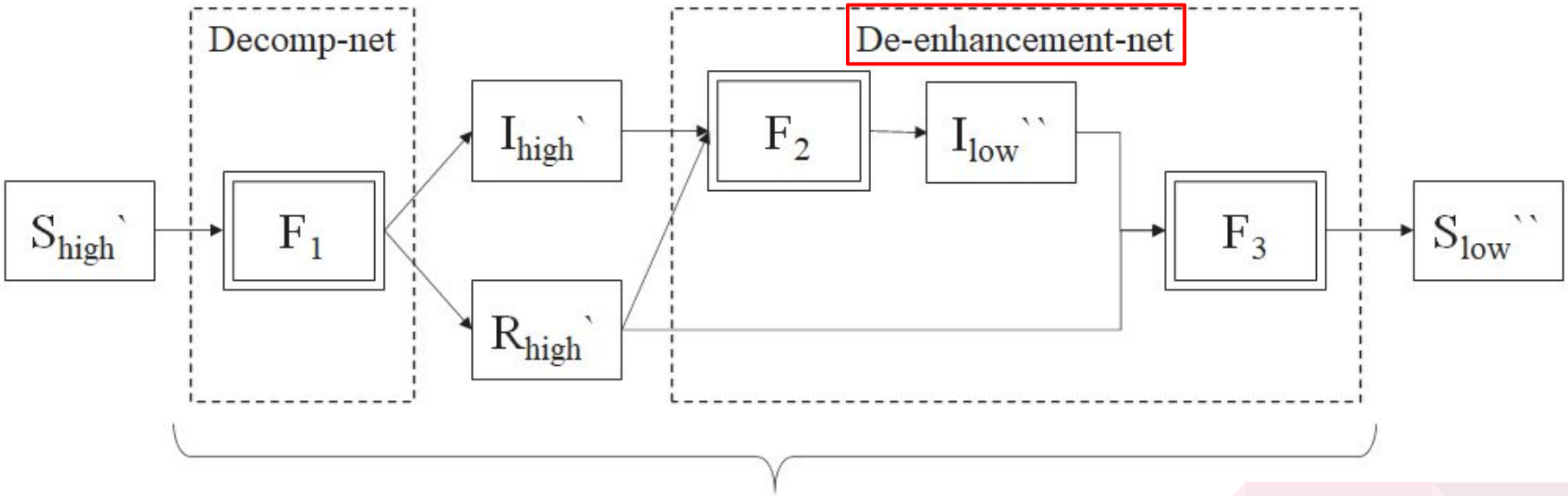

Low-light dataset
$\tilde{S}_{low}$
sample
S_{low}
G_{LH}
G_{LH} loss
D_H
y_L
$\mathcal{L}_L$ loss
G_{HL} loss
G_{HL}
$S_{low`}$
$S_{high`}$
$\tilde{S}_{high}$
Well lit dataset
sample
S_{high}
D_H
y_H
$\mathcal{L}_H$ loss

Component analysis: Forward generation



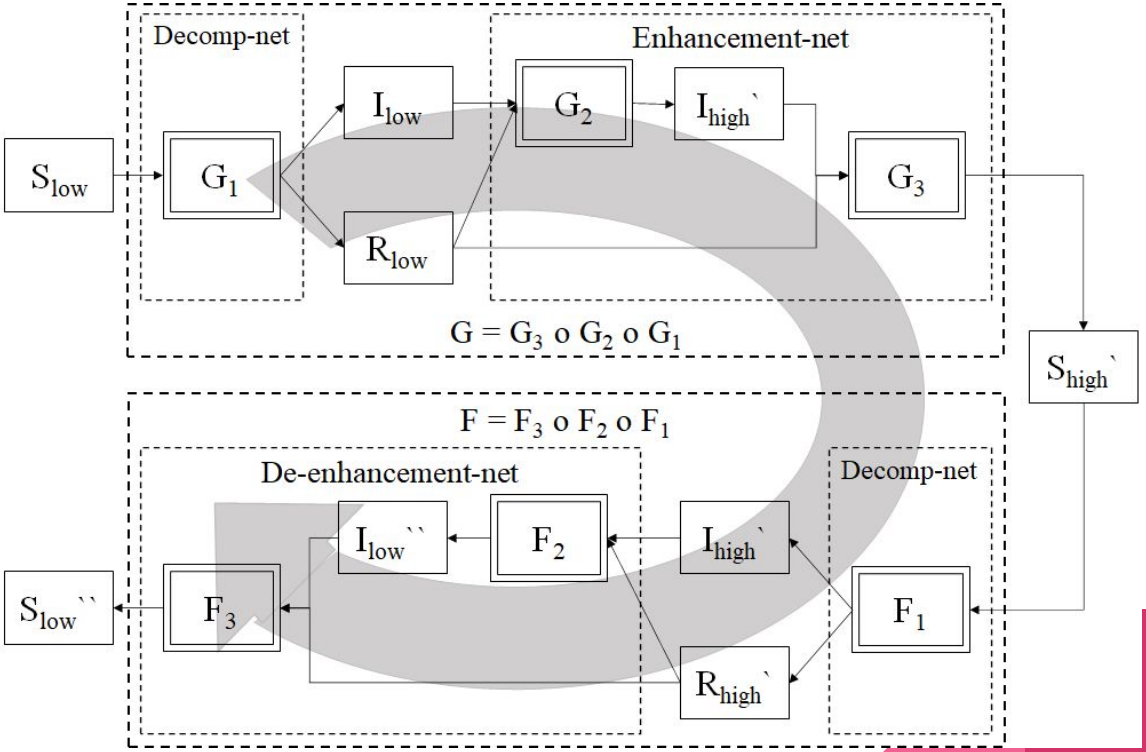
$$G = G_3 \circ G_2 \circ G_1$$

Component analysis: Reverse generation

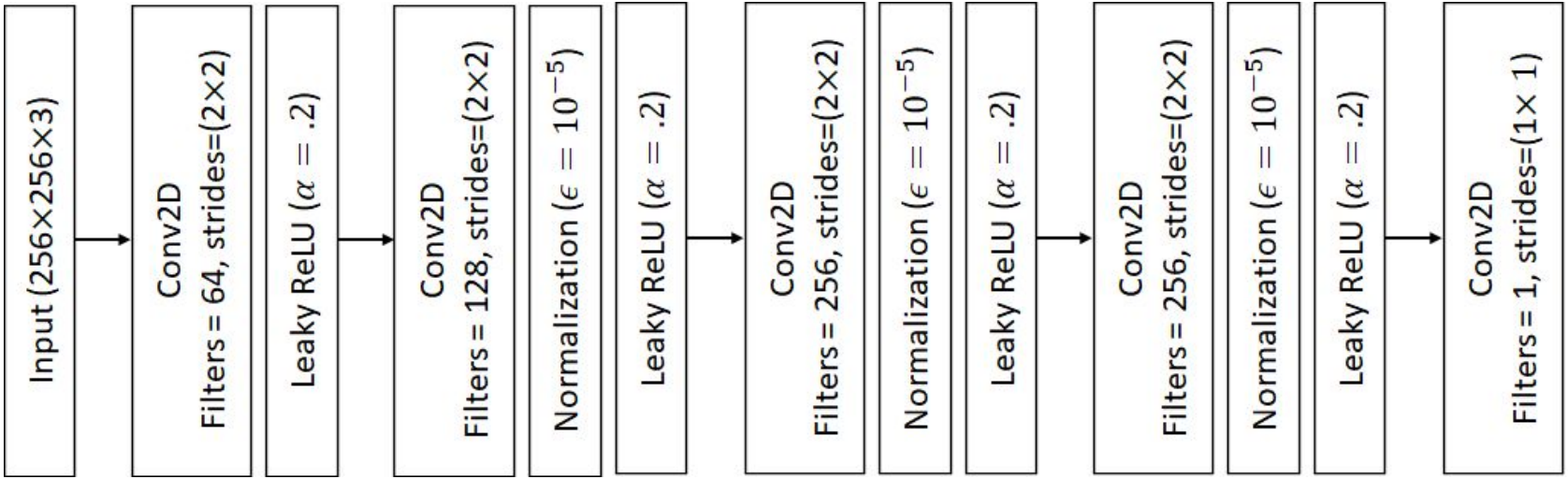


$$F = F_3 \circ F_2 \circ F_1$$

Component analysis: GAN cycle



Component analysis: Discriminator



Component analysis: Loss function

$$\mathcal{L}_{cyc_s} = \mathbb{E}_{S_{low}\ p(S_{low})}\left[\left\|F\left(G(S_{low})\right) - S_{low}\right\|_1\right] + \mathbb{E}_{S_{high}\ p(S_{high})}\left[\left\|G\left(F(S_{high})\right) - S_{high}\right\|_1\right]$$

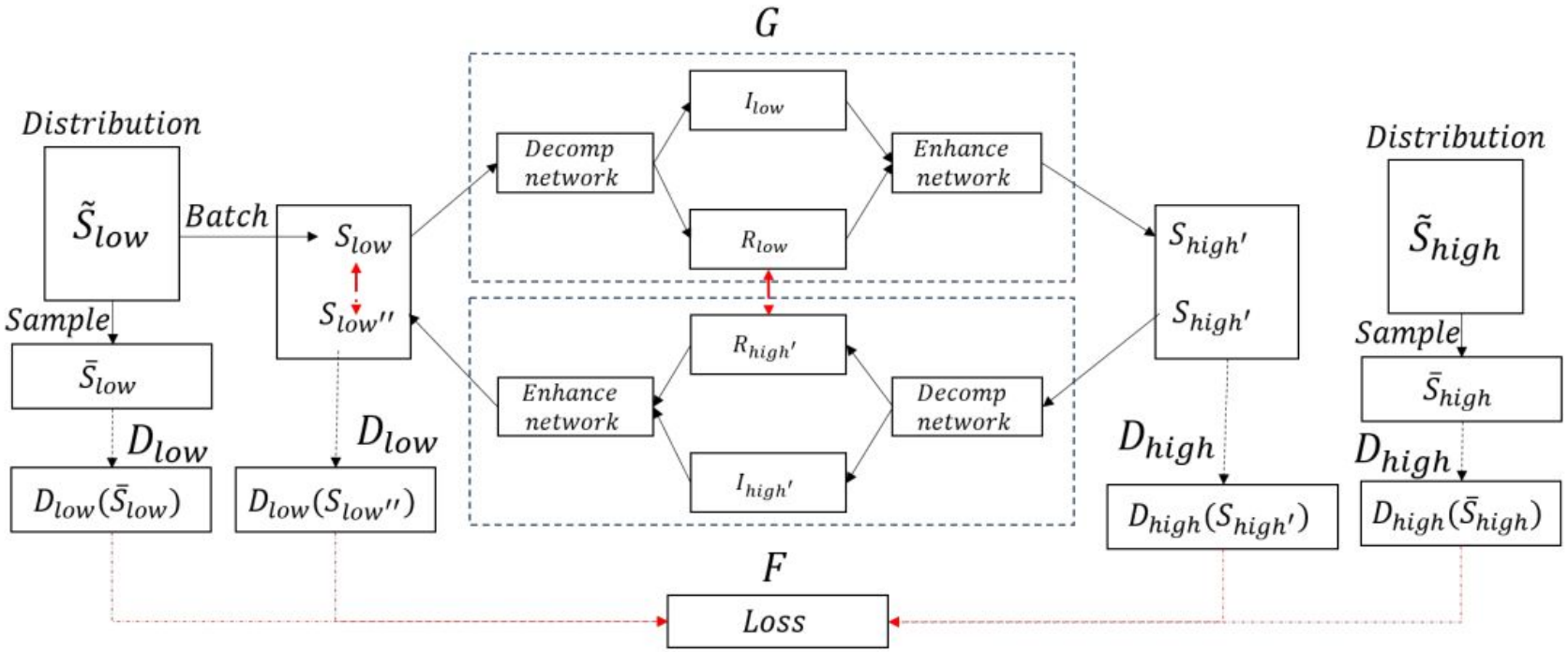
$$\mathcal{L}_{cyc_R} = \left\|R_{low} - R_{high`}\right\|_2 + \left\|R_{high} - R_{low`}\right\|_2$$

$$\mathcal{L}_{cyc} = \mathcal{L}_{cyc_s} + \mathcal{L}_{cyc_R}$$

$$\mathcal{L}_{gen} = \mathcal{L}_{cyc} + H\left(1, D_{high}\left(G(S_{low})\right)\right) + H\left(1, D_{low}\left(F(S_{high})\right)\right)$$

$$\mathcal{L}_{disc} = H\left(1, D_{high}\left(G(S_{low})\right)\right) + H\left(0, D_{high}(S_{high})\right)$$
$$+H\left(1, D_{low}\left(F(S_{high})\right)\right) + H\left(0, D_{low}(S_{low})\right)$$

