UNIVERSITAS INDONESIA
Veritas, Probitas, Iustitia
FAKULTAS ILMU KOMPUTER

Image Restoration and Image Quality

Dr. Eng. Laksmita Rahadianti, Muhammad Febrian Rachmadi, Ph.D.,
Dr. Dina Chahyati, Prof. Dr. Aniati M. Arymurthy

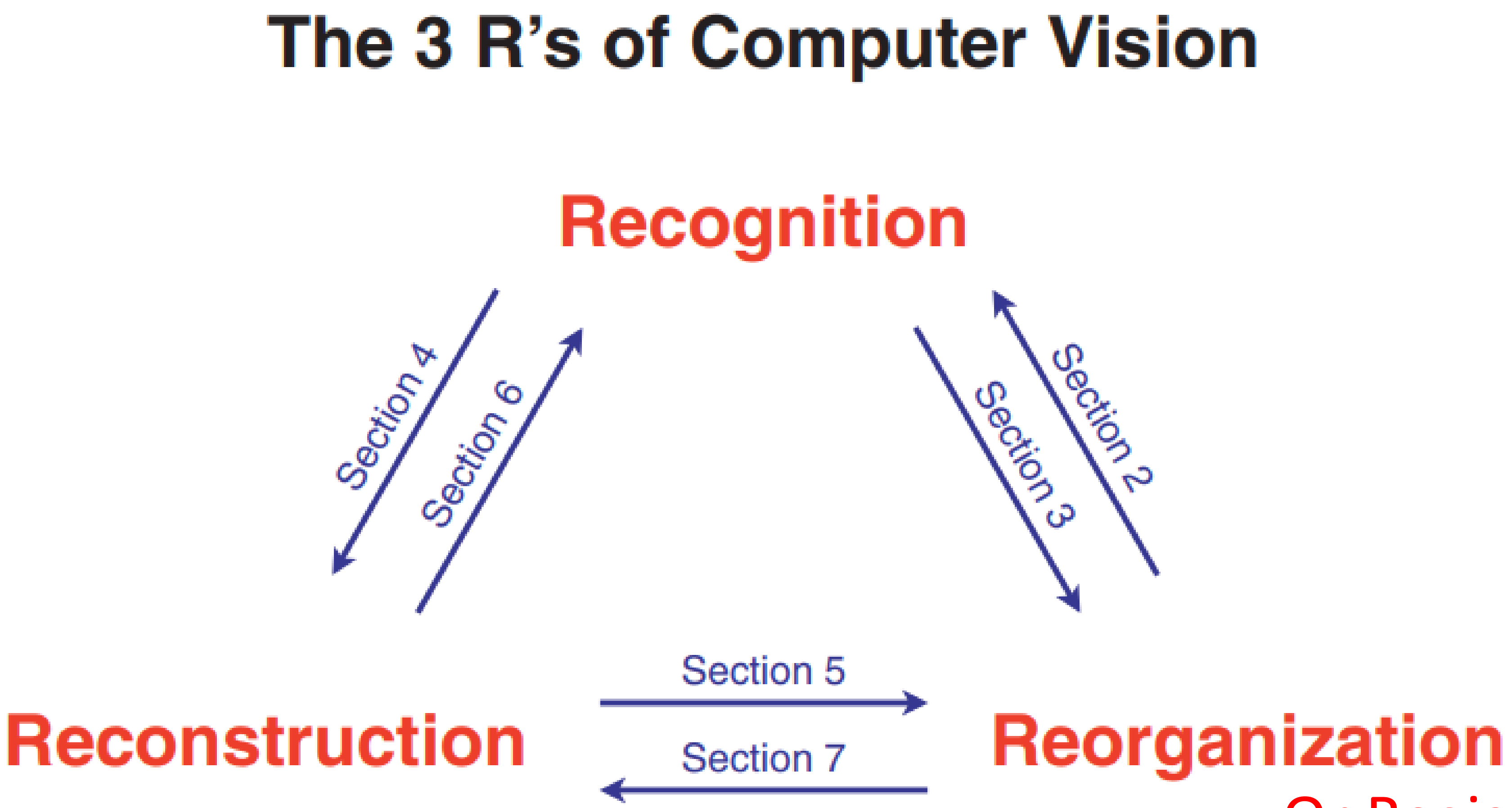
CSCE604133 Computer Vision
Fakultas Ilmu Komputer Universitas Indonesia

Acknowledgements

- These slides are created with reference to:
  - Computer Vision: Algorithms and Applications, 2nd ed., Richard Szeliski https://szeliski.org/Book/
  - Digital Image Processing, Gonzales and Woods, 3rd ed, 2008.
  - Course slides for CSCE604133 Image Processing – Faculty of Computer Science, Universitas Indonesia
  - Introduction to Computer Vision, Cornell Tech https://www.cs.cornell.edu/courses/cs5670/2024sp/lectures/lectures.html
  - Computer Vision, University of Washington https://courses.cs.washington.edu/courses/cse576/08sp/

The 3 R's of Computer Vision

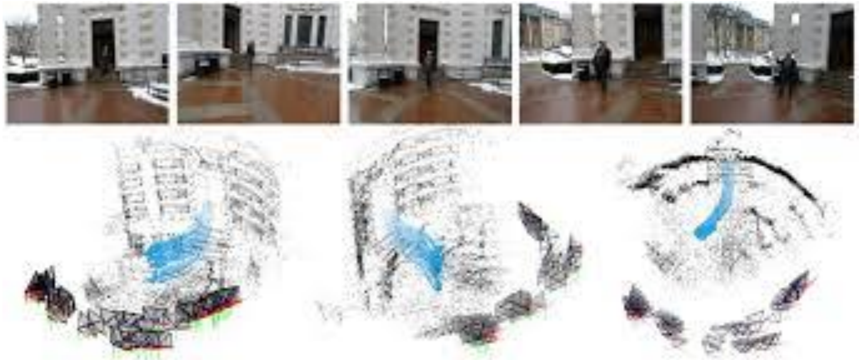
Malik, Jitendra, et al. "The three R's of computer vision: Recognition, reconstruction and reorganization." *Pattern Recognition Letters* 72 (2016): 4-14.



Or Registration

Reconstruction

- Reconstructing info from images to another form
  - 3D reconstruction



  - Image restoration / recovery

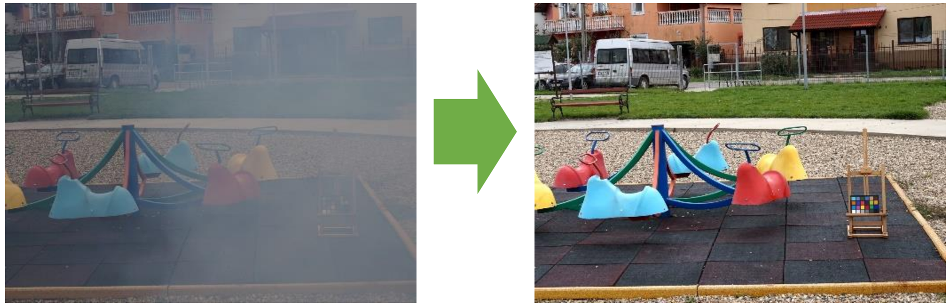


Image Restoration

Image Restoration vs Enhancement

Both restoration and enhancement do this!

- Image enhancement is done heuristically, via **trial and error,** until we obtain a good image
- In image restoration **we attempt to estimate the model of distortion**

Image Restoration

Estimate

- Estimate the distortion / degradation

Restore

- Using that pre-knowledge, recover the original image

Why is Restoration Needed?

- Computational approaches to understand the 3D scene from 2D images

Are there any objects in the scene?

What objects are in the scene?

Where (how far from the camera) is the object?

How is the scene structure?

- Often using geometric cues such as lines or corners

Why is Restoration Needed?(2)

- Computational approaches to understand the 3D scene from 2D images

Are there any objects in the scene?

What objects are in the scene?

Where (how far from the camera) is the object?

How is the scene structure?

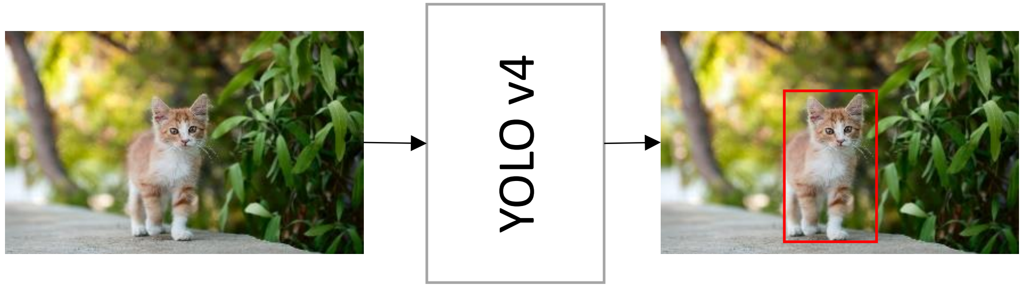
- Degraded images make computer vision difficult!!!

[1]Bochkovskiyet, et al. (2020)l; [2]Wang, et al. (2024)

YOLOv4 for Object Detection in Hazy Images

Ferro Geraldo Hardian, Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2020)

- YOLOv4[1] is an object detection model – since has evolved to YOLOv9[2]
- YOLOv4 is supposed to be powerful.. can we use it for degraded (hazy) images?



- YOLOv4 is still able to detect objects, but will fail on very dense haze:

Best→

Worst→

Fail→

YOLOv4 for Object Detection in Hazy Images

Ferro Geraldo Hardian, Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2020)

- What is the best combination to train the model?
  - Trained only with clear – is unable to detect objects
  - Trained with Hazy images and Clear+Hazy is successful in **some** cases
  - Multiple combinations tried
- *The model is very overfit* to the dataset
  - **Very difficult to find a single solution for all images**

Clear→

Hazy→

Clear+
Hazy→

Trust Issues in CV: Automated Vehicles

- Studies show a basic resistance in the majority of car users to handing over control of their vehicle to a machine. Conversely, other current polls have demonstrated that, particularly for drivers aged 19 to 31 years, driving itself is often bothersome.
- Perceived negative features: a. Social consequences, **b. Safety risks,** c. Data abuse, d. Rise in costs, and e. Uncertainties;
- How can we incite trust in using automated vehicles?

A possible mitigation is to display the restored model with the inference to the human

Image Degradation

Model of Image Degradation

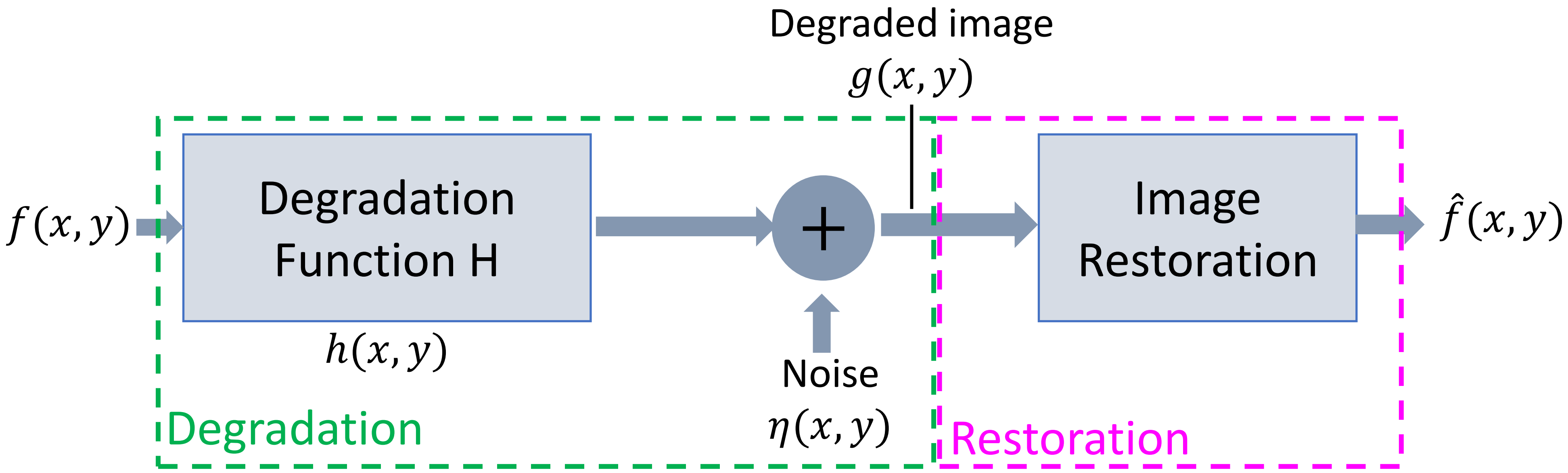
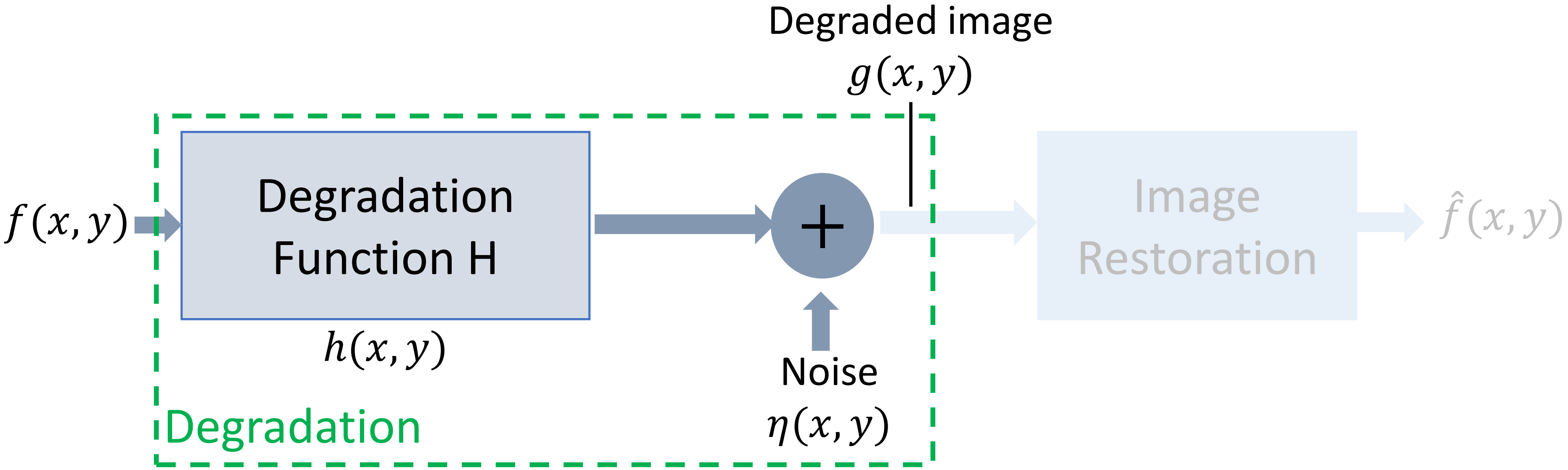


Image Degradation



- Spatial domain

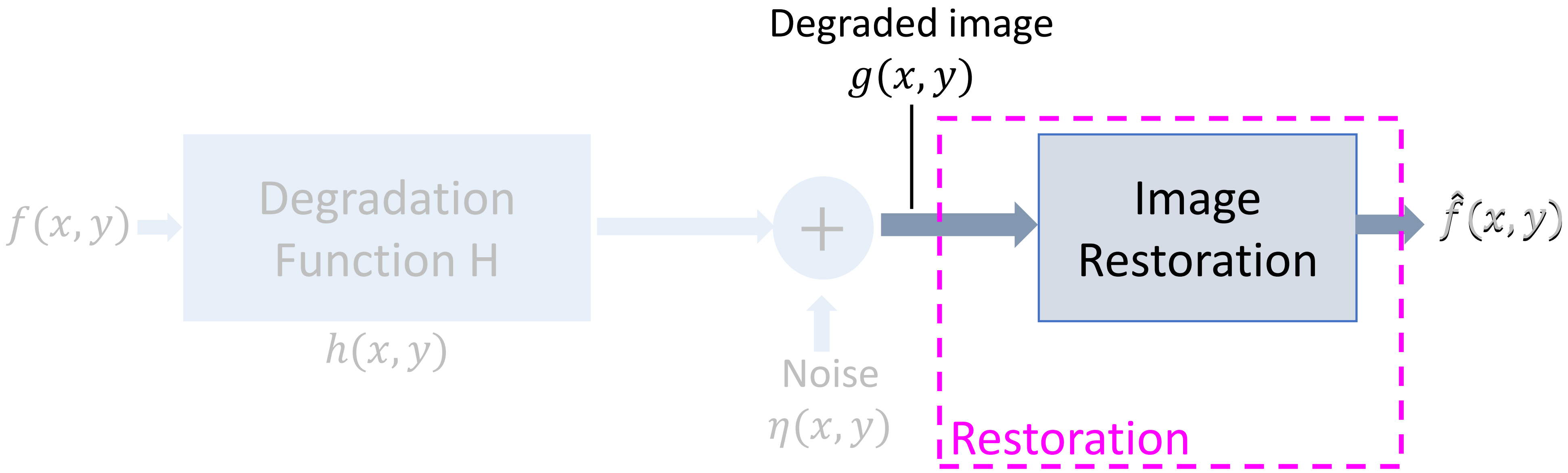
$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$$

- Frequency domain

$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

- $g/G(x,y)$: degraded image
- $h/H(x,y)$: degradation function
- $f/F(x,y)$: original image
- $\eta/N(x,y)$: noise

Image Restoration



- Image restoration attempts to recover a degraded image using **prior knowledge** of the distortion
- We can estimate the **noise** $\eta(x,y)$ or the **degradation** $h(x,y)$

Estimating the Noise Model

$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$$

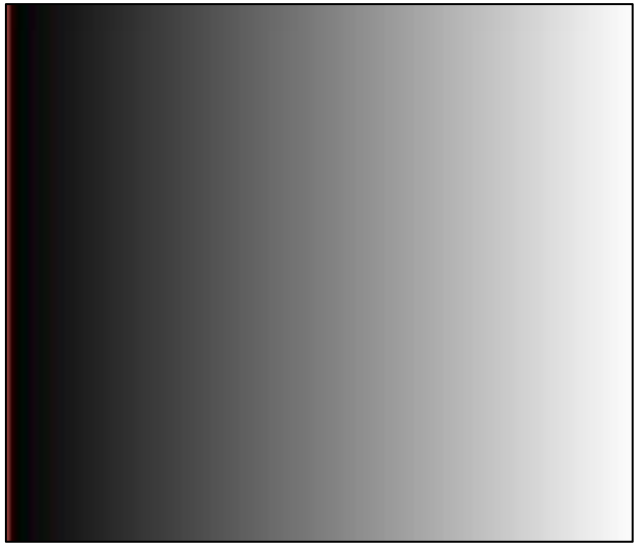
$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

- Recall:
  - Image restoration attempts to recover a degraded image using **prior knowledge** of the distortion.

We can attempt to estimate the noise model $\eta/N(u,v)$

Noise Estimation- If the imaging system is available

- Capture a plain **known** surface
- We can then obtain the pattern of the noise

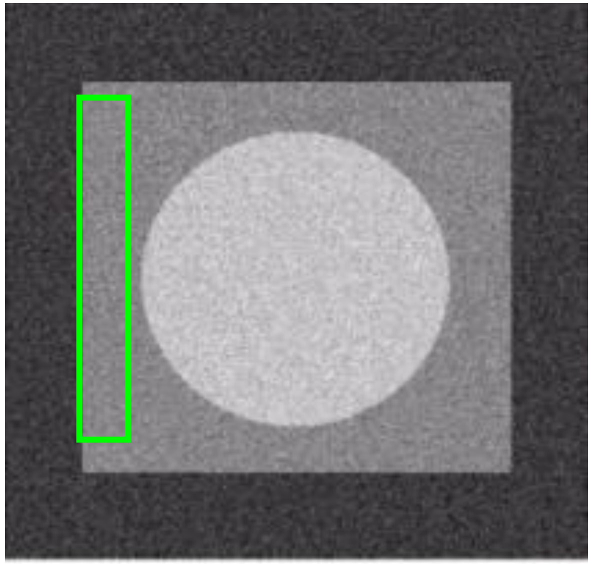


Clear Image

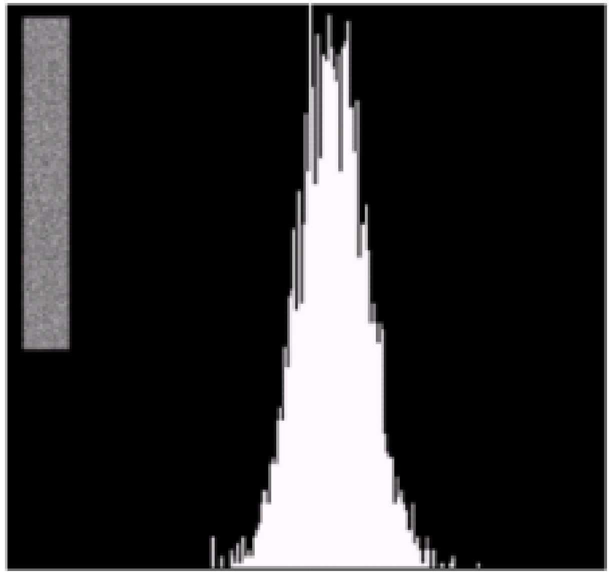
Captured Image

Noise Estimation- If the imaging system is not available

- Use the captured images
- Take strips of a reasonably constant intensity
- Observe the histogram of the strips



Captured Image

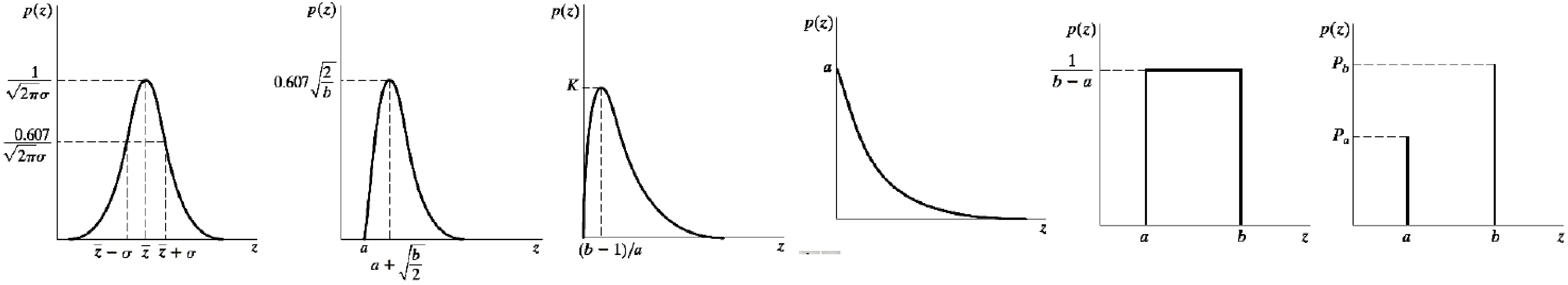


Histogram

Noise model $\eta(u,v)$ - Spatial Noise

- Spatial Noise
  - Can be described statistically as random variables, described by their probability density functions (PDF)

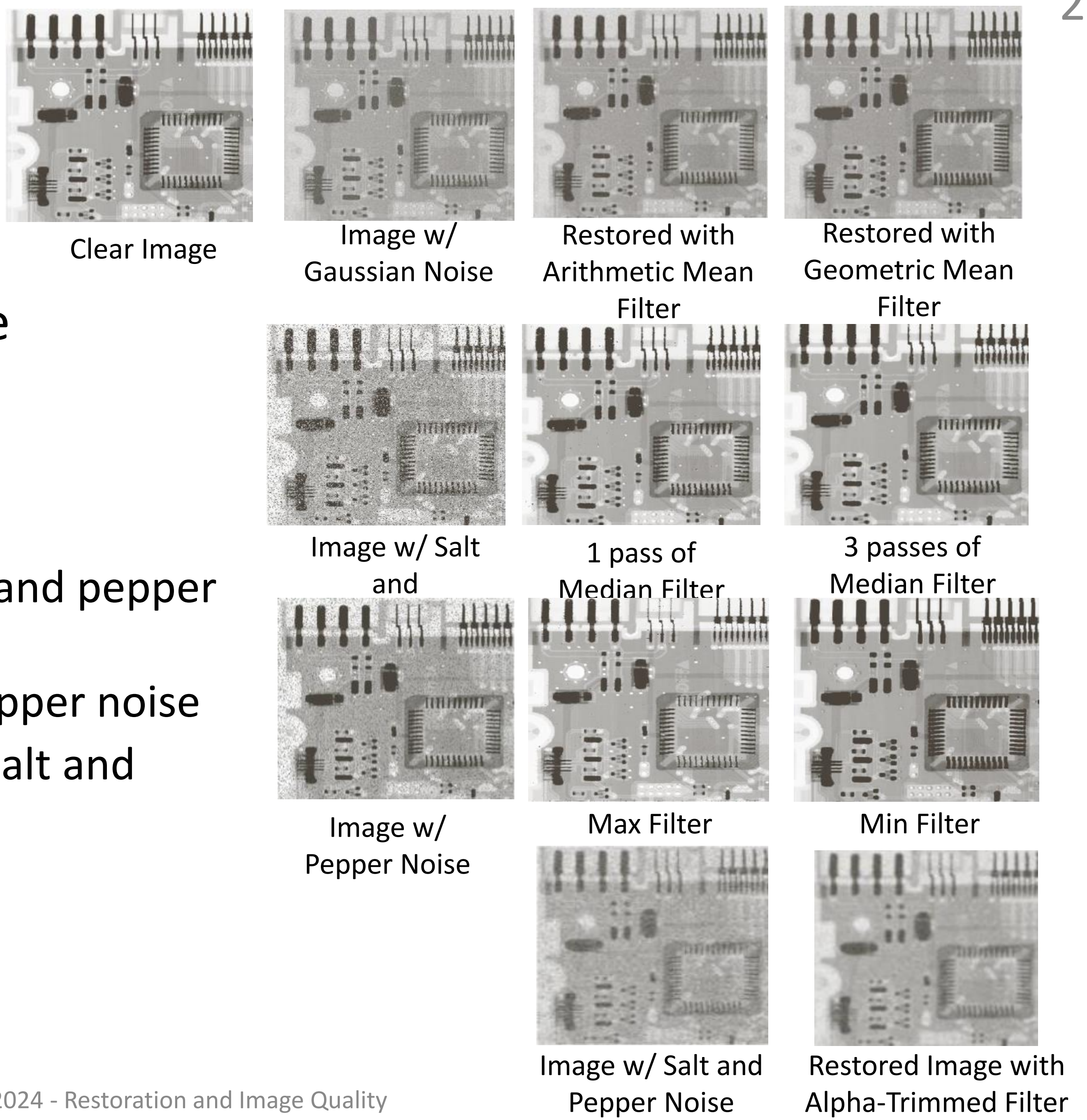
Spatial noise – can be corrected with **spatial filters.**



Gaussian/Normal Rayleigh Erlang/Gamma Exponential Uniform Impulse

Spatial Filtering

- Mean Filters – for gaussian noise
  - Arithmetic Mean Filter
  - Geometric Mean Filter
- Order Statistic Filters
  - Median Filters – for gaussian/salt and pepper noise
  - Min / Max Filters – for salt and pepper noise
  - Alpha Trimmed Mean Filter – for salt and pepper noise



Wiener Filter (N. Wiener, 1942)

- Or the minimum mean square/least square error filter
- The objective is to find the estimate $\hat{f}$ for a clear image f with the mean square error between them is minimized

$$\boxed{\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] G(u,v)}$$

$H(u,v)$: degradation function
$H^*(u,v)$: complex conjugate of $H(u,v)$
$|H(u,v)|^2$: $H^*(u,v)\,H(u,v)$
K: constant chosen for the best visual result