Proposed model: Architecture



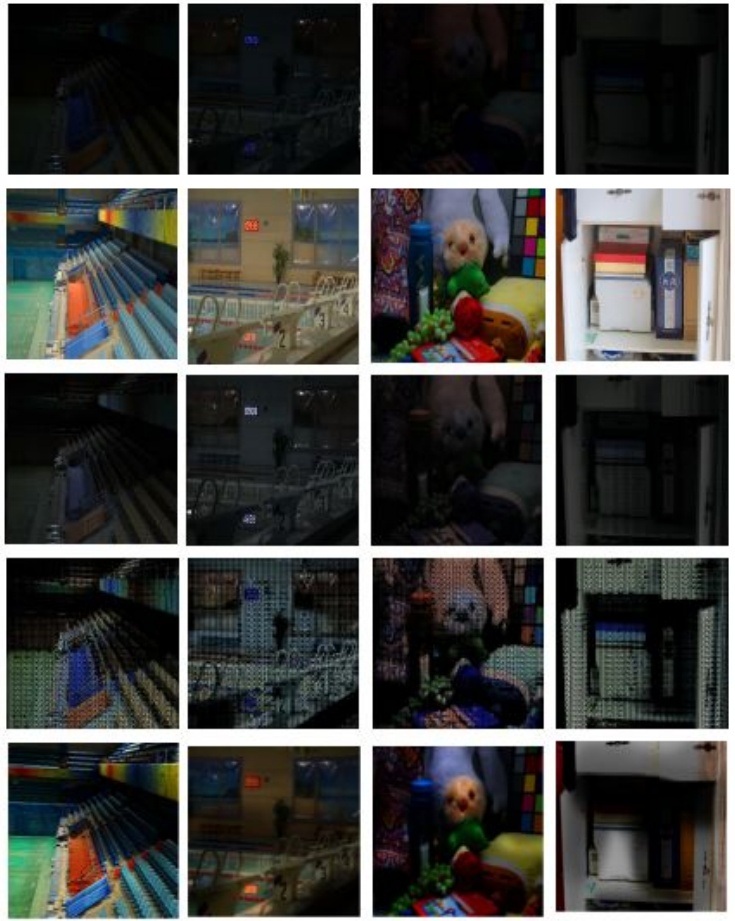
Low lit image (S_{low})

Corresponding well lit image (S_{high})

Enhancing low light images using a **generic GAN**

Enhancing low light images using a **generic CycleGAN**

Proposed model



Comparison of evaluation metrics

Algorithm	**MSE**	**NIQE**
Ground truth	0.0000	1.0000
DCGAN	0.2171	1.4592
Retinex based DCGAN	0.0514	1.6967
Retinexnet	**0.0090**	1.7896
Proposed algorithm	0.0173	**1.7921**

Conclusion and Futureworks

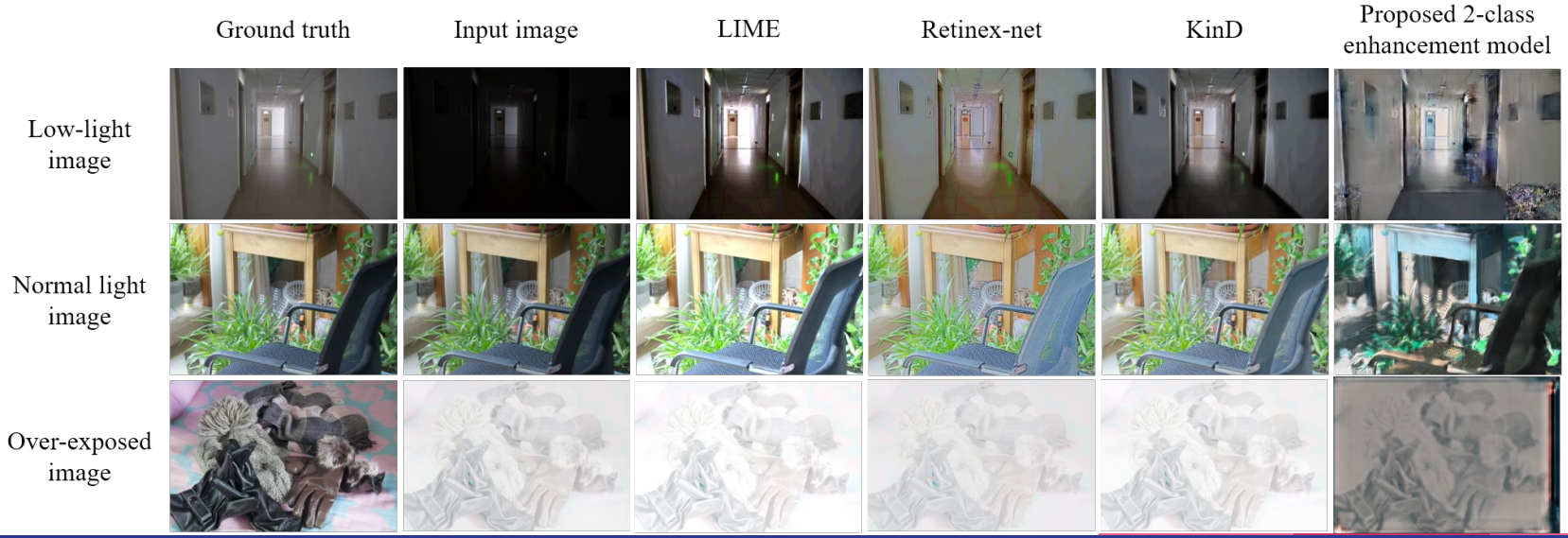
- The proposed model **combines** existing ideas from Retinex theory, CNN, and CycleGAN.
- Using **both** paired (synthetic + non-synthetic) and unpaired (non-synthetic) images, the model provides better performance in comparison.
- Certain images show issues with respect to **smoothness** similar to other related works. This must be analyzed for further improvements.

Multi-level controlled illumination enhancement

Presented by E/14/339

Two-class enhancement - Shortcomings

- Low-light image enhancement as a **two-class problem** is well defined.
- However, when translated to real world applications, these models often encounter **performance degradation**.



Multi-class Controlled illumination enhancement

- Two main reasons for lower performance of two-class enhancement models:
  - Translation of images under **two type of lightning conditions** rather than a range of lighting in real world applications.
  - Assumes spatial uniform distribution of exposure, and **low exposure distortion**.

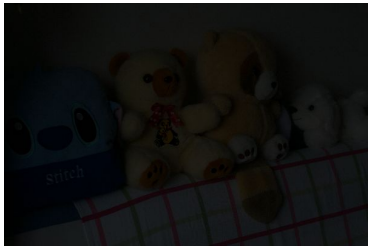
- Therefore, we motivate ourselves through the following questions:

- **Improve the performance of the proposed model.**
- **Produce images of varying illumination levels.**

Related works

- Prior works address two interconnected subproblems :
  - **Exposure adjustment** : Normalize the distribution of light in the image.

  - **Illumination enhancement** : Adjust (increase) the overall light incident on the image.



Datasets

- Dataset requires high variation in terms of both illumination and exposure distortion which is not available.
- Therefore, multiple datasets were combined:
  - Training set : 500 images
    - (synthetic, paired)
  - Test set: 100 images
    - (synthetic + non-synthetic)

Ex-dark images

Well-lit images

Synthetic images

Multi-illumination level paired data

- Supervised learning requires **paired data** with different illumination levels.
- **No labeled dataset** since the problem is not well defined. So, we produce synthetic paired dataset. (Explained next)
- **Labeling** for controlled illumination adjustment:
  - Images can be labeled using **continuous values** proportional to illumination:
    - Value (HSV), **Luma (YCbCr),** Lightness (LAB).
  - However, generic supervised learning models cannot handle continuous data. Thus, images are assigned **discrete labels** and separated to 10 classes.

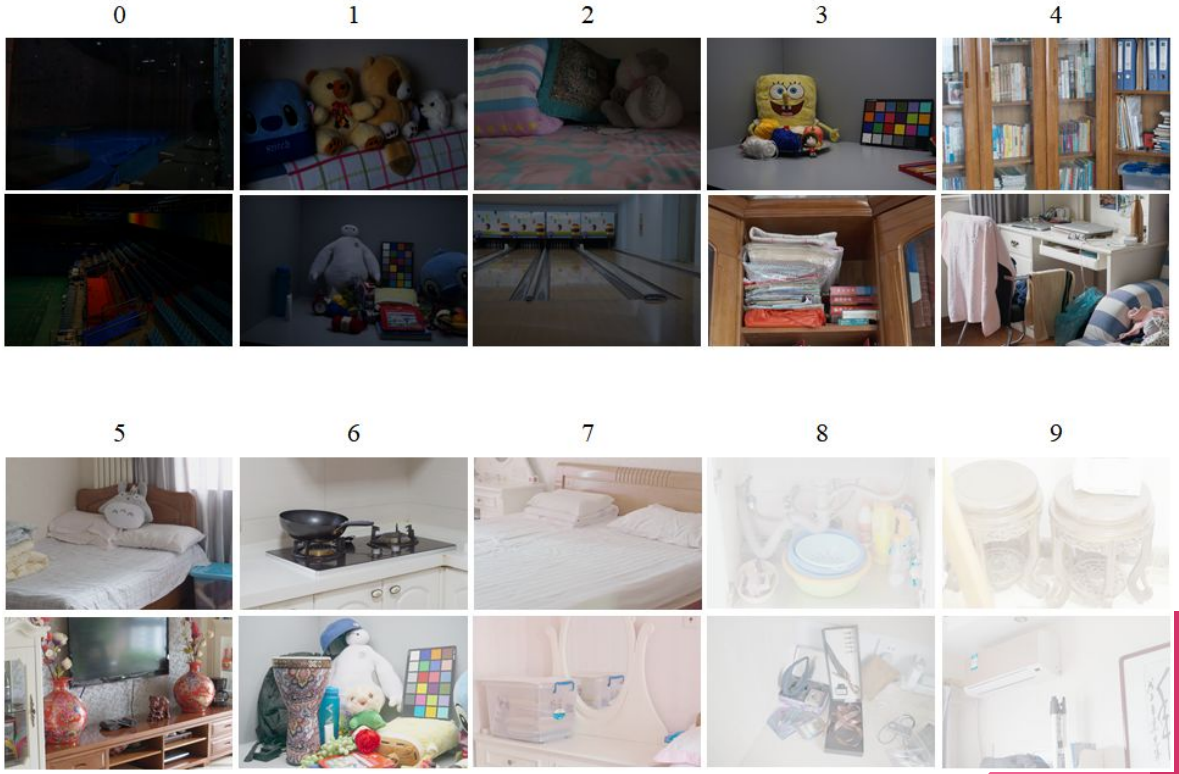
Image synthesis

- Image synthesis is based on linear and gamma transformation:

$$\boxed{I_{out} = \beta\ (\alpha\ I_{in})^{\gamma}}$$

  - I_{out}, I_{in} are the input and output images (scaled between 0, 1)
  - Low-light image : $\alpha \sim$ U(0.9, 1.0) , $\beta \sim$ U(0.5, 1.0) , $\gamma \sim$ U(1.5, 5.0)
  - High-exposure image : $\alpha \sim$ U(0.9, 1.0) , $\beta \sim$ U(0.9, 1.0) , $\gamma \sim$ U(0.3, 1.0)

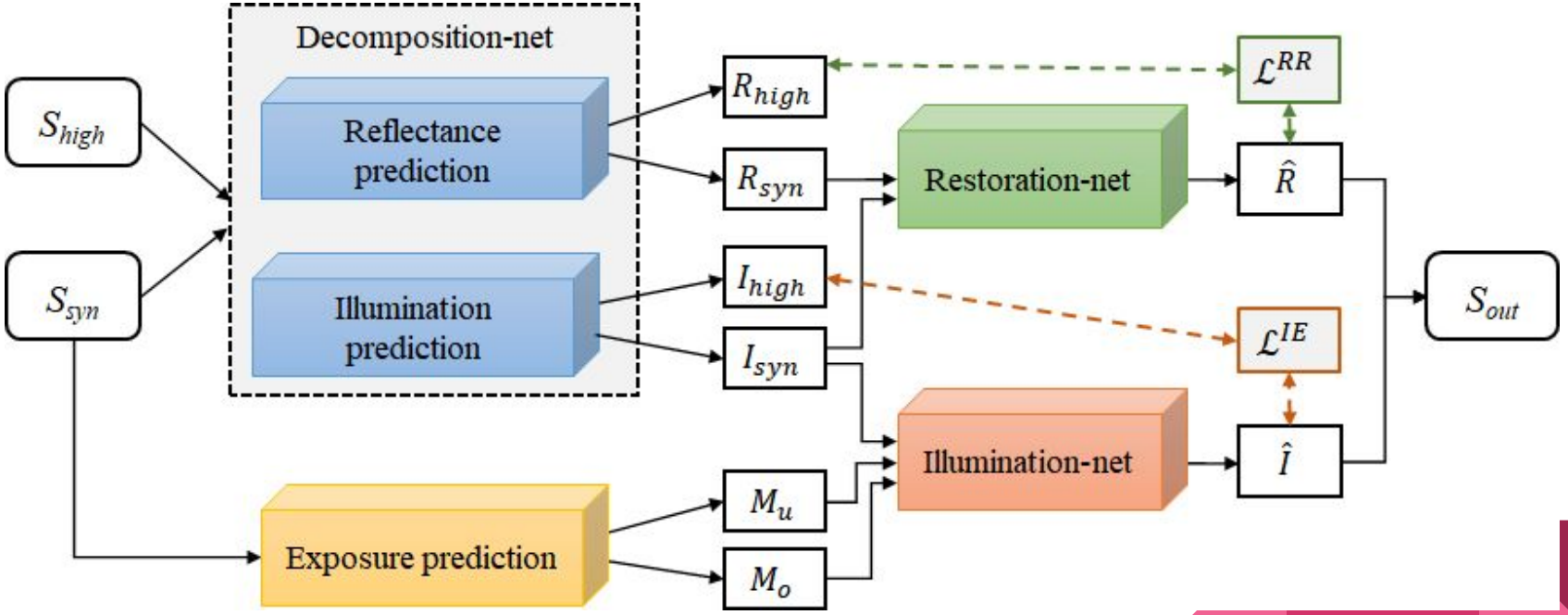
Dataset - sample images



Proposed method: Steps

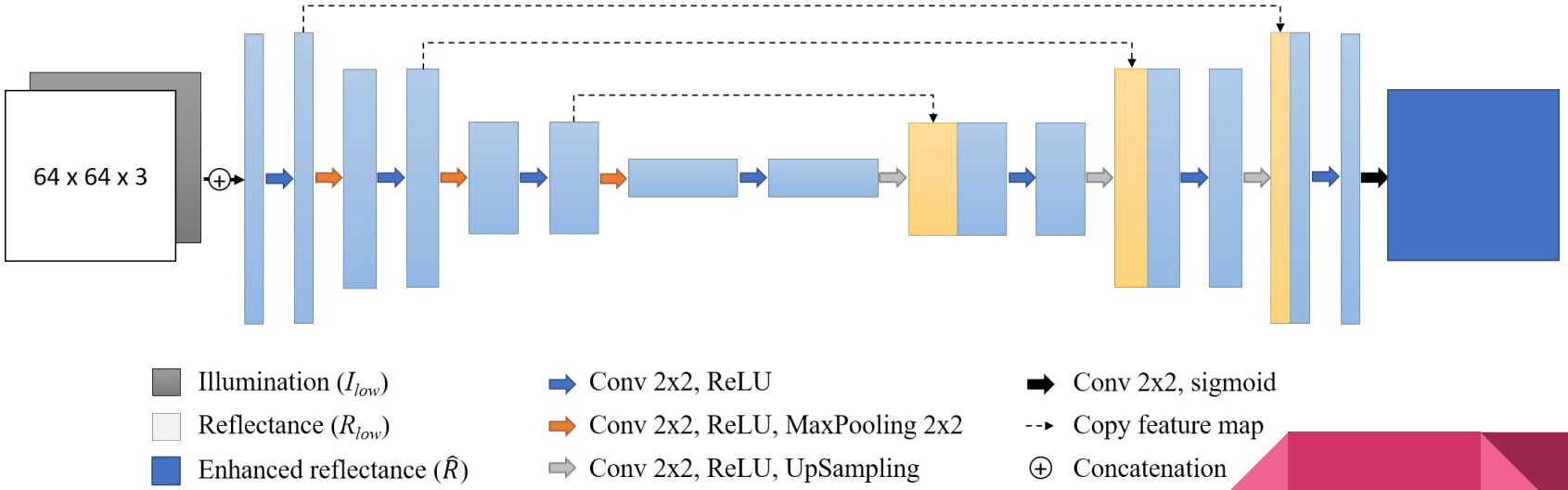
1. Decompose the input image as Reflectance (R) and illumination (I).
2. Predict over-exposure (M_o) and under-exposure (M_u) maps using the exposure prediction model.
3. Enhance reflectance using “restoration-net”.
4. Adjust illumination using “illumination-net”.
5. Fuse enhanced reflectance and illumination to form the output image.

Component analysis: Architecture



Component analysis: Restoration-net

- Illumination-aware reflectance restoration.



$$L^{RR} = \left\|\hat{R} - R_{high}\right\|^2 + \left\|\nabla\hat{R} - \nabla R_{high}\right\|^2 - \mathrm{SSIM}(\hat{R}, R_{high})$$

Proposed method: Exposure prediction

- Overexposure M_o and underexposure M_u maps are predicted based on
  - Saturation : $\sqrt{(a^2 + b^2)}$
  - Intensity : G(L)

- Overexposed = Low saturation + High intensity
- Underexposed = Low saturation + Low intensity

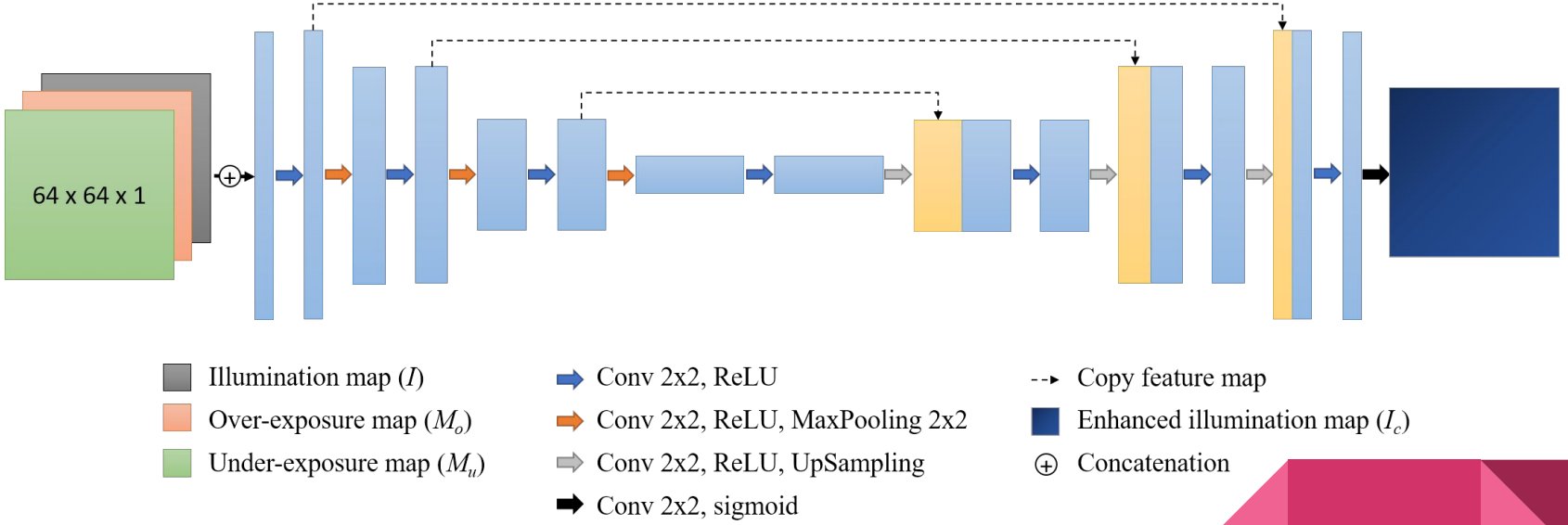
$$M_u = 0.5 + 0.5 \tanh\left(\delta_u \left(L_{max} - \left(\sqrt{a^2 + b^2} + G(L)\right)\right)\right)$$

$$M_o = 0.5 + 0.5 \tanh\left(\delta_o \left(L_{min} - \left(\sqrt{a^2 + b^2} - G(L)\right)\right)\right)$$

LAB values scaled (0,1)