Clear Image

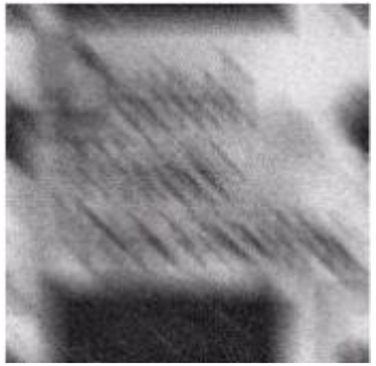


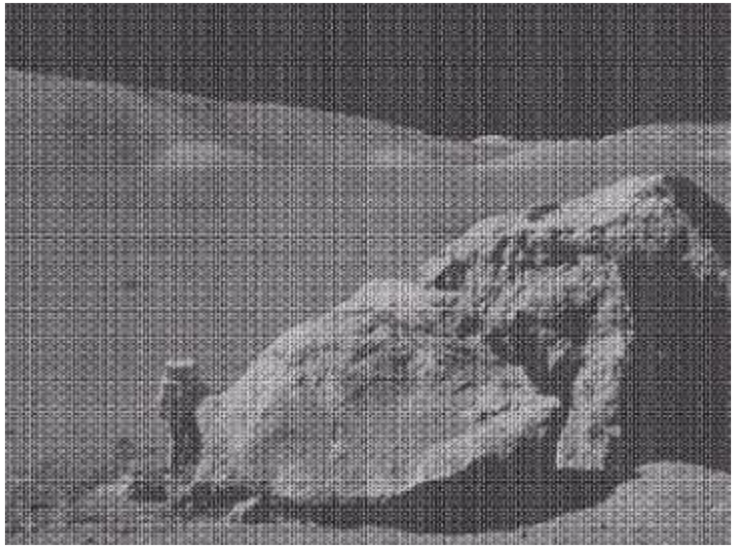
Image w/ Motion Blur and Additive Noise



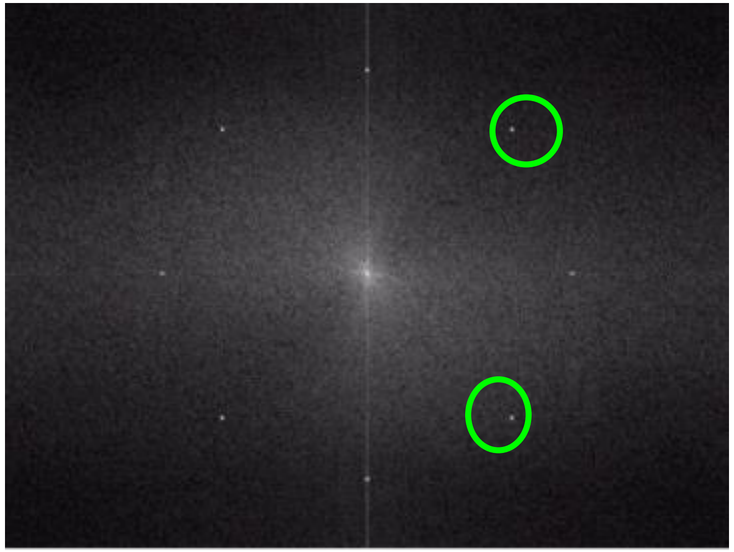
Wiener Filter Results with various noise variance

Noise model $N(u,v)$ – Periodic Noise

- Unlike previously, periodic noise is spatially dependent



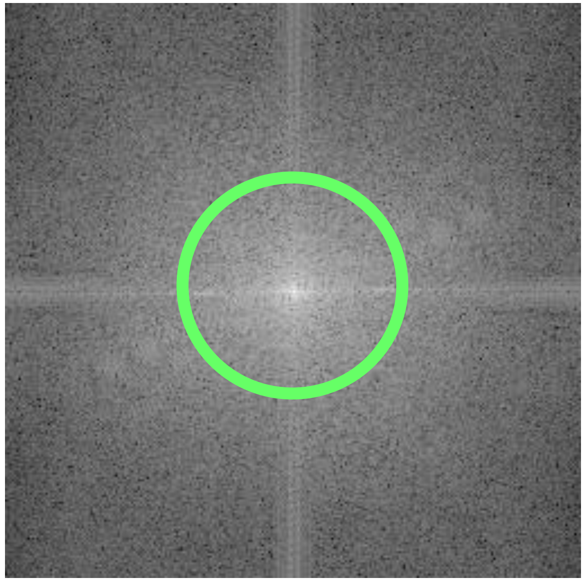
Spatial Image



Fourier Transform

Periodic noise – can be corrected with **frequency domain filtering.**

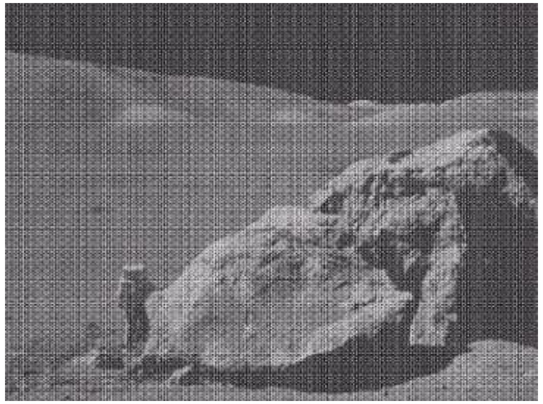
Recall: Images in the Frequency Domain



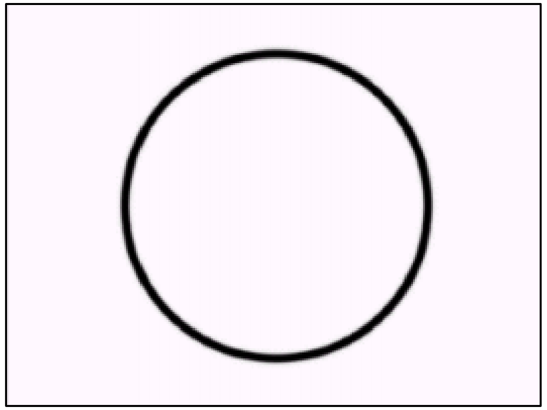
The central part of FT, i.e. the low frequency components are responsible for the general gray-level appearance of an image.

The outer part of the FT, i.e. the high frequency components of FT are responsible for the detail information of an image.

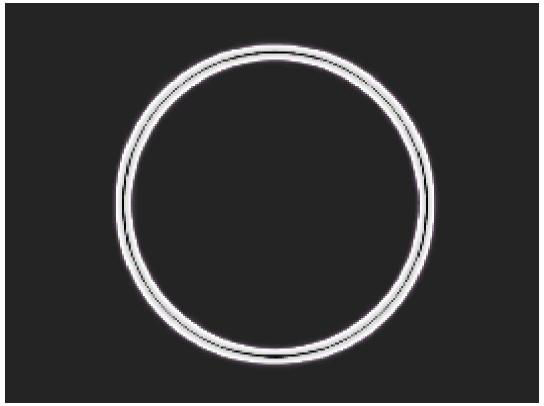
Band Filters



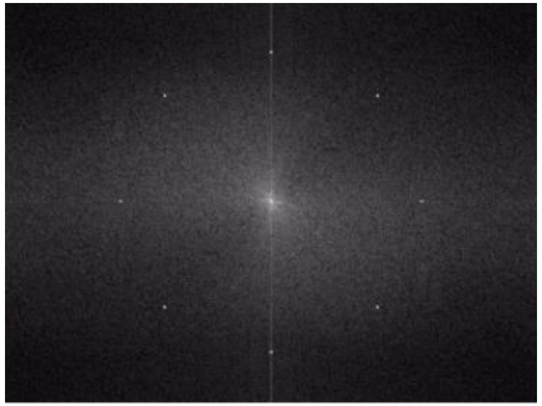
Spatial Image



Band Reject Filter

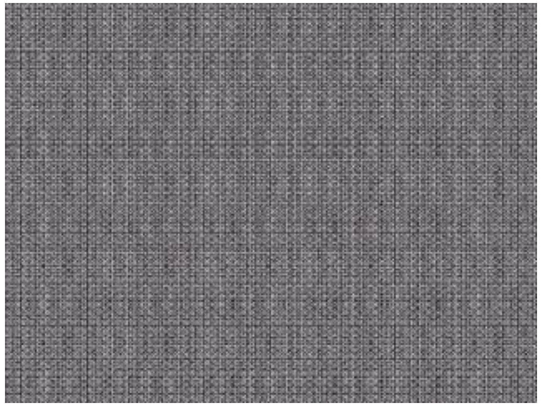


Band Pass Filter



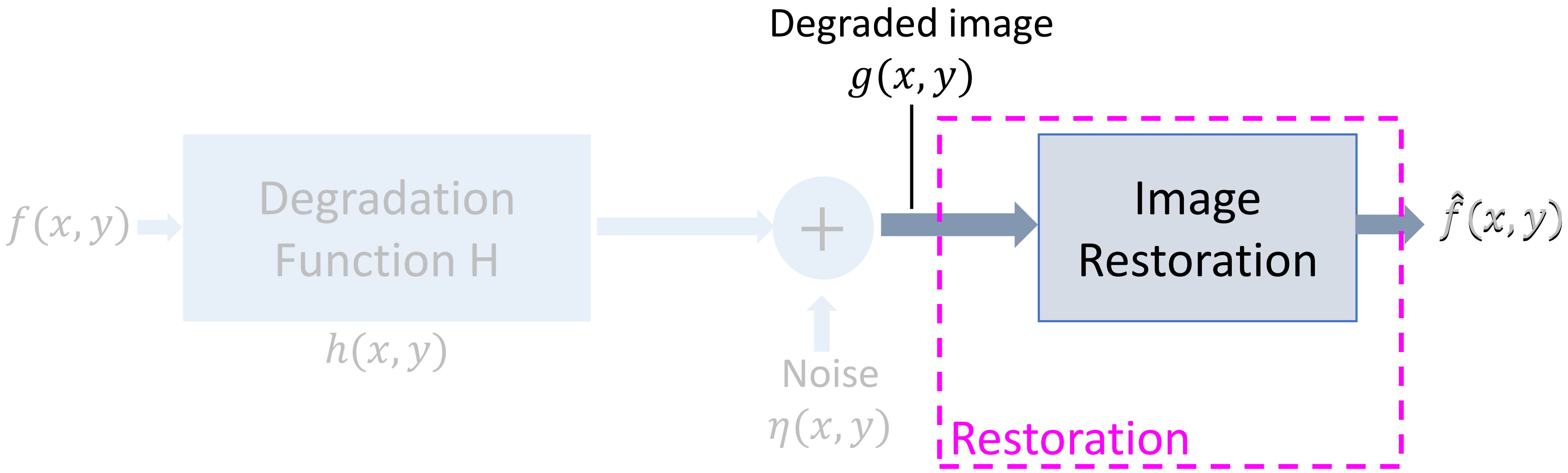
Fourier Transform

Restored Image



Noise Pattern of Image

Image Restoration



- Image restoration attempts to recover a degraded image using **prior knowledge** of the distortion
- We can estimate the **noise** $\eta(x, y)$ or the **degradation** $h(x, y)$

Estimating the Degradation Model

$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y)$$

$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

- Recall:
  - Image restoration attempts to recover a degraded image using **prior knowledge** of the distortion.

We attempt to estimate the degradation model $h/H(u,v)$

- This is often very difficult! – underdetermined, multiple unknowns!

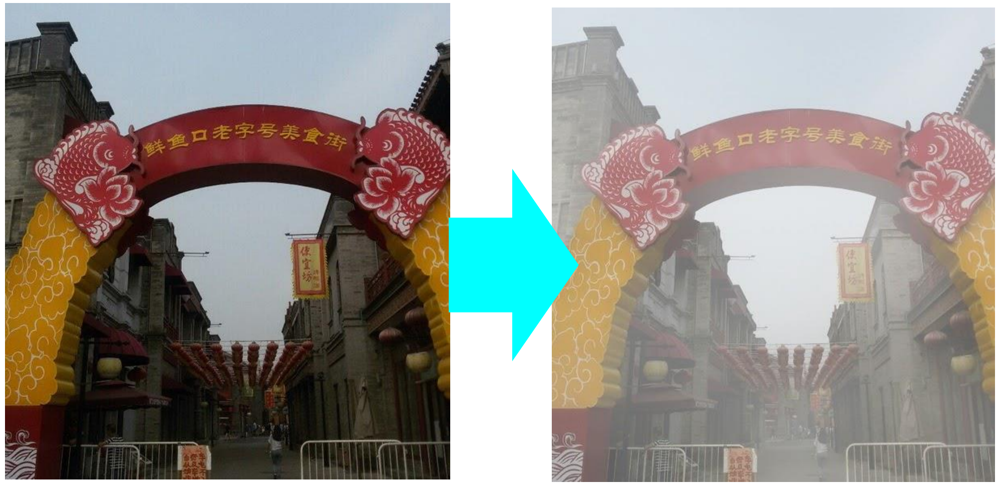
Estimation By Mathematical Modeling

- Study the actual physical process that results in the degradation!
- Can be used to estimate environmental effects to image capture
- How did this image happen??

Model the Degradation

- Study the Degradation Model

J $\qquad\qquad$ I



$$I = Jt + A(1 - t)$$

J: original intensity
A: airlight – color of surrounding media
t: transmission

Use the Degradation Model

Image Restoration

I $\quad$ J

****Computer Vision Task**

I $\quad$ $e^{-\beta d}$ or d

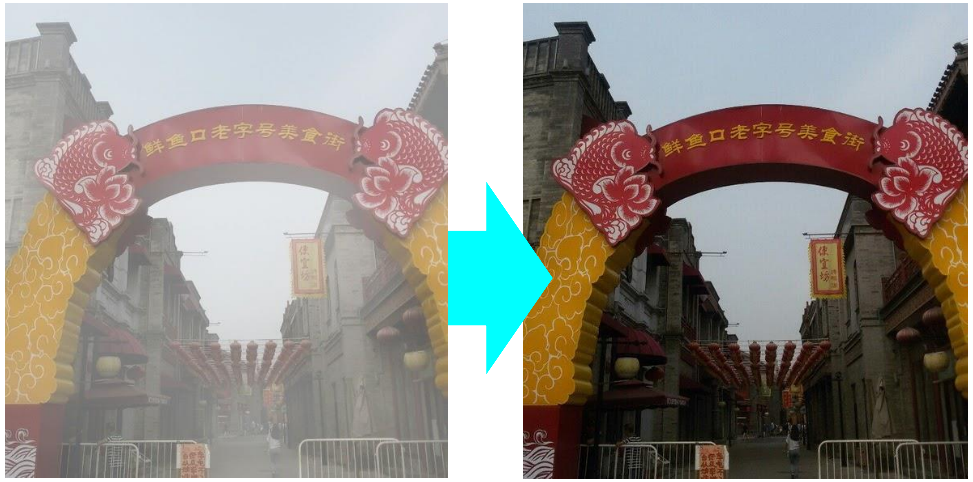
$$I = Jt + A(1 - t)$$

**Depth estimation in this case, depends on the model used

Image Restoration

After the images are restored – conventional CV methods can be used

I J



1. **Pixel intensity-based**
   - Assumptions & Constraints
   - Statistical Priors
   - Estimating Variables
2. **Deep learning based**
   - Directly map degraded to clear image

$$I = Jt + A(1 - t)$$

Statistical Priors – To Solve Inverse Models

- The image degradation model is usually very underdetermined if only the degraded image is known. Therefore, it is necessary to use additional constraints such as statistical priors.
- Statistical priors are observations of images of the same nature that are statistically significant so we can assume it is a prior (known) about that nature.

$$I = Jt + A(1 - t)$$

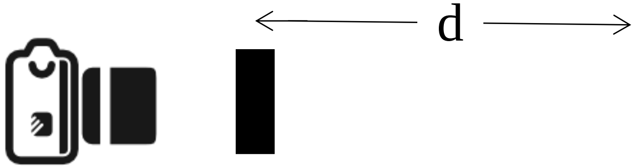
He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." IEEE transactions on pattern analysis and machine intelligence 33.12 (2010): 2341-2353.

Dark channel Prior (DCP) *

- For **clear** non-sky local patches Ω, at least 1 channel {R,G,B} is very small ≈ 0

$$DCP(x) = \min_{s \in \{R,G,B\}} \left(\min_{y \epsilon \Omega(x)} (J^s(y)) \right) \approx 0$$

- For hazy/foggy images



$DCP \approx 0$ $DCP \uparrow$

- Transmission based on DCP

$$\tau_{DCP}(x) = 1 - \min_{s \in \{R,G,B\}} \left(\min_{y \epsilon \Omega(x)} \left(\frac{I^s}{A^s} \right) \right)$$

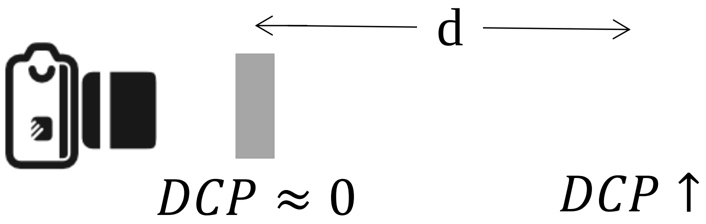
I: Observed Intensity
A: Airlight

$$I = Jt + A(1 - t)$$

Airlight Estimation

Airlight can be found at maximum distance

- Statistical Priors
  - Maximum distance is indicated by maximum DCP
- This is not foolproof as ambiguities exist *



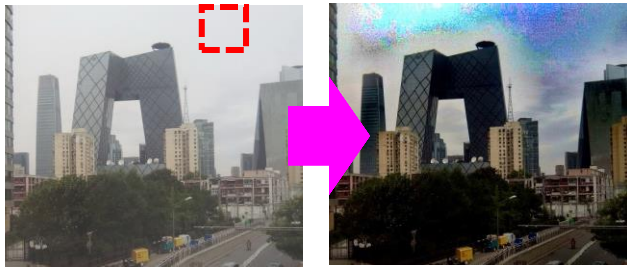
$$I = Jt + A(1 - t)$$

Image Restoration using Statistical Priors

Aziz Fikri Hudaya, Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2021)

- Studied various statistical priors for image restoration

Dark Channel Prior[1]

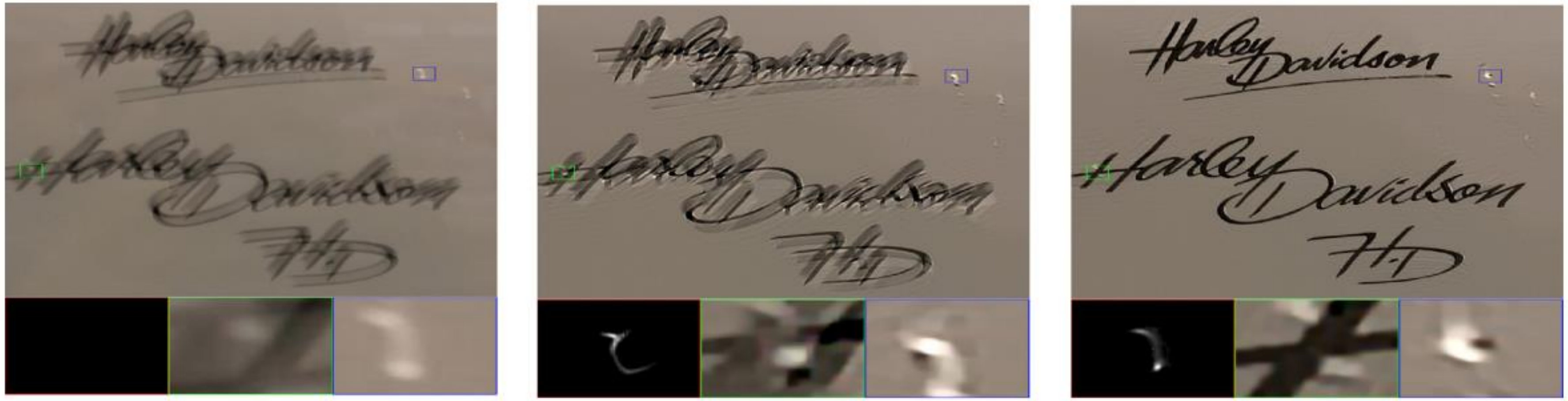


Two Peak Channel Prior[2]

- Proposed framework for optimized hazy image restoration: Separate sky and object

Blind Image Deblurring Based on the Sparsity of Patch Minimum Information[1]

- Using Dark Channel Prior[2], patch minimum information[1], and edge regularizer[3]
- Deblurring in two steps:
  - Estimating the kernel by approximation
  - Restoration



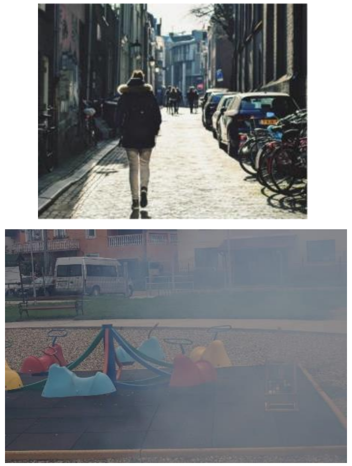
(a) Blurred (k) Pan-17 [35] (l) Ours

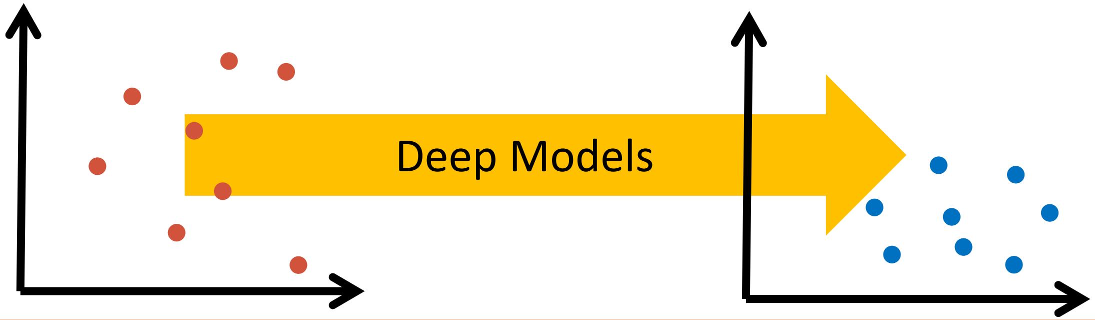
1)Hsieh, Po-Wen, and Pei-Chiang Shao. "Blind image deblurring based on the sparsity of patch minimum information." *Pattern Recognition* 109 (2021).
2)He, Kaiming, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior." IEEE PAMI. (2010).
3)Hsieh, Po-Wen, Pei-Chiang Shao, and Suh-Yuh Yang. "A regularization model with adaptive diffusivity for variational image denoising." Signal Processing 149 (2018).

Image-to-Image Translation for CV tasks

- By establishing the relationship between 2 image sets, we can transform/ translate images from 1 set to the other as a pixel regression problem
- Deep learning networks are very powerful to establish image to image translation

Input

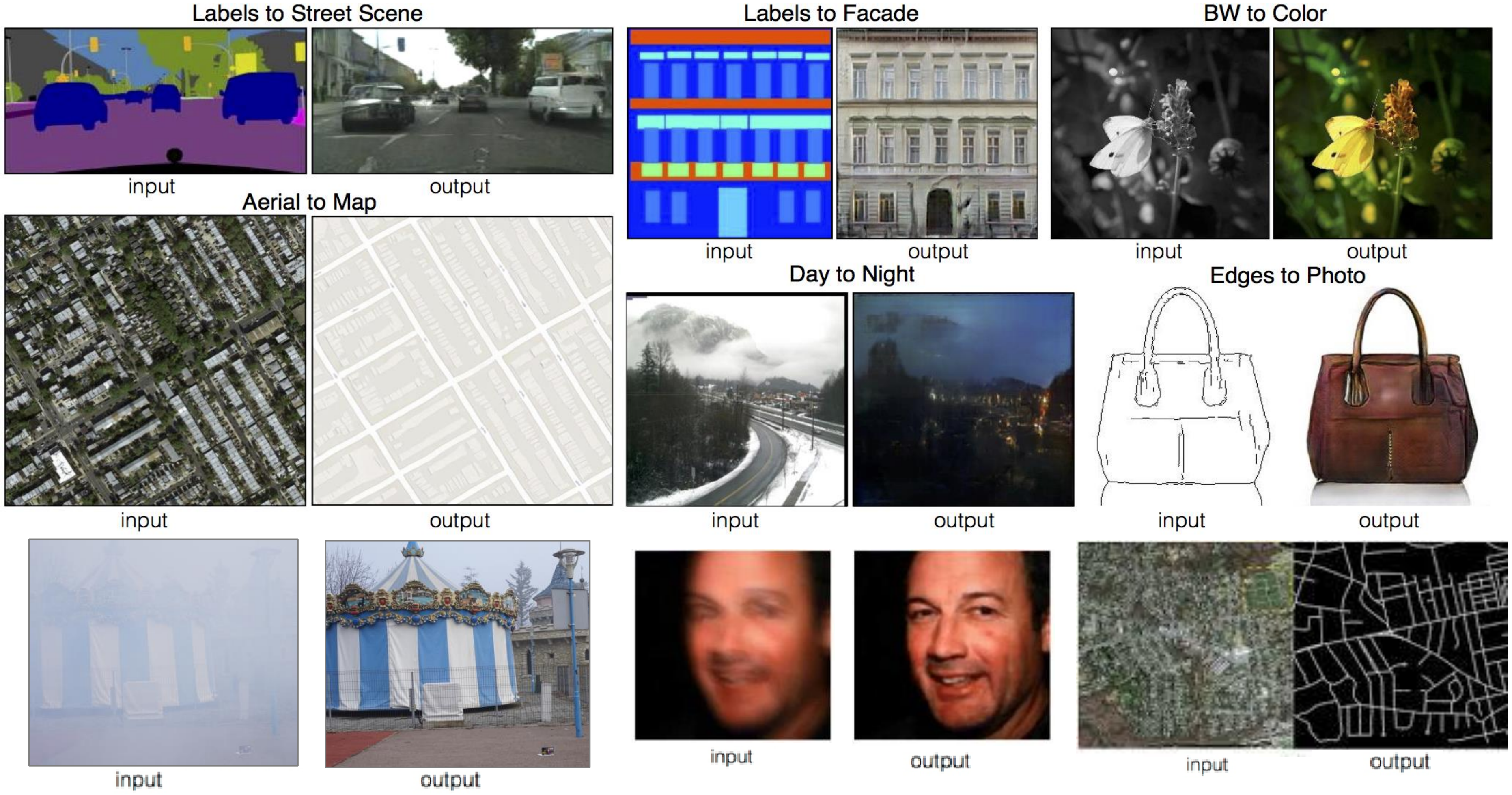




Output

- Image Segmentation
- Image Restoration
- Pixelwise Information Extraction

Image-to-Image Translation Cases



[1] Li, et.al (2019); [2] Liu, et al. (2019); [3]Qin, et al. (2020).

Mod PDR-Net

Cahyo Adhi Hartanto, Tesis, Magister Ilmu Komputer Universitas Indonesia (2021).
Hartanto, Rahadianti. "Single Image Dehazing Using Deep Learning." JOIV: International Journal on Informatics Visualization 5.1 (2021).

- Mod PDR-Net – a modification to the original PDR-Net[1]

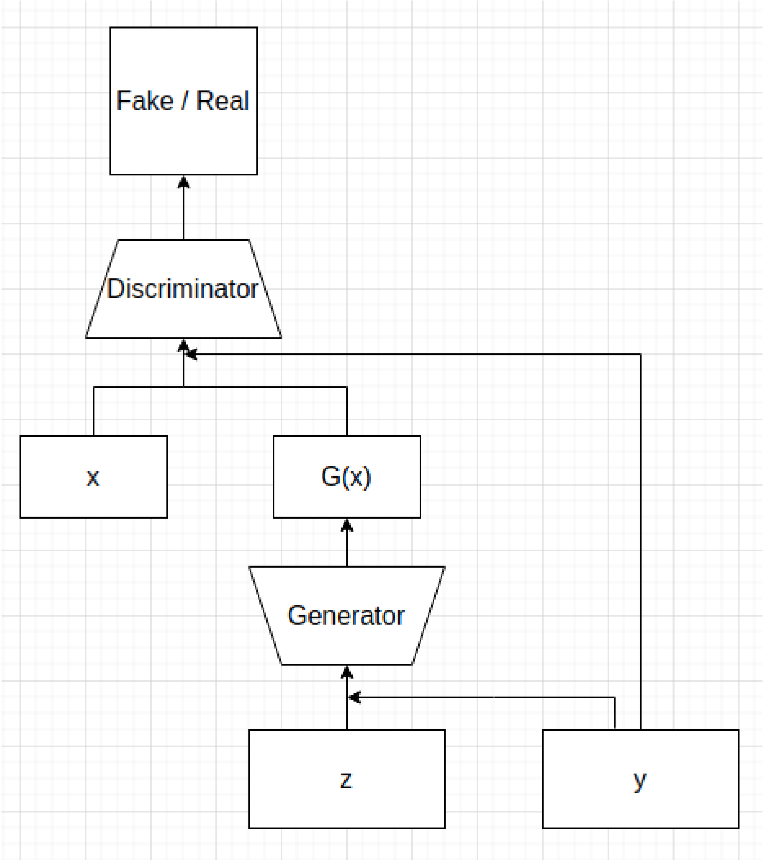
- Better dehazing results on test set **and** other datasets - more robust!