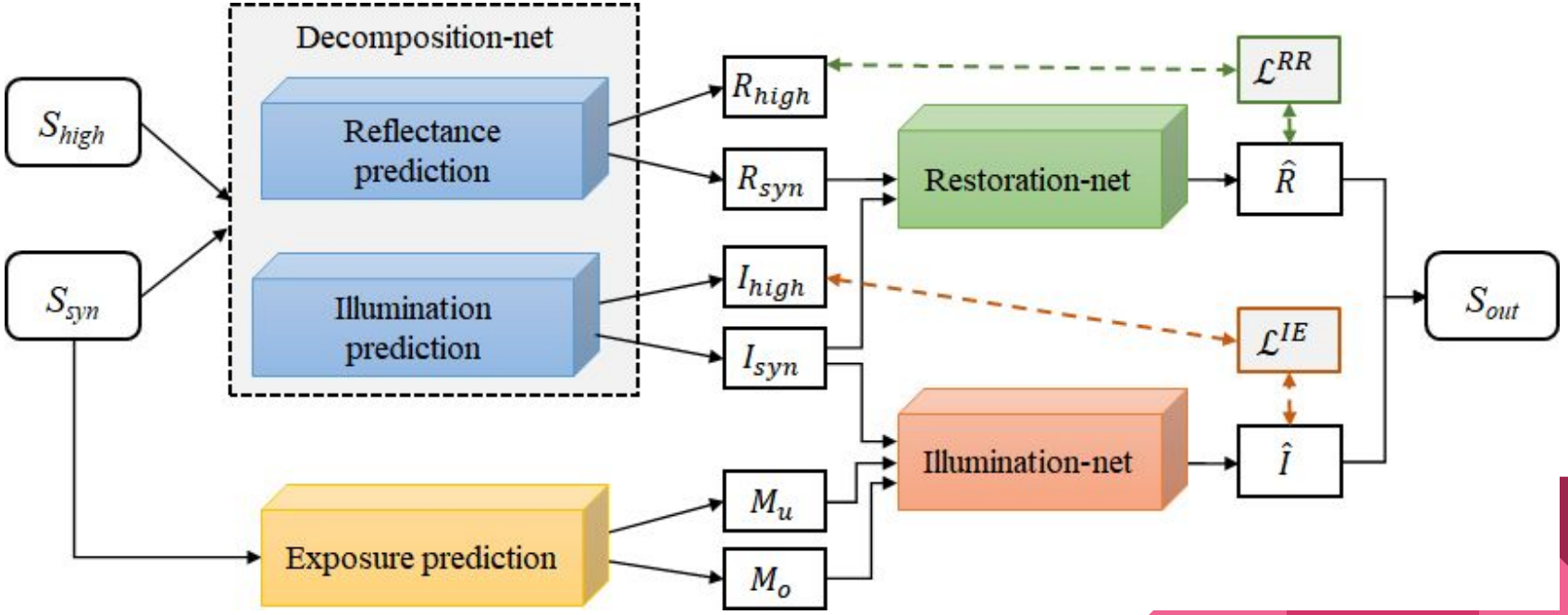
Component analysis: Illumination-net

- Exposure-aware illumination enhancement.



$$L^{IE} = \left\|\hat{I} - I_{high}\right\|^2 + \left\|\nabla\hat{I} - \nabla I_{high}\right\|^2$$

Component analysis: Architecture

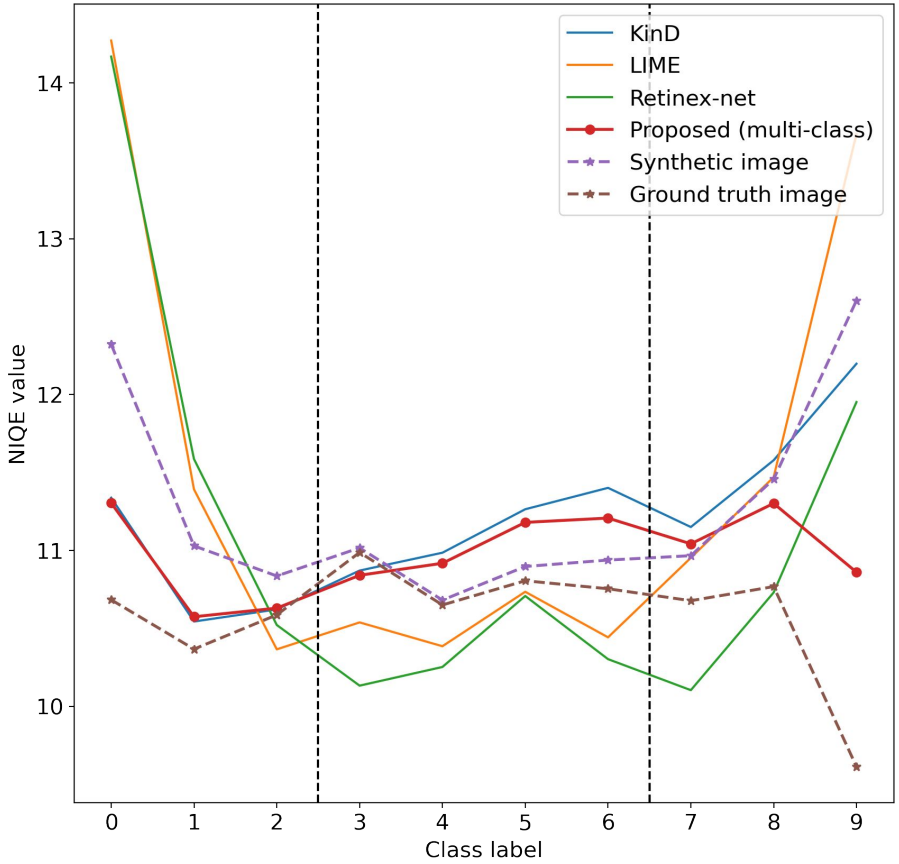


Output for input images of varying illumination

- Low-light = class 0-2
- Normal light = class 3-6
- Well-lit = class 7-9

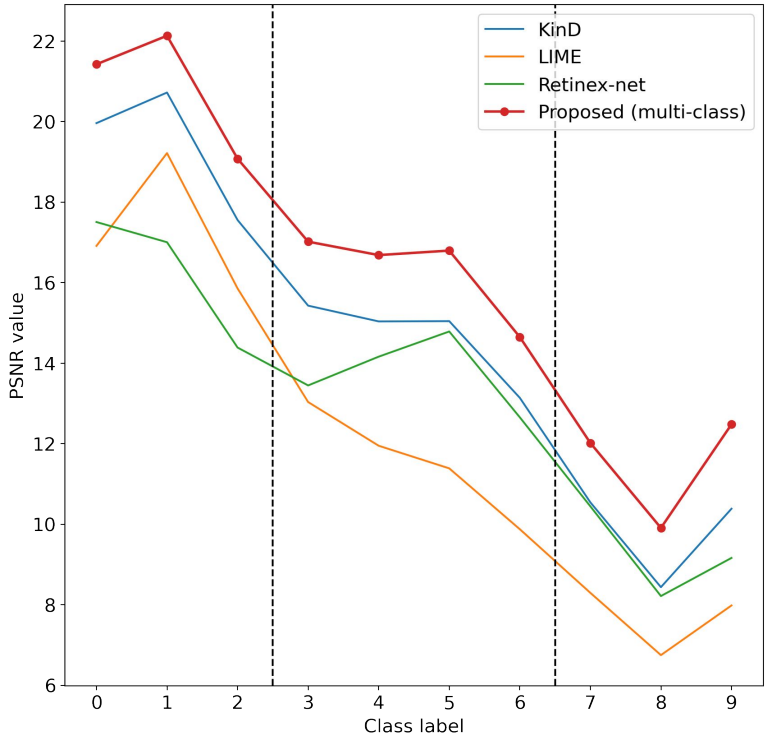
Variation of NIQE with illumination class

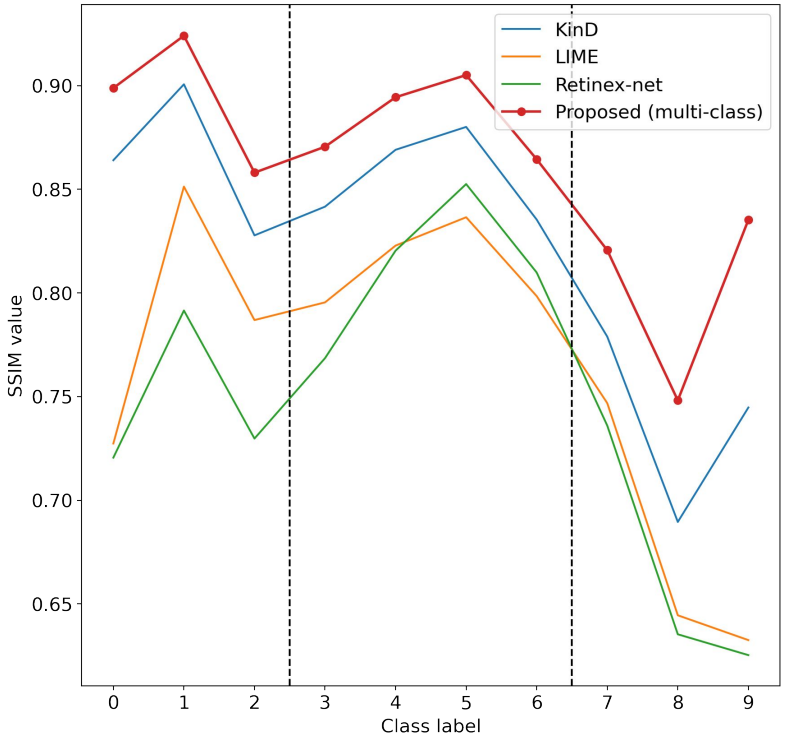
- NIQE - Lower is better
- Proposed model improves performance in both extreme cases (0 and 9)



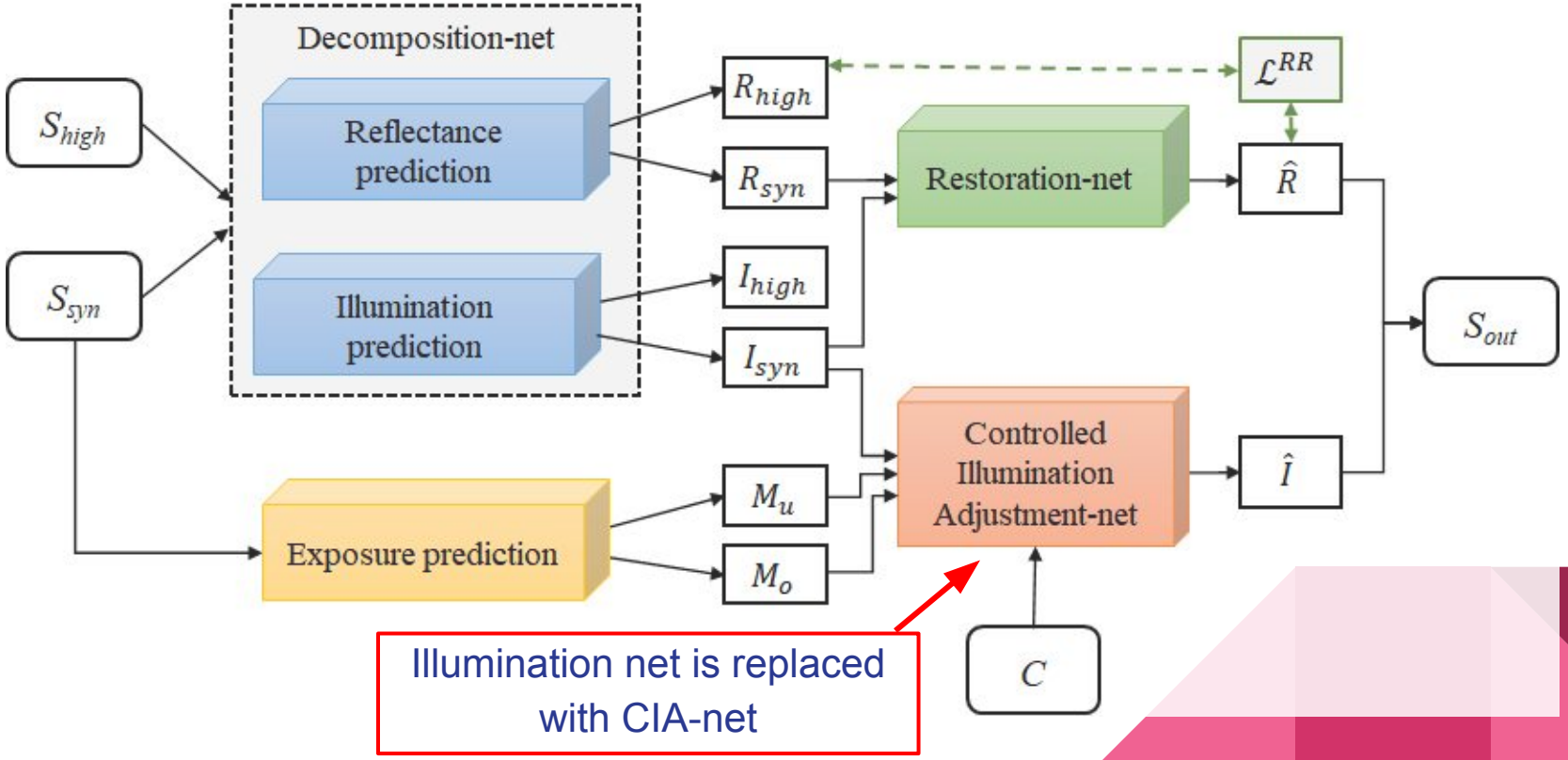
Similarity between ground truth and enhanced image