Hazy Image | PDR-Net | Proposed Mod PDR-Net | Ground Truth

[1]Hartanto, Rahadianti. "Single Image Dehazing Using Deep Learning." JOIV: International Journal on Informatics Visualization 5.1 (2021).
[2]Hartanto, **Rahadianti**, Rachmadi (*under review)*

CGAN with Module Based PDR-Net

Andrew Theodore T., Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2022).



- Generator yang digunakan adalah Module Based PDR-Net[1,2]
- Discriminator yang digunakan PatchGan
- Loss function yang digunakan: Binary Cross Entropy Loss, L1 Loss, Perceptual Loss, dan SSIM Loss.

Metrik	Mod PDR Net	CGAN Mod PDR Net (Usulan)
SSIM ↑	0.772 ± 0.103	0.785 ± 0.094
BRISQUE ↓	31.436 ± 8.967	28.375 ± 7.691


Hazy


Ground Truth

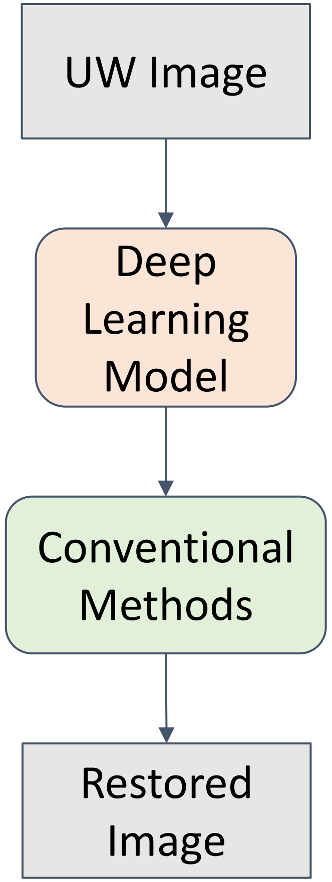

Mod PDR-Net


CGAN with Mod PDR-Net

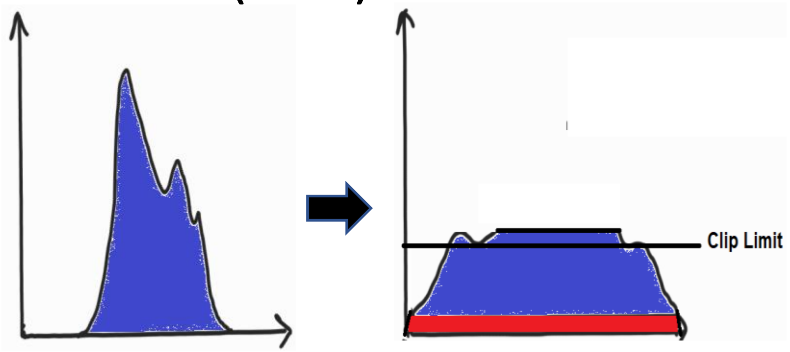
[1]Fu, X., & Cao, X. (2020)

Underwater Image Restoration

Jahroo N. V, Tesis, Magister Ilmu Komputer Universitas Indonesia (2022).
Marvi, Jahroo Nabila, and Laksmita Rahadianti. "Towards Robust Underwater Image Enhancement." International Conference on Soft Computing in Data Science. Singapore: Springer Nature Singapore, 2023.



- High variance of the underwater appearance, difficult to get ground truth   banyak variance kalau underwater
- Conventional methods are not enough – deep learning is prone to overfitting
- Hybrid method: GL-Net[1]: takes the generic restoration of DL model, and further corrects the result with conventional Compressed Histogram Equalization (CHE)



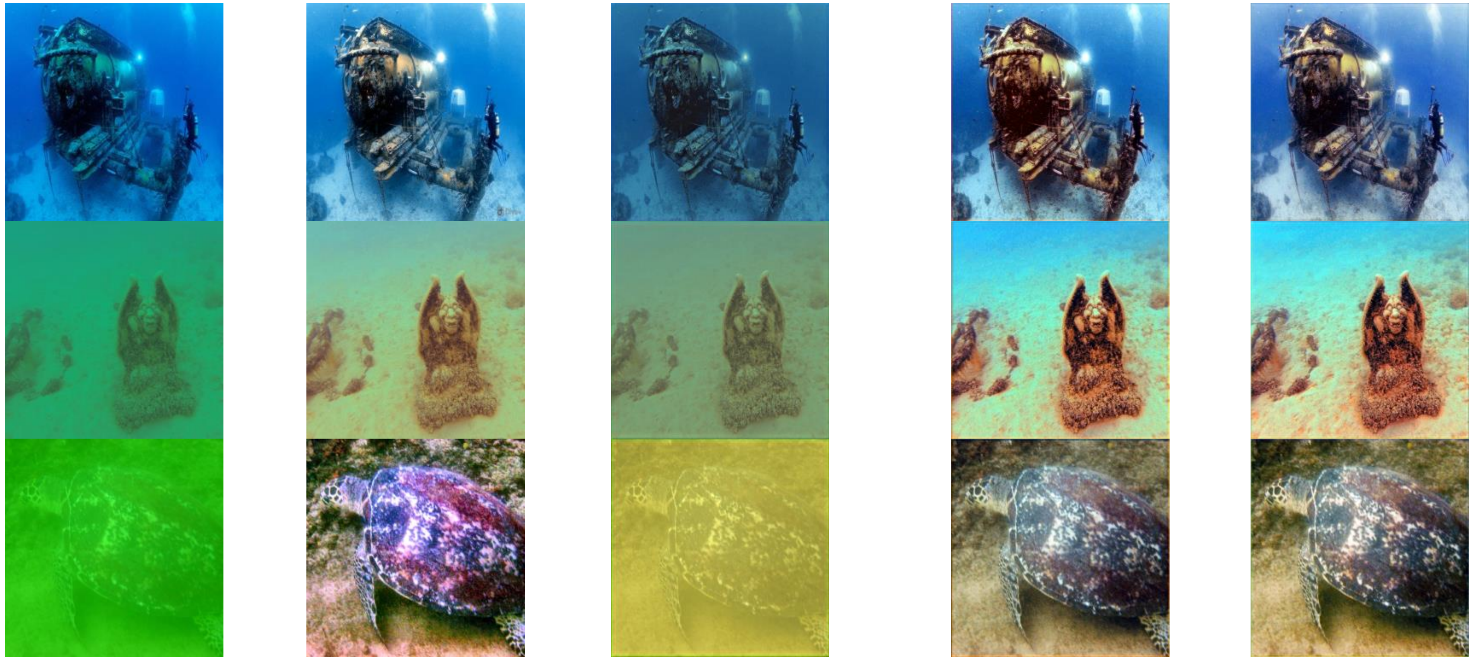
Underwater Image Restoration (2)

Jahroo N. V, Tesis, Magister Ilmu Komputer Universitas Indonesia (2022).
Marvi, Jahroo Nabila, and Laksmita Rahadianti. "Towards Robust Underwater Image Enhancement." International Conference on Soft Computing in Data Science. Singapore: Springer Nature Singapore, 2023.

UW Image → **Deep Learning Model** → Conventional Methods → Restored Image

- We propose a modified deep learning architecture: Mod GL-Net

Input | Referensi | UWCNN | GL-Net | Mod GL-Net



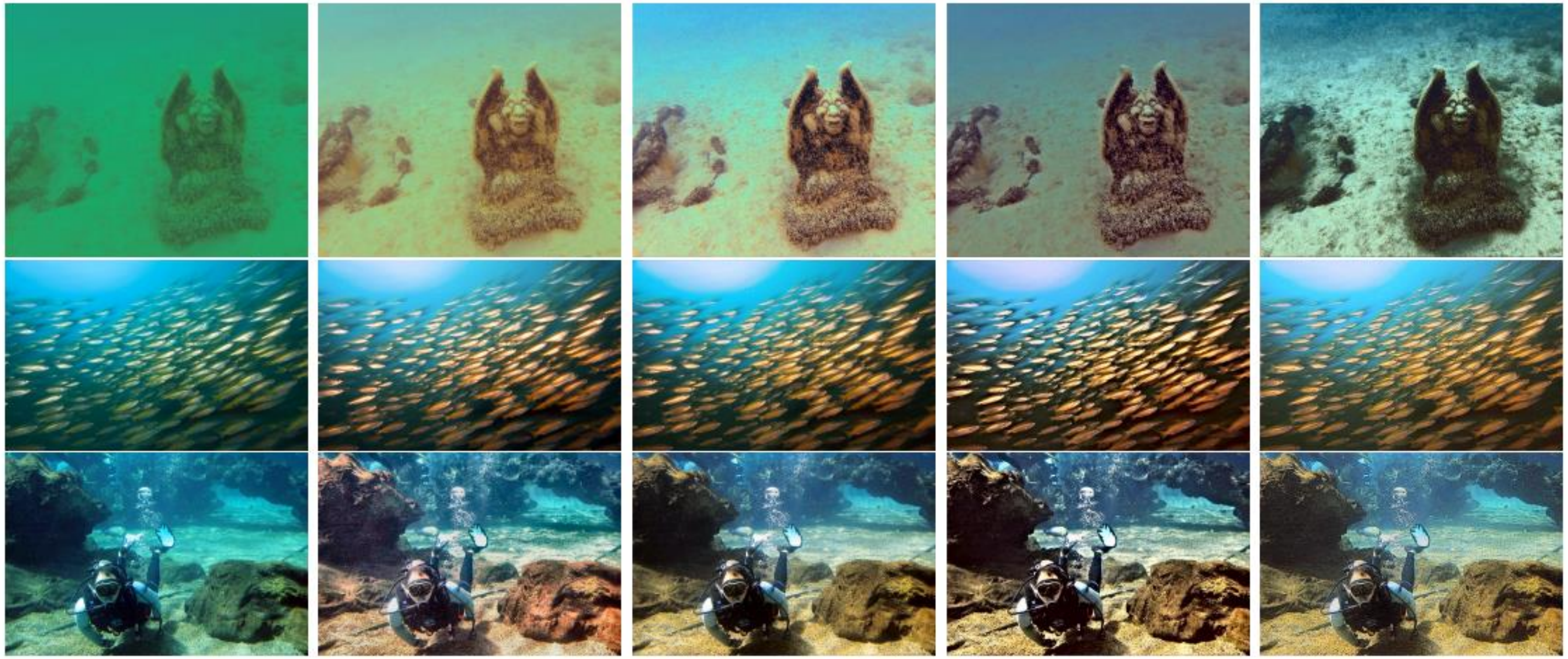
[1]Fu, X., & Cao, X. (2020)

Underwater Image Restoration with GL-Net

Virdian H. P., Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2024)

UW Image → Deep Learning Model → **Conventional Methods** → Restored Image

- This time we use GLNet[1] and tried alternate conventional methods:

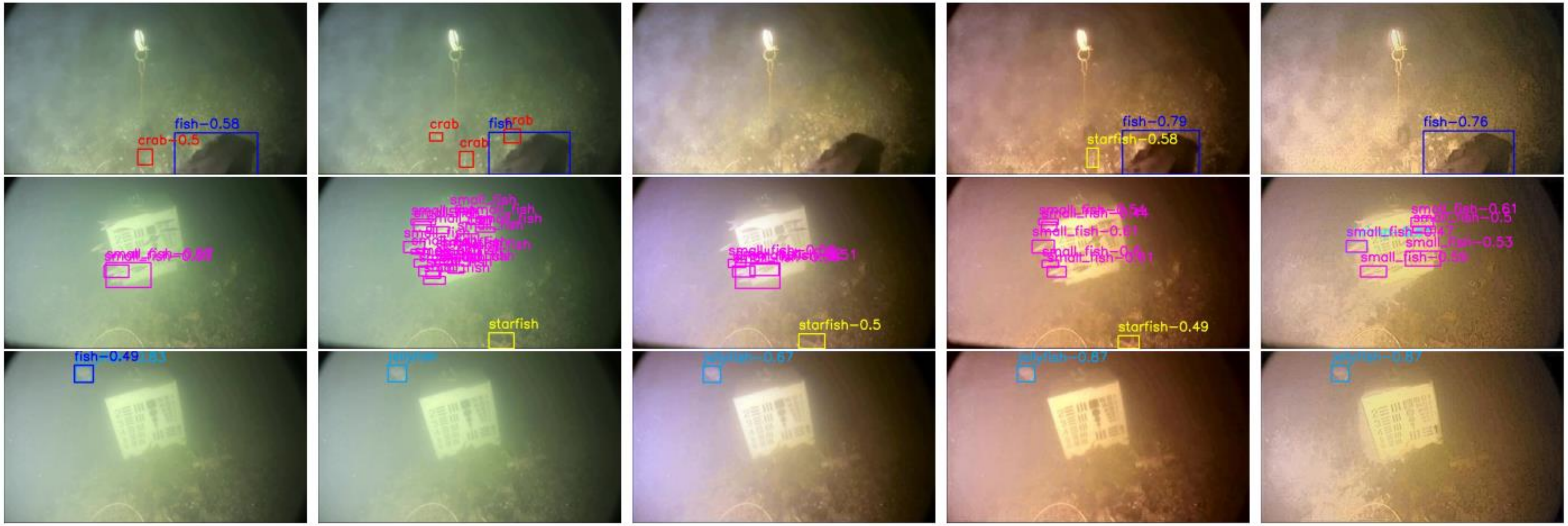


(a) Original (b) *Ground Truth* (c) CHE (d) LACE (e) Bayesian Retinex

Underwater Image Restoration with GL-Net (2)

Virdian H. P., Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2024)

- Does restored images work better for object detection?
- Case study: UW SAR, autonomous UW vehicles



(a) Original (b) *Ground Truth* (c) CHE (d) LACE (e) Bayesian Retinex

Deblurring for Non-Uniform Blur

Made P. N., Tesis, Magister Ilmu Komputer Universitas Indonesia (2024)

- Image deblurring with deep learning

Deblurring with DL

- Models treat images as a whole – what happens when blur is not uniform?