- SSIM, PSNR - Higher is better





Multi-class image generation - Architecture



Illumination net is replaced with CIA-net

Controlled illumination adjustment

- Objective : Produce images of desired (multi-)illumination levels.
- Since reflectance stays same (∵ retinex theory), only illumination-net is replaced.
- However, the dataset does not contain single image taken in multiple illumination levels. ∴ **Supervised learning is NOT possible**.
- Therefore, conditional generative model with function : $\mathrm{CIA}(\cdot)$

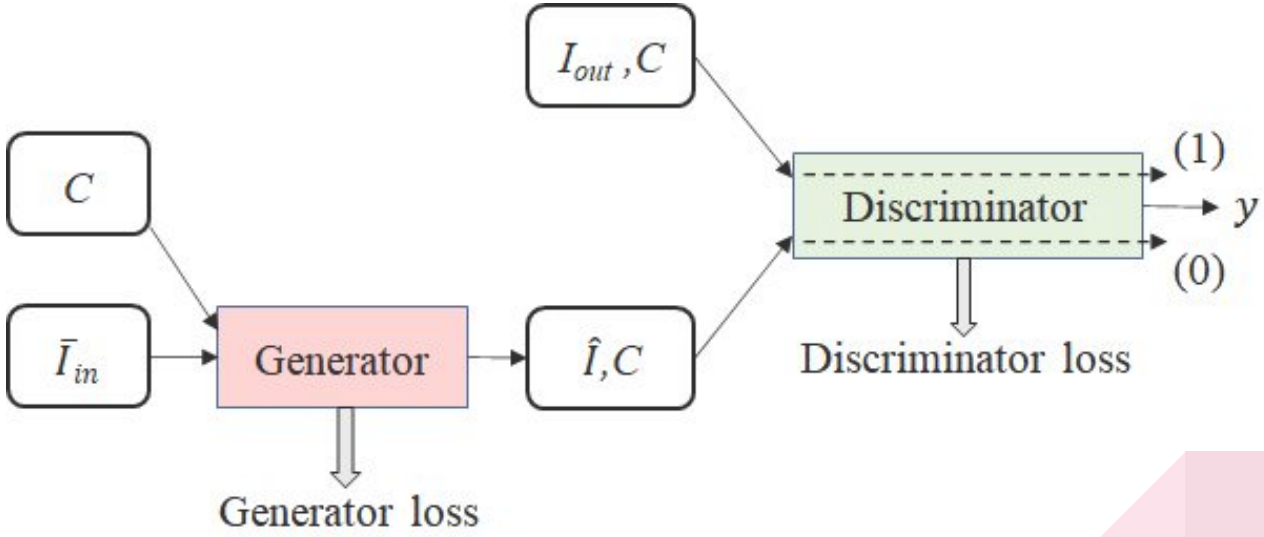
$$\mathrm{CIA}(\bar{I}_{syn}, C_{high}) = I_{high}$$

$$\mathrm{CIA}(\bar{I}_{high}, C_{syn}) = I_{syn}$$

Input information — Class label — Output image

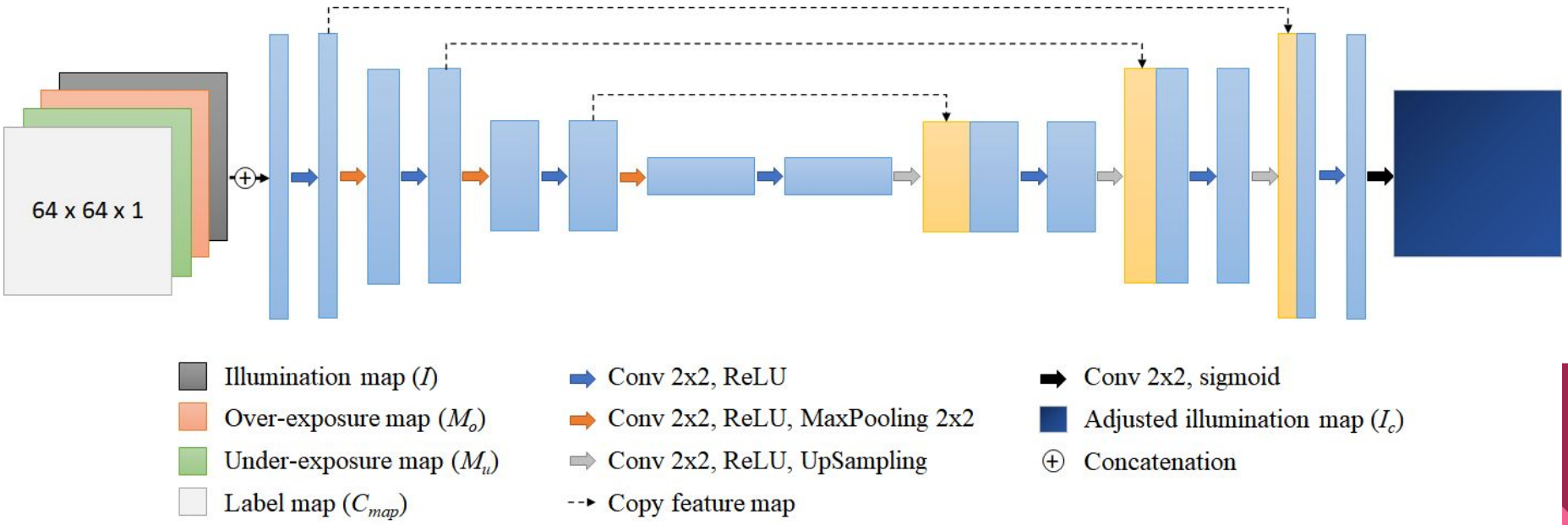
CIA-net - cGAN

- cGAN - Both generator and discriminator take condition (C) as input.

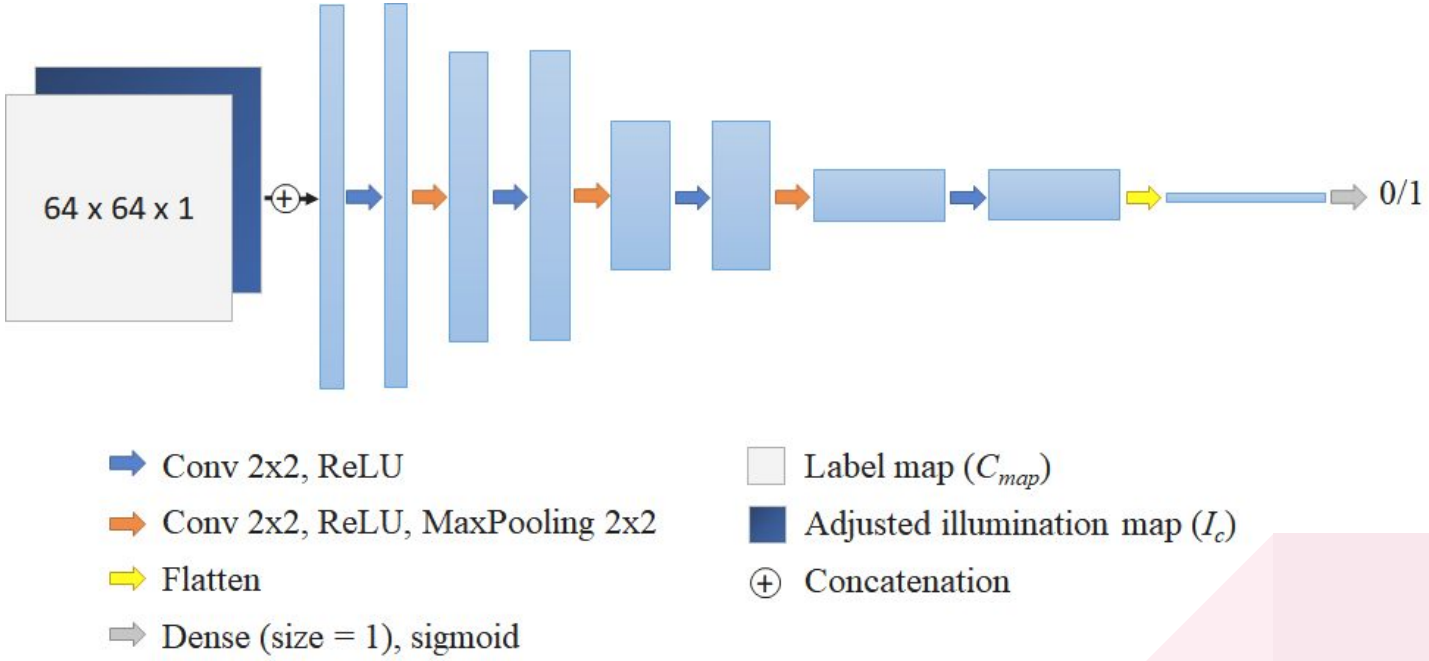


CIA-net - Generator

- Label map : created by passing the label vector of size (10,1) into an embedding layer and then reshaping it to the required shape (64, 64, 1).