- DL models for restoration
  - Restormer[1] – for most restoration, fast and reliable, not precise for non-uniform
  - BA-Net[2] – with attention mechanism, but very very slow….
- Can we find a balance between the two?

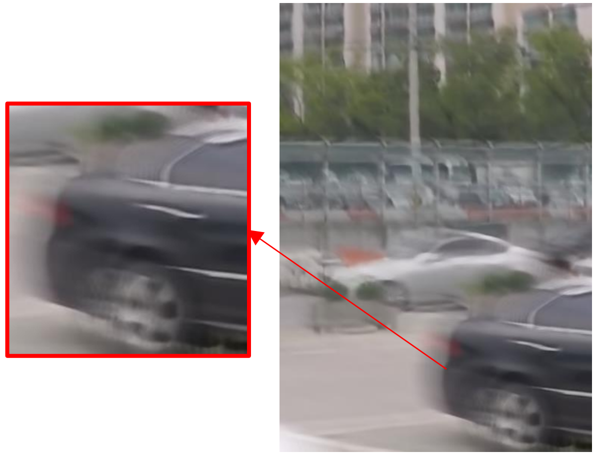
Deblurring for Non-Uniform Blur (2)

Made P. N., Tesis, Magister Ilmu Komputer Universitas Indonesia (2024)

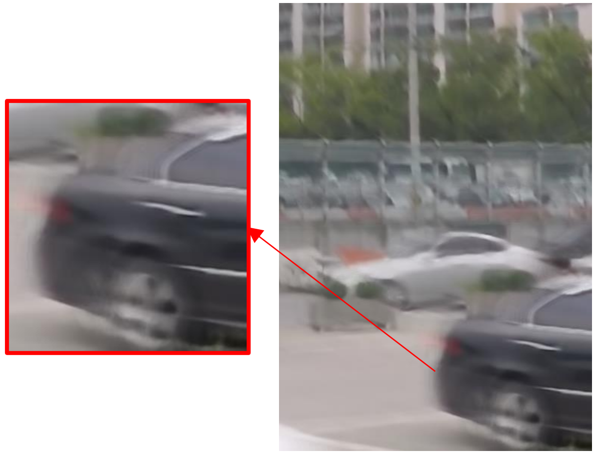
- Restormer-BA: adds blur attention to restormer model

Output Restormer

Output RestormerBA



Output Restormer



Output RestormerBA

Guo, Chunle, et al. "Zero-reference deep curve estimation for low-light image enhancement." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.

Image Restoration for Low-Light Images

- Can we use dark images for computer vision?

- DCE-Net[1] for dark image restoration uses multiple losses:
  - Color constancy
  - Spatial constancy
  - Exposure control
  - Illumination smoothness

$$\boxed{\text{Total Loss} = 5L_{col} + 1L_{spa} + 10\,L_{exp} + 100L_{tvA}}$$

Image Restoration for Low-Light Images (2)

Mardianto, Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2023)

- Modifing the color constancy using the minkowski norm and tuning weights w_n

Citra gelap

Citra *ground truth*

Hasil peningkatan kualitas dengan L2 *color constancy loss* dan *bilateral filter*

Hasil peningkatan kualitas dengan L4 *color constancy loss* dan *bilateral filter*

Hasil peningkatan kualitas dengan L6 *color constancy loss* dan *bilateral filter*

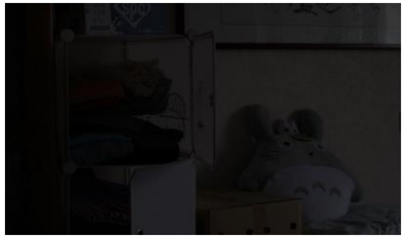
$$\text{Total Loss} = 5L_{col} + 1L_{spa} + 10\,L_{exp} + 100L_{tvA}$$

$$\text{Total Loss}^{**} = 5L'_{col} + 1L_{spa} + 10\,L_{exp} + 100L_{tvA}$$

Image Restoration for Dark Images

Adam S. M., Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2024)

- DCE-Net[1] for dark image restoration uses 4 aspects of loss
- Meanwhile, structure[2] (SSIM) is an important factor not yet accounted for



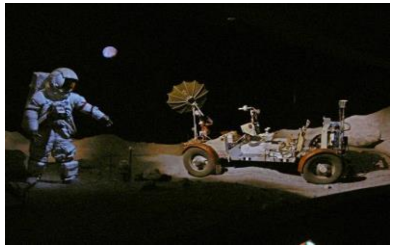
Sumber

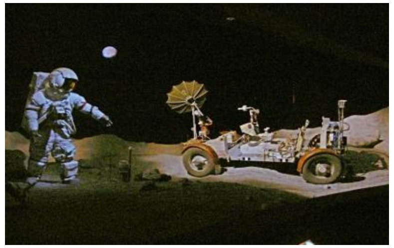
Ground Truth

Zero-DCE

Zero-DCE + Structure

N/A





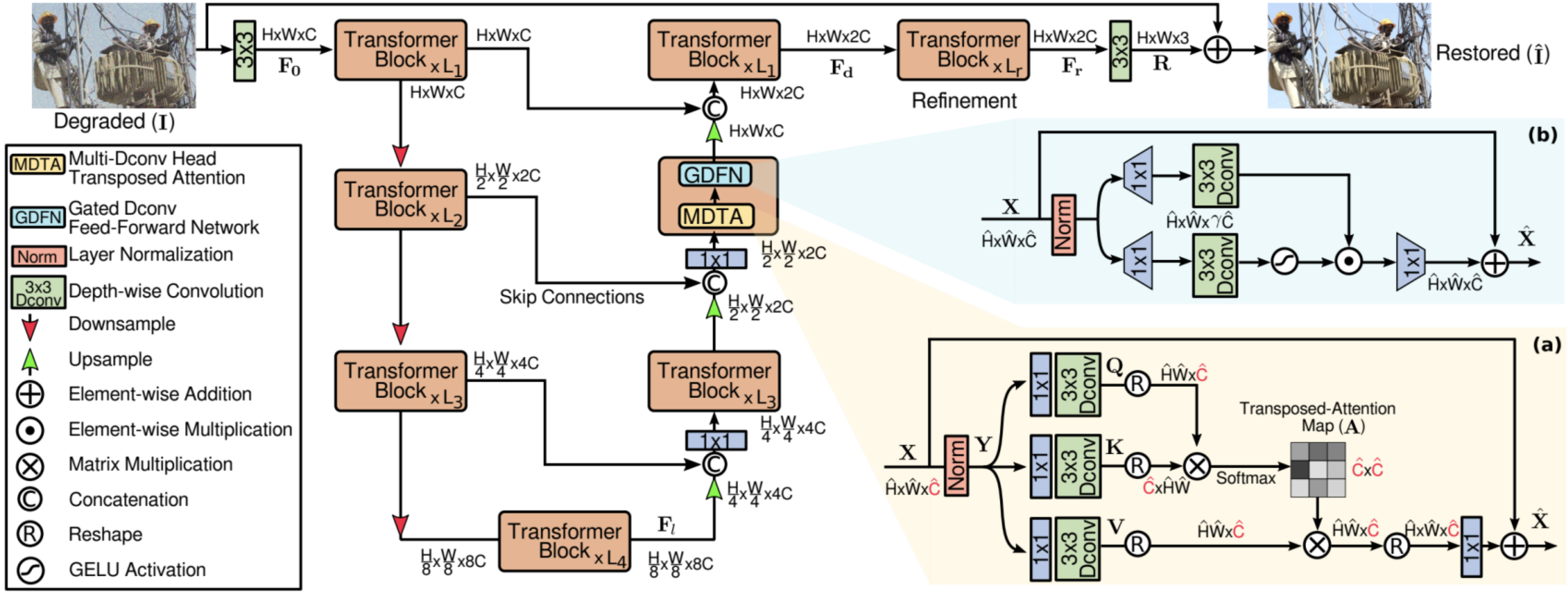
$$\text{Total Loss} = 5L_{col} + 1L_{spa} + 10\,L_{exp} + 100L_{tvA}$$

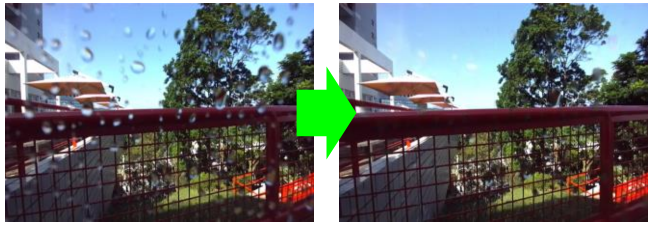
$$\text{Total Loss**} = 5\,L_{col} + 1\,L_{spa} + 10\,L_{exp} + 200\,L_{tvA\,+} + w_s\,L_{SSIM}$$

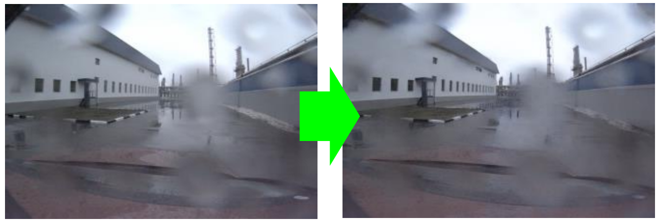
Deraining Images with Restormer

Stephen H. C , Skripsi, Sarjana Ilmu Komputer Universitas Indonesia (2022).

- Deraining, Motion Deblurring, Defocus Deblurring, and Denoising







Deblurring by Realistic Blurring[1]

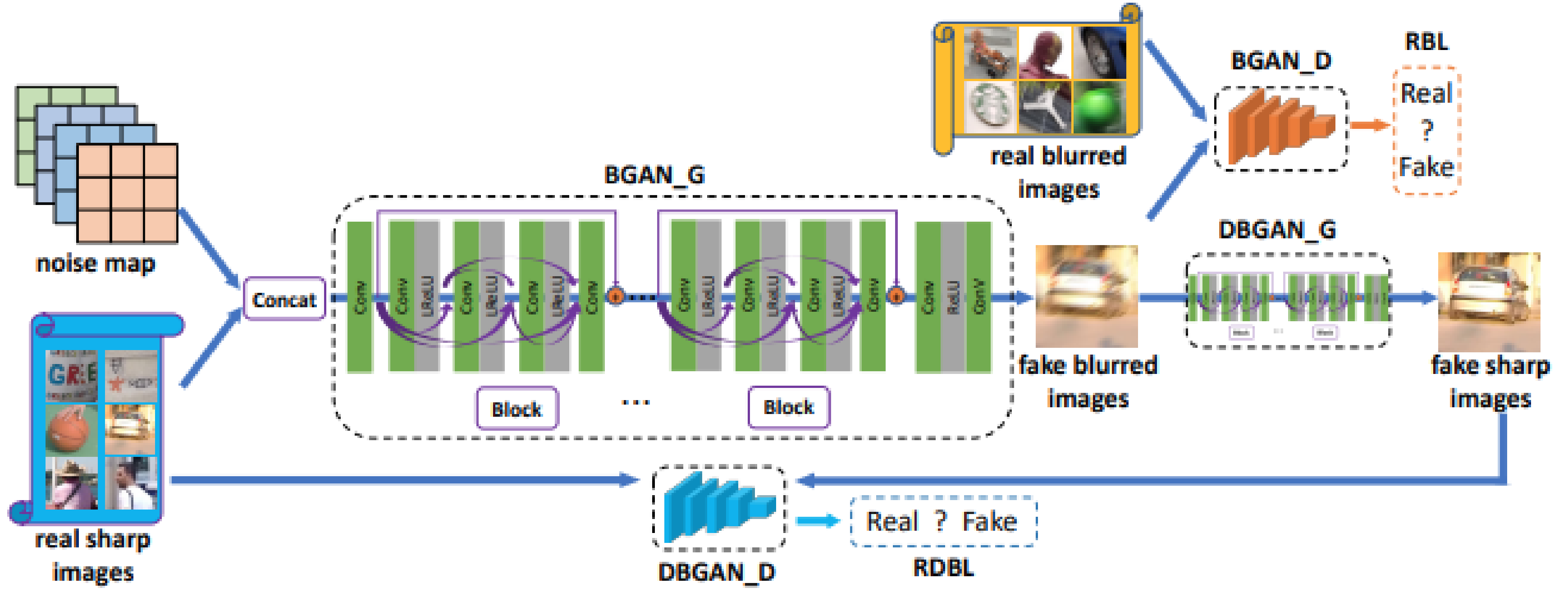


Figure 2. **The proposed framework and training process.** This framework contains two main modules, a BGAN and a DBGAN. D and G denote discriminator and generator networks, respectively. The BGAN takes sharp images as input and outputs realistic blurry images, which are then fed into the DBGAN in order to learn to deblur. During the inference stage, only the DBGAN is applied.

Blur Image | Proposed

[1]Zhang, Kaihao, et al. "Deblurring by realistic blurring." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

- Pedersen, et al. (2012). Full-Reference Image Quality Metrics: Classification and Evaluation. Foundations and Trends in Computer Graphics and Vision.
- Ahmed, et al. "A Survey of Recent Approaches on No-Reference Image Quality Assessment with Multiscale Geometric Analysis Transforms.