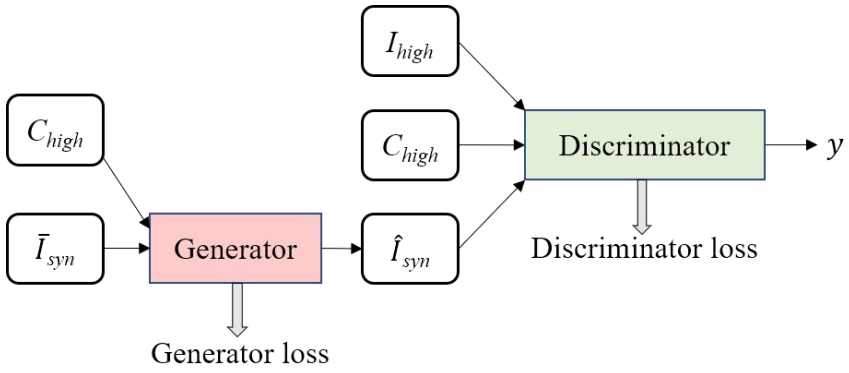
CIA-net - Discriminator

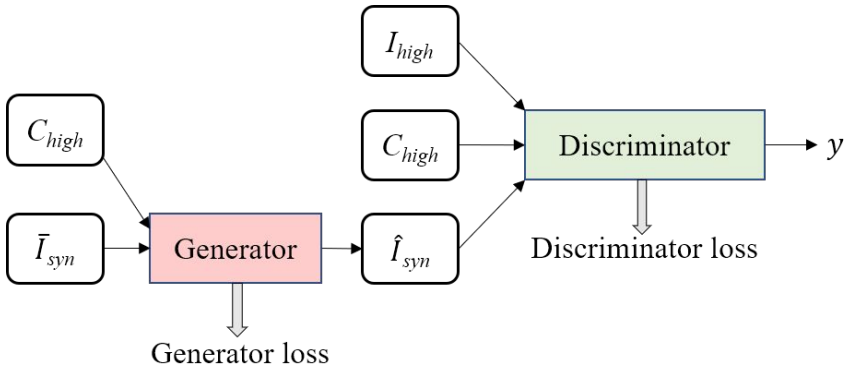


CIA-net - Loss function

$$\mathrm{CIA}(\bar{I}_{syn}, C_{high}) = I_{high}$$

$$\mathrm{CIA}(\bar{I}_{high}, C_{syn}) = I_{syn}$$





Discriminator

$$L_D = H\left(D_{CIA}\left(\mathrm{CIA}(\bar{I}_{syn}, C_{high}), C_{high}\right), 0\right) + H\left(D_{CIA}(I_{high}, C_{high}), 1\right)$$

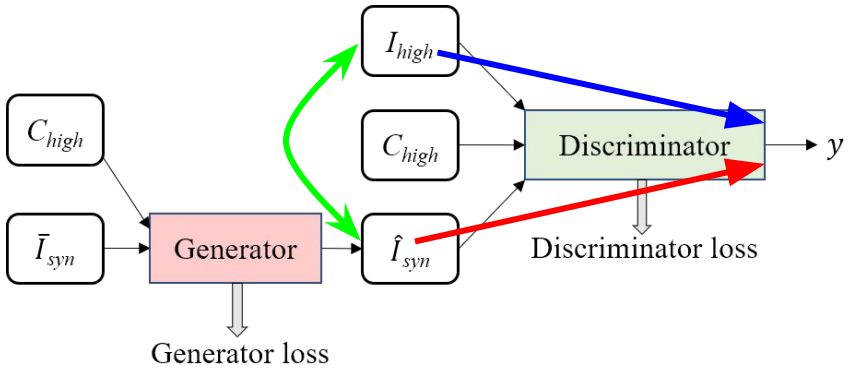
$$L_G = E\left[\left\|\mathrm{CIA}(\bar{I}_{syn}, C_{high}) - I_{high}\right\|^2\right] + L_D$$

Generator

CIA-net - Loss function

$$\mathrm{CIA}(\bar{I}_{syn}, C_{high}) = I_{high}$$

$$\mathrm{CIA}(\bar{I}_{high}, C_{syn}) = I_{syn}$$



$$L_D = H\left(D_{CIA}\left(\mathrm{CIA}(\bar{I}_{syn}, C_{high}), C_{high}\right), 0\right) + H\left(D_{CIA}(I_{high}, C_{high}), 1\right)$$

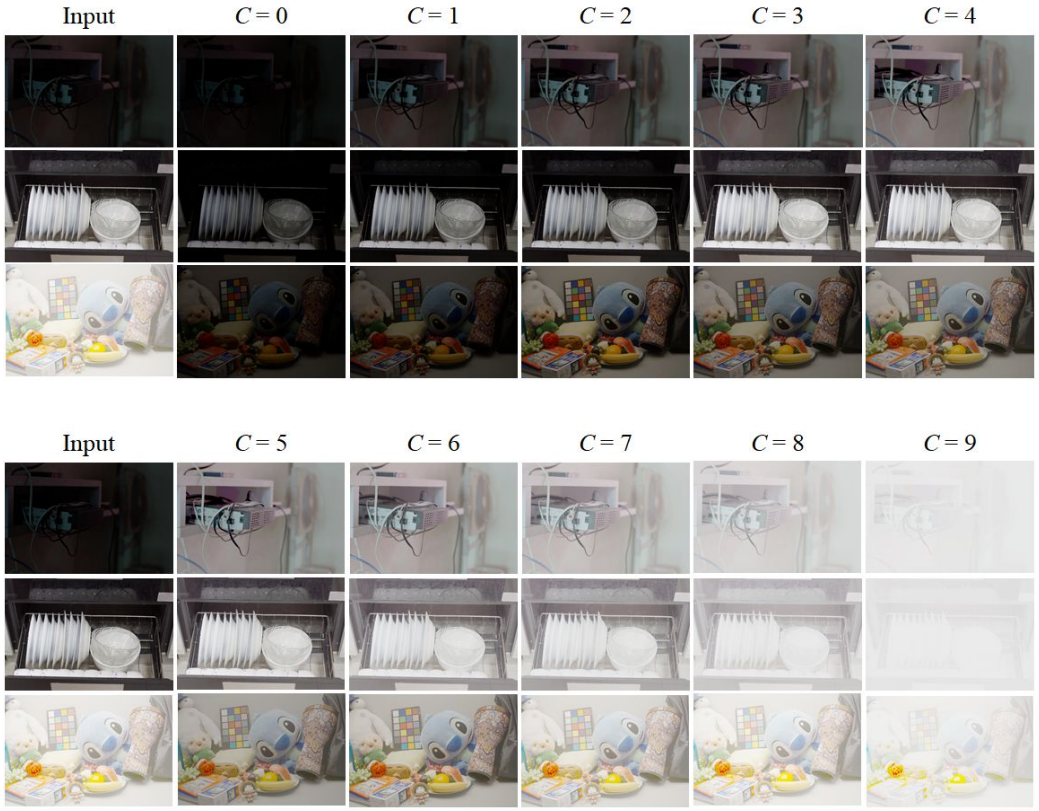
Discriminator

$$L_G = E\left[\left\|\mathrm{CIA}(\bar{I}_{syn}, C_{high}) - I_{high}\right\|^2\right] + L_D$$

Generator

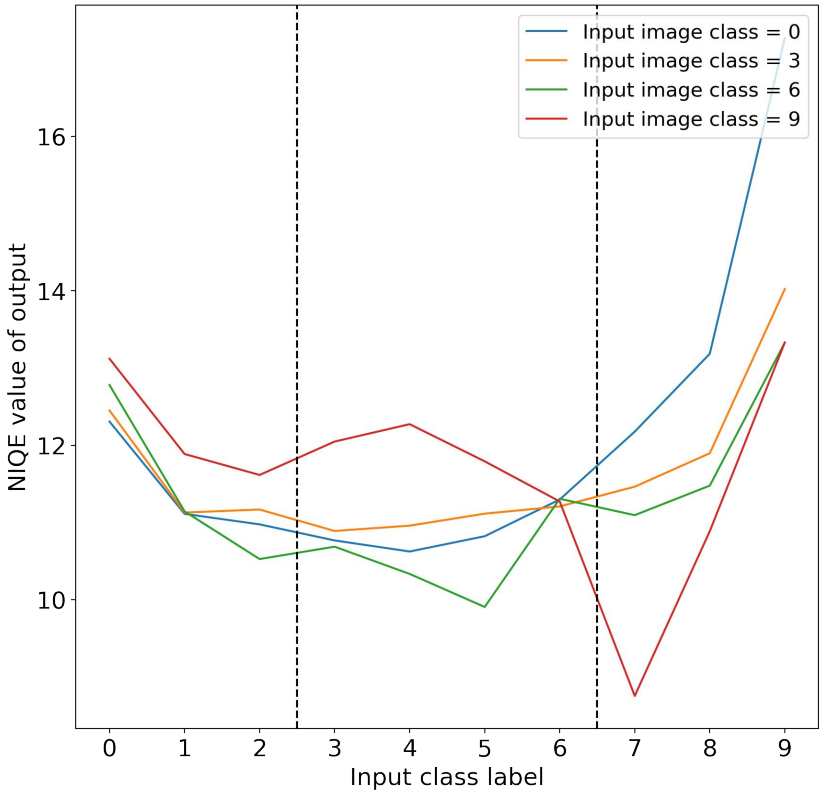
Multi-class images from CIA-net

- Images of each classes are generated from
  - Low light (Row 1, 4)
  - Normal light (Row 2, 5)
  - High light (Row 3, 6)
- Classes are shown as columns (C=1, 2, …, 9)



Variation of NIQE for different input classes

- NIQE - Lower is better
- The image specially shows poor quality in the following cases
  - Low-light images from well-lit images (from 9 to 0)
  - High-light images form low-light images (from 0 to 9)



Conclusion and Futureworks

- The proposed multi-class enhancement model improves efficacy compared to the two-class enhancement models.
- The conditional illumination enhancement model produces sufficiently good results. The nature of the output images must be further studied.

- Fully supervised models are dependant on paired dataset. The dependency on synthetic data must be eliminated for better performance.
- Cycle-consistency should be studied and incorporated.
- Lack of related literature acts as a barrier for comparison. Thus evaluation metrics must be established.

Interpretation of Low Light Images

Presented by E/14/158

High level questions

- Can we classify images in low light conditions?
- Can we classify images in low light conditions when we have training datasets in well light conditions?
- Can we detect suspicious behaviour (anomalies) in low light conditions when we have know what is normal in well light conditions?

Problem definition : One class anomaly detection in low light conditions

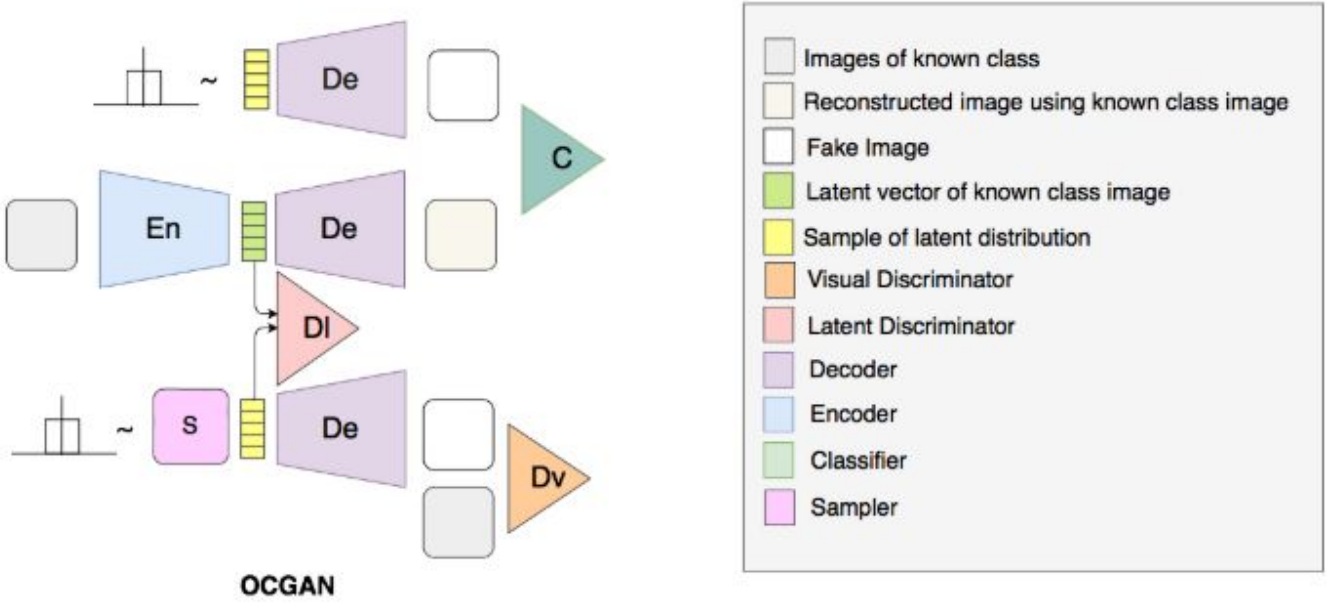
Detecting whether a testing image comes from a given training class or not when the training class in in well light condition and the testing image is in low light conditions

Training data : dogs (well light)

Related work (1/2)

1. **Retinex decomposition network (discussed earlier)** : used to extract the lighting condition invariant information (reflectance) from images.
2. **One Class Generative Adversarial Network (Perera et. al. 2019)**. OCGAN transforms the input images to a latent representation which can be used for either reconstruction or classification. Then we train a discriminator to classify images looking at the latent representation. The discriminator is trained with a training dataset (in class) and artificially data (anomaly) generated through negative mining (an optimization technique).