Image Quality Assessment

Evaluation of Image Restoration

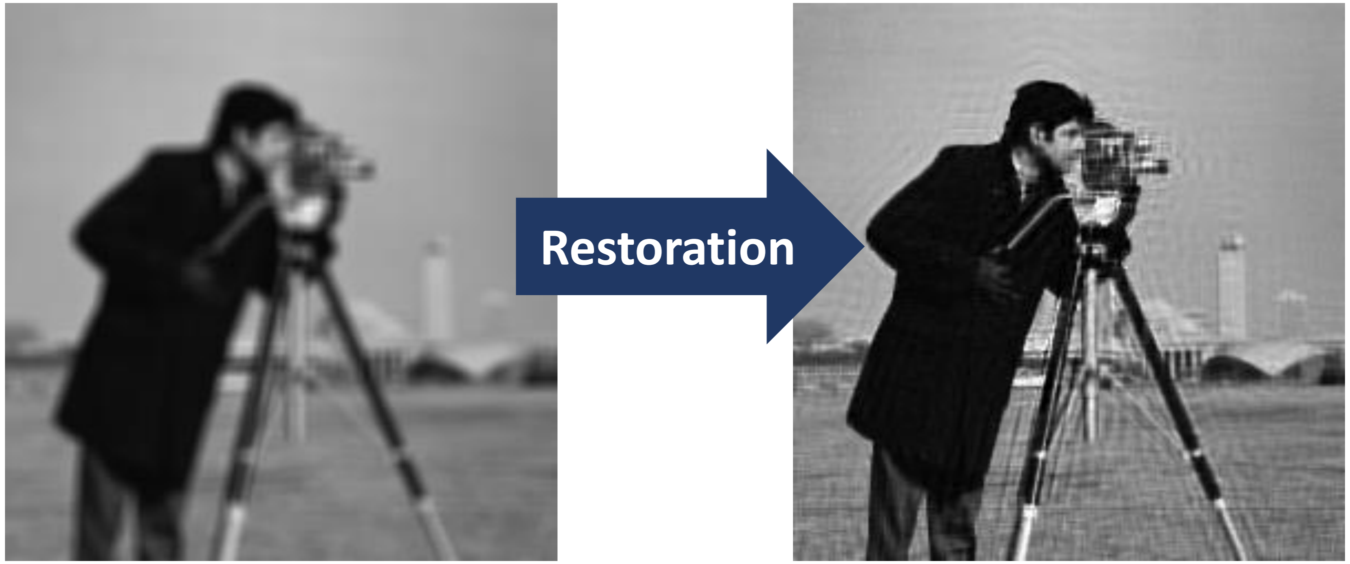
How **good** is this image?

Restoration

Compare with original – **if available!**

Evaluation of Image Restoration

How **good** is this image?



No original image available -
how can we measure quality?

Image Quality

- The weighted combination of all the visually significant attributes of an image – visual pleasingness
- The level of accuracy in which digital systems display the image.

Gu, Ke, et al. "Subjective and objective quality assessment for images with contrast change." *2013 IEEE International Conference on Image Processing*. IEEE, 2013.

Image Quality Assessment

ini yang paling reliable karena orang adalah end user

Subjective Image Quality Assessment

- The most reliable method for assessing the quality of images
- Uses human observers, since human observers are users in multimedia applications.
- Can describe **visual pleasingness**

Objective Image Quality Assessment

- The goal of objective IQA is to design mathematical models that are able to predict the quality of an image accurately and automatically
- Usually expressed **numerically**
- Aimed to mimic the quality predictions of an average human observer.

Objective Image Quality Metrics

Full-reference (FR) methods

The image quality is computed by comparing it with a reference image that is assumed to have perfect quality

Reduced-reference (RR) methods

The image quality is measured based on a features extracted from the reference image.

No-reference (NR) methods

The image quality is measured without any reference to the original one

Pedersen, Marius, and Jon Yngve Hardeberg. "Survey of full-reference image quality metrics." (2009).
Pedersen, Marius, and Jon Yngve Hardeberg. "Full-reference image quality metrics: Classification and evaluation." *Foundations and Trends® in Computer Graphics and Vision* 7.1 (2012): 1-80.

Full Reference Image Quality Metrics

- Mean Square Error (MSE) Based

$$MSE = \frac{\sum_{M,N}(I(m,n) - J(m,n))^2}{M.N}$$

Root Mean Square Error (RMSE) $RMSE = \sqrt{MSE}$

Signal to Noise Ratio $SNR = 10 \cdot log10\left(\frac{\sum_{y=0}^{M}\sum_{x=0}^{M} S(x,y)^2}{MN \cdot MSE}\right)$

Peak Signal to Noise Ratio (PSNR) $PSNR = 10\log_{10}\frac{R^2}{MSE}$

Derivations of error metrics (so many!) WMSE, CSF weighted MSE, DECOR-WSNR, etc

1) Wang, Zhou, and Alan C. Bovik. "A universal image quality index." *IEEE signal processing letters* 9.3 (2002): 81-84.
2) Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13.4 (2004): 600-612.

Full Reference Image Quality Metrics

- Structure Based

ada berapa perbedaan, ada berapa persamaan, nah itu yang dihitung

Universal Quality Index (UIQ)[1]

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\,\bar{x}\,\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\,\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} .$$

Correlation coefficient ,difference of luminance, and similarity of contrast between x and y.

strukturnya seberapa mirip

Structural Similarity (SSIM) [2]

$$\text{SSIM}(x,y) = \left[l(x,y)^{\alpha} \cdot c(x,y)^{\beta} \cdot s(x,y)^{\gamma}\right]$$

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \qquad c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \qquad s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

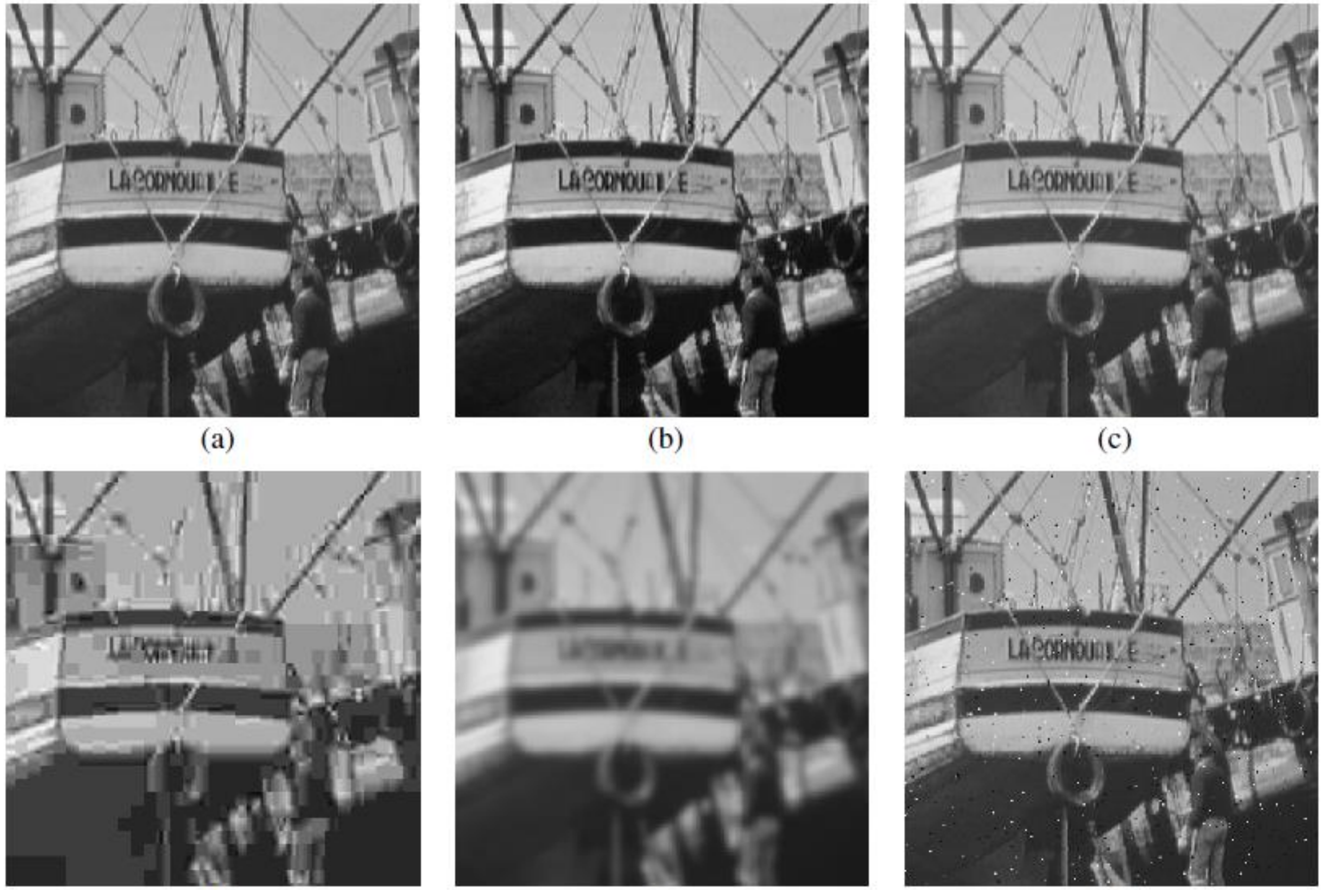
Various derivations of SSIM

Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13.4 (2004): 600-612.

Structural Similarity (SSIM)

dalam konteks gambar
kalau kita lihat berdasarkan error, semua gambar tersebut punya error yang sama

SSIM berusaha menangkap struktur, sehingga shape-shape (luminence) gambar sebelumnya ada, akan terlihat.

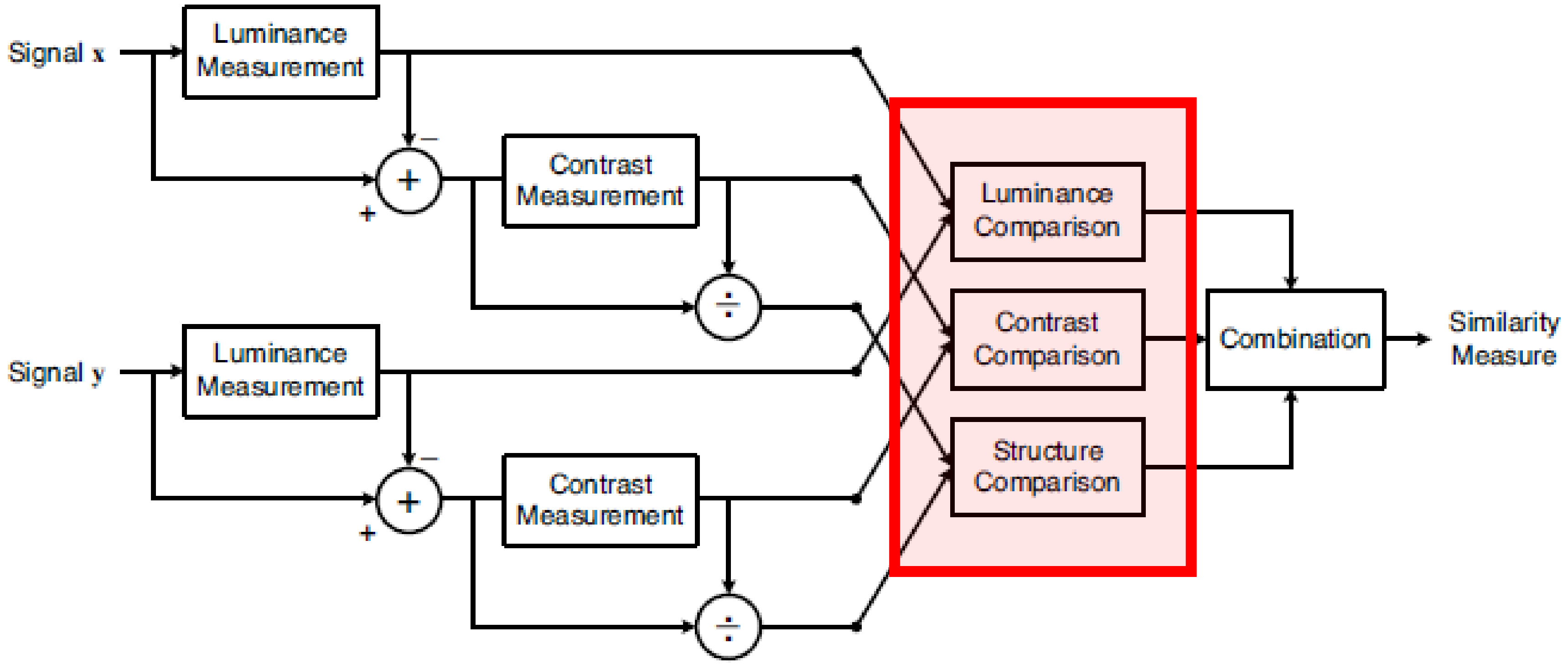


(a) (b) (c) (d) (e) (f)

- Comparison of images with different types of distortions, all with MSE = 210.
- But the images have drastically different perceptual quality.
  a) Original image
  b) Contrast stretched
  c) Mean-shifted
  d) JPEG compressed
  e) Blurred
  f) Salt-pepper noise