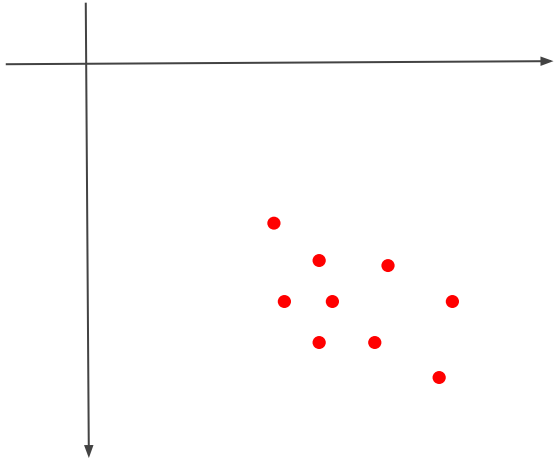
Related work (2/2) : OCGAN



Proposed Solution (1/3)

Retinex Decomp Network (encoder)

Intermediate representation

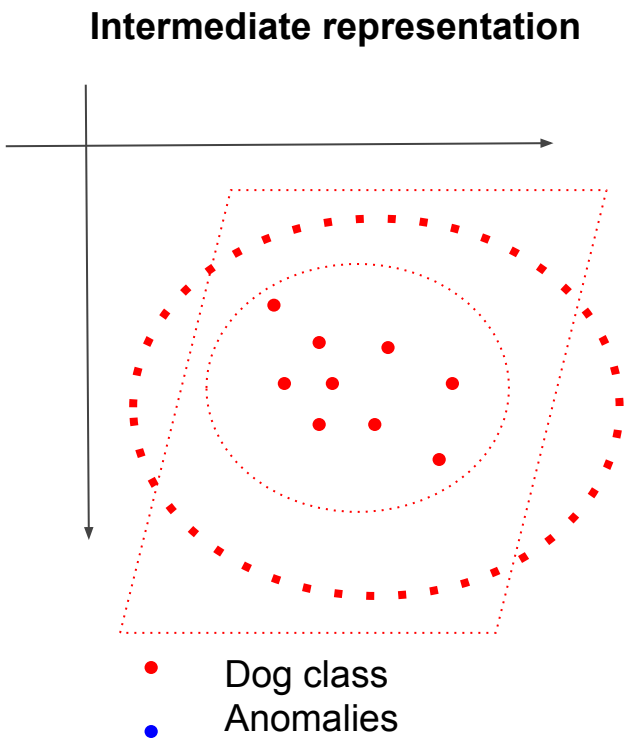


Dog class

Anomalies

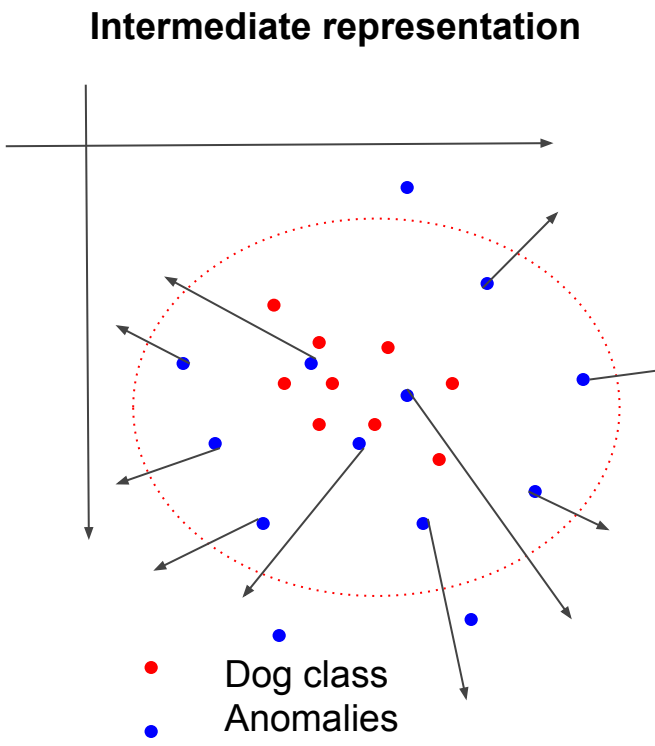
Proposed Solution (2/3)

- When we only have one class of datapoints, we cannot determine what the decision boundary should be.
- Assume that the Discriminator initializes with the thick decision boundary.



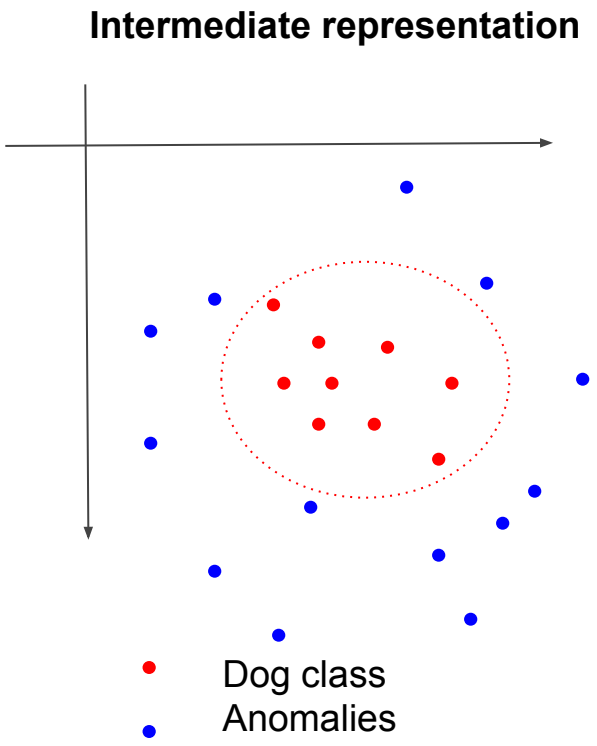
Proposed Solution (3/3)

- Using negative mining, we generate anomalies.
- We use two negative mining algorithms
    - Gradient descent inspired.

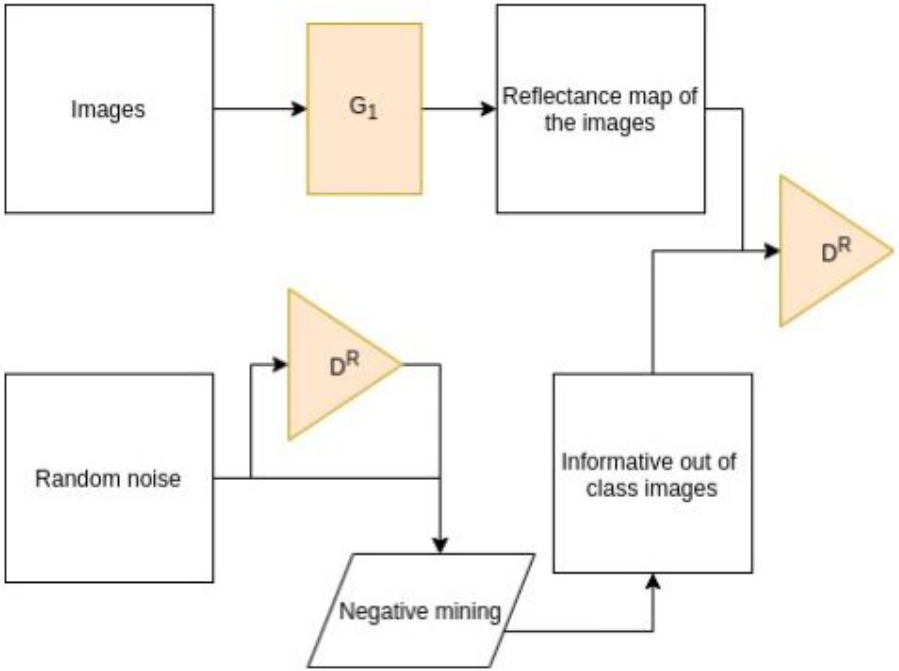


Proposed Solution (3/3)

- We update the decision boundary based on the training data and artificial data.
- These steps are repeated until we have a good decision boundary, capable of detecting out of class anomalies successfully.



Neural Network Architecture

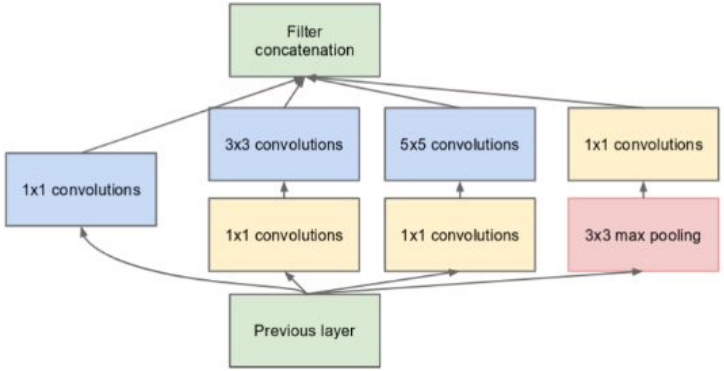


G_1 is the encoder (notation is consistent with the light enhancement section)

D^R is the Discriminator working on the reflectance maps. (This is inspired by the inception v3 architecture)

Discriminator : Modified Inception Architecture

- We use only 2 inception blocks.
- We use only one output neuron.
- NN is trained to minimize binary cross entropy.