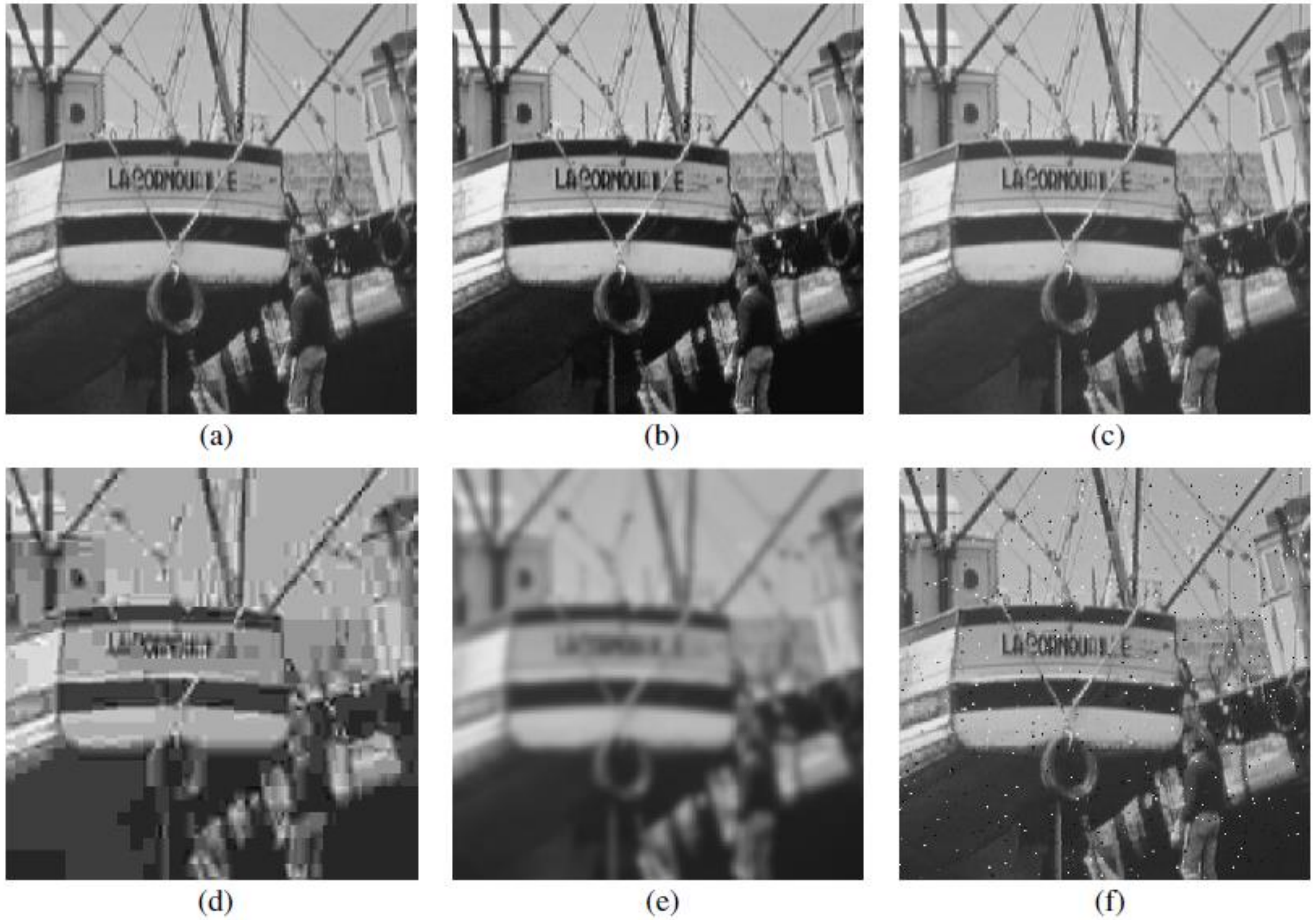
Structural Similarity (SSIM) (2)

- Image degradations as perceived changes in structural information variation.
- SSIM quality measure is constructed from the perspective of image formation.



Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13.4 (2004): 600-612.

Structural Similarity (SSIM) (3)



- Comparison of images with different types of distortions, all with MSE = 210.
- SSIM can show the perceptual quality

a) Original image
b) Contrast stretched, **MSSIM = 0.9168**
c) Mean-shifted, **MSSIM = 0.99**
d) JPEG compressed, **MSSIM = 0.6949**
e) Blurred, **MSSIM = 0.7052**
f) Salt-pepper noise, **MSSIM = 0.7748**

Reduced Reference Image Quality Metrics

- Estimating Known FR Metrics
- Edge Based
- Image Statistics Based
- And many more…

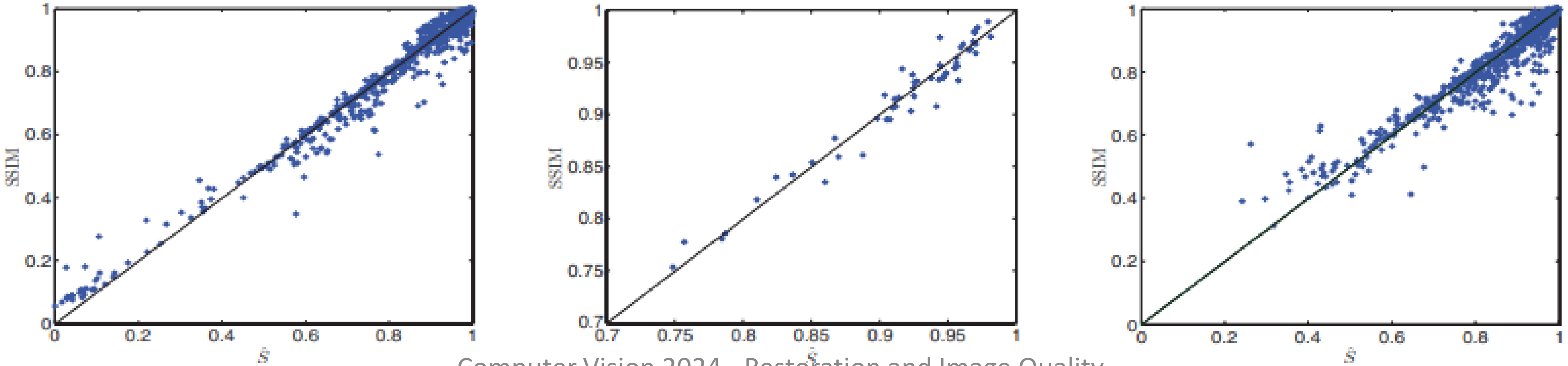
Structural Similarity Estimation

- Proposed distortion measure D_n

$$D_n = g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d) \log\left(1 + \frac{1}{D_0} \sum_{k=1}^{K} \left|\hat{d}^k(p^k || q^k)\right|\right). \qquad g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d) = \frac{\|\boldsymbol{\sigma}_r\|^2 + \|\boldsymbol{\sigma}_d\|^2 + C}{2(\boldsymbol{\sigma}_r \cdot \boldsymbol{\sigma}_d) + C}$$

$$\hat{S} = 1 - \alpha D_n$$

- The distortion model has a near perfect linear relationship with SSIM
- Estimated SSIM $\hat{S}$

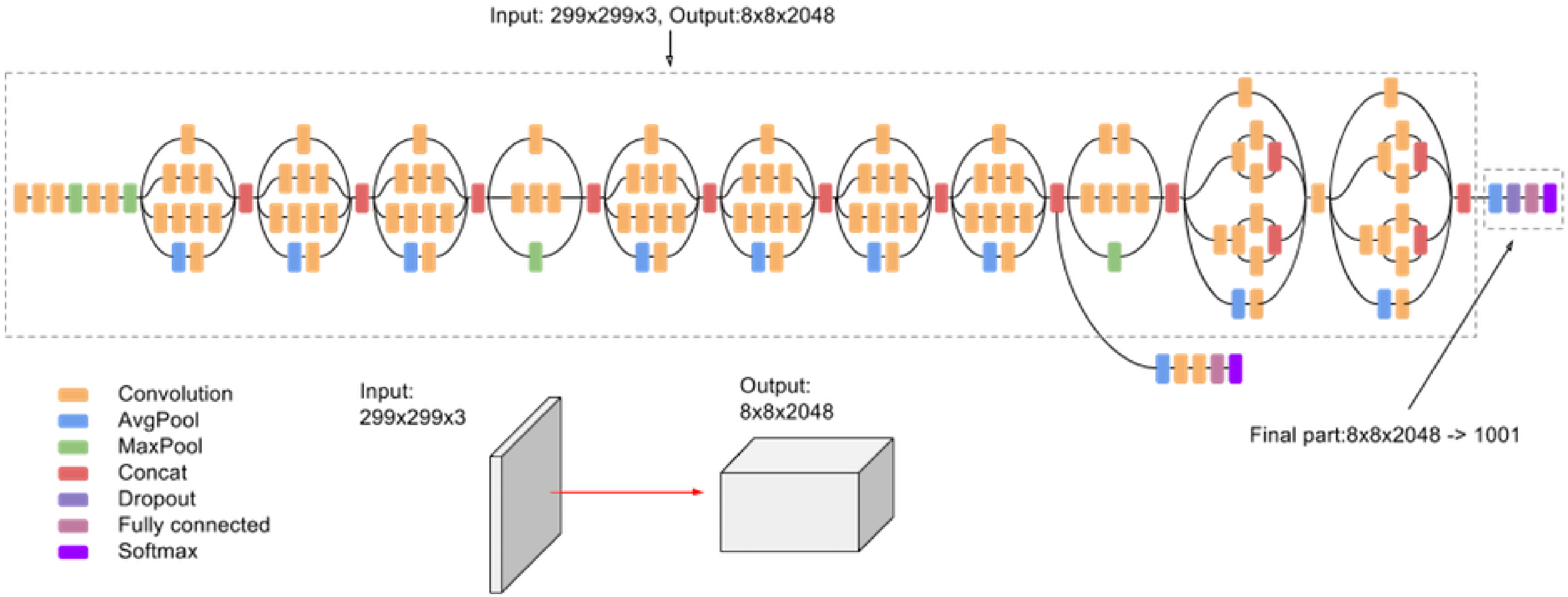


[1]Rehman, Abdul, and Zhou Wang. "Reduced-reference image quality assessment by structural similarity estimation." *IEEE transactions on image processing* 21.8 (2012).

1) Heusel, Martin, et al. "Gans trained by a two time-scale update rule converge to a local nash equilibrium." *Advances in neural information processing systems* 30 (2017).
2) Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

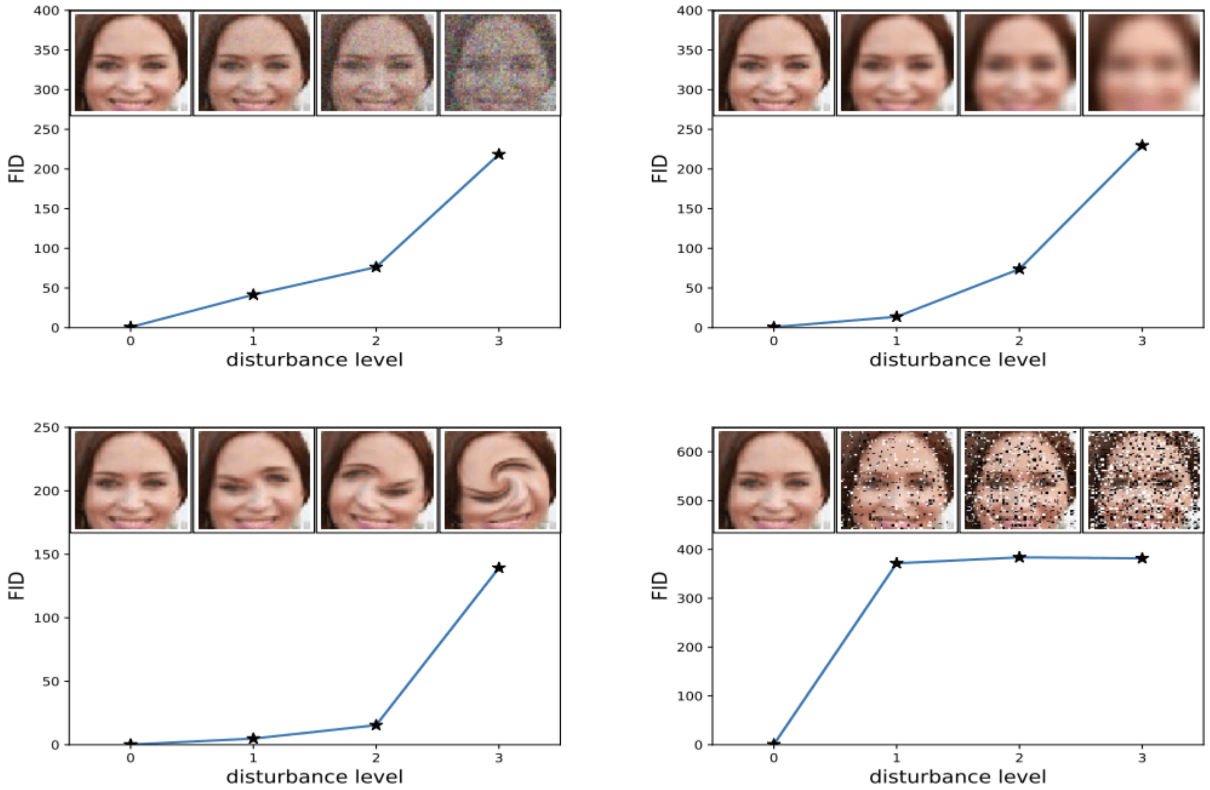
Fréchet Inception Distance (FID)[1]

- A metric for evaluating the quality of generated images and specifically developed to evaluate the performance of generative adversarial networks (GAN)
- Uses a pre-trained Inception V3[2] model (for image recognition)



Fréchet Inception Distance (FID)[1] (2)

- Takes the output of the last layer of Inception V3[2]
- FID score uses the last coding layer (the last pooling layer prior to classification) of Inception V3[2]
- These activations are calculated for a collection of real and generated images.



1) Heusel, Martin, et al. "Gans trained by a two time-scale update rule converge to a local nash equilibrium." *Advances in neural information processing systems* 30 (2017).
2) Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

Fréchet Inception Distance (FID)[1] (3)

- A lower FID score indicates better quality of reconstruction
- How low though?
- Not enough benchmarking
- No consensus as to what value is **good**

susah benchmarking bagus atau engga

Paper	IS	FID
Real CIFAR-10 data (Salimans et al., 2016)	11.24	–
Unsupervised representation learning with deep convolutional generative adversarial networks (DCGAN) (Radford, Metz, & Chintala, 2015)	6.16[4]	37.1[5]
Conditional image generation with PixelCNN decoders (van den Oord et al., 2016)	4.60[6]	65.9[6]
Adversarially learned inference (ALI) (Dumoulin et al., 2016)	5.34[7]	–
Improved techniques for training GANs (Salimans et al., 2016)	6.86	–
Improving generative adversarial networks with denoising feature matching (Warde-Farley & Bengio, 2016)	7.72	–
Learning to generate samples from noise through infusion training (Bordes, Honari, & Vincent, 2017)	4.62	–
BEGAN: Boundary equilibrium generative adversarial networks (Berthelot, Schumm, & Metz, 2017)	5.62	–
MMD GAN: Towards deeper understanding of moment matching network (Li, Chang, Cheng, Yang, & Póczos, 2017)	6.17	–
Improved training of Wasserstein GANs (Gulrajani, Ahmed, Arjovsky, Dumoulin, & Courville, 2017)	7.86	–
Coulomb GANs: Provably optimal Nash equilibrium via potential fields (Unterthiner et al., 2017)	–	27.3

[1] Mittal, A, et al. "No-reference image quality assessment in the spatial domain." *IEEE Transactions on image processing* 21.12 (2012).
[2] Mittal