Inception block

Negative Mining Algorithm : Based on gradient

Data: D^R

Result: R_0

$\mathrm{R}_0 \leftarrow sample\_from(\mathbb{R}^{H \times W \times 3})$

while $D^R(R_0) < 0.5$ **do**

  $R_0 \leftarrow sample\_from(\mathbb{R}^{H \times W \times 3})$

end

for $i = 0;\ i < 5;\ i = i + 1$ **do**

  $Pred_0 \leftarrow D^R(R_0)$

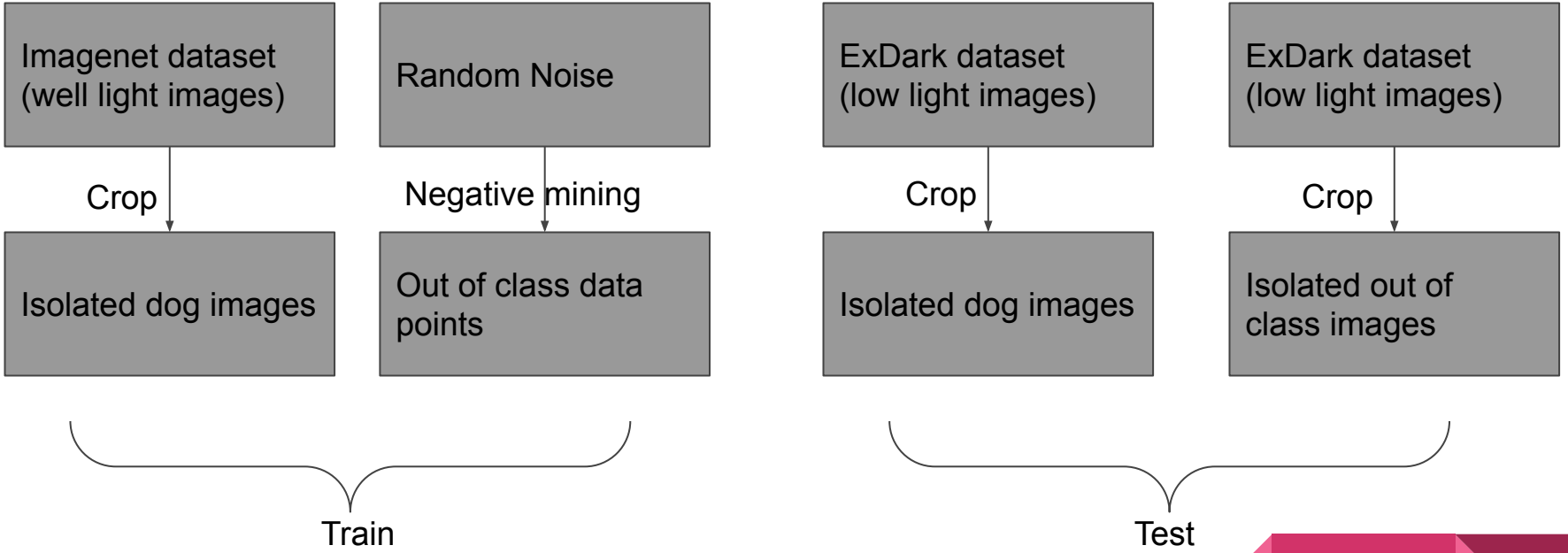
  $R_0 \leftarrow gradient\_descent(Pred_0, R_0, D^R)$

end

Negative mining algorithm : based on random

Data: D^R
Result: R_0
$\mathrm{R}_0 \leftarrow sample\_from(\mathbb{R}^{H \times W \times 3})$
while $D^R(R_0) < 0.5$ **do**
  $R_0 \leftarrow sample\_from(\mathbb{R}^{H \times W \times 3})$ **end**
  $Pred_{00} = D^R(R_0)$ **for** $i = 0;\ i < 5;\ i = i + 1$ **do**
    $step \leftarrow sample\_from(\mathbb{R}^{H \times W \times 3})$
    $Pred_{01} \leftarrow D^R(R_0 + step)$
    if $Pred_{01} < Pred_{00}$ **then**
      $R_0 \leftarrow R_0 + step$
      $Pred_{00} \leftarrow D^R(R_0)$
    end
  end

Experiments



We do similar experiments with other classes (dogs are used as an example here)

Results

True Class	Predicted class	
	Dog	Other
Dog	98.35	1.65
Other	95.10	4.90

Results using the gradient descent based negative mining.

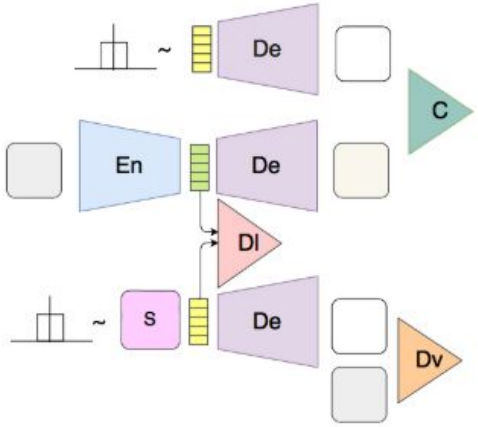
True Class	Predicted class	
	Dog	Other
Dog	83.60	26.40
Other	87.00	13.00

Results using the random walk based negative mining.

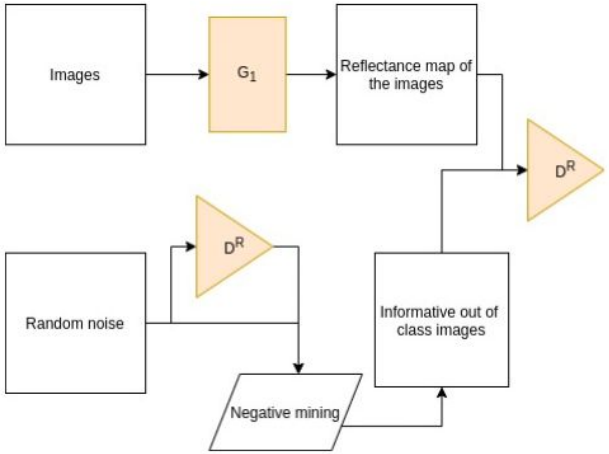
The results are not sufficient for a usable anomaly detection system.

Discussion and Future Work (1/4)

- The intermediate representation used in this work is 300x300x3 (reflectance). OCGAN has proposed for a 4x4x64 intermediate latent representation. It appears as if we need to further encode the images before discrimination step.

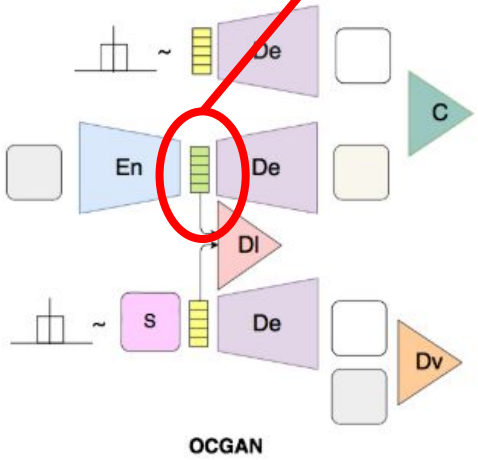


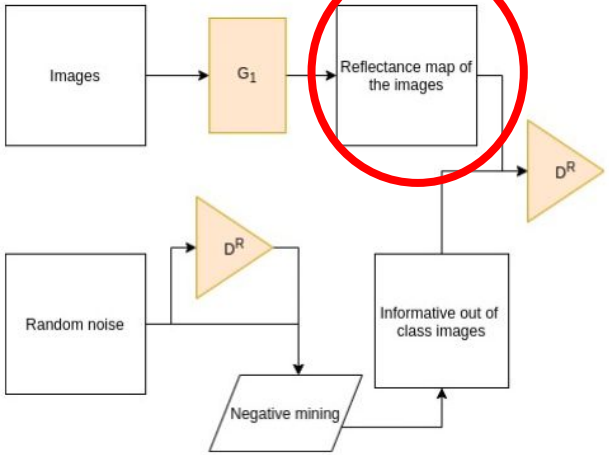
OCGAN



Discussion and Future Work (2/4)

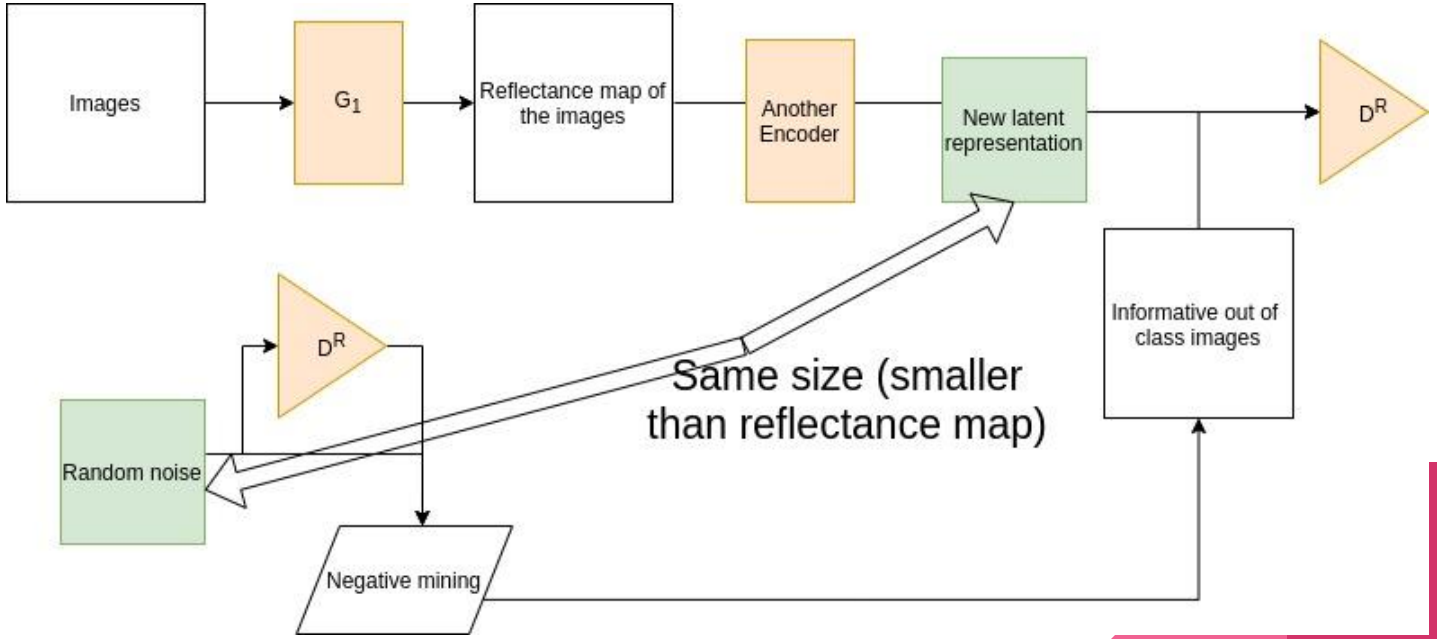
- The intermediate representation used in this work is **300x300x3** (reflectance). OCGAN has proposed for a **4x4x64** intermediate latent representation. It appears as if we need to further encode the images before discrimination step.





Discussion and Future Work (3/4)

Future work: Introduce another encoding step.



Discussion and Future Work (4/4)

- Negative mining is not good enough to get meaningful out of class images.

  = The “generated anomalies” are not representative of the actual anomalies
  = Cats are closer to dogs than the generated anomalies that were supposed to represent cats.

- **Future work:** Initialize the negative mining process with some anomaly examples.

Concluding remarks

- The proposed model for enhancement combines existing ideas from **Retinex theory, CNN, and CycleGAN.**
- Using **both** paired (synthetic + non-synthetic) and unpaired (non-synthetic) images, the model provides better performance in comparison.
- The multi-class illumination model further improves the performance in comparison to the two-class problem.
- The proposed model for one class anomaly detection (incorporating retinex theory, decomp net, negative mining and inception network) should be improved.

Thank you

Q & A

Summary ---> Q and A

- Image enhancement algorithms are important for 2 reasons:
  - **Enhancement** (Improving image aesthetics)
  - **Interpretation** (Application of computer vision algorithms)
- **Retinex theory :** Image = Reflectance x Illumination
- **Light enhancement** : Retinex decomp (CNN) + enhancement (GAN)
- **Controlled enhancement** : Retinex decomp + cGAN
- **One class novelty detection** : Retinex decomp + Discriminator (Inception based) uses negative mining for training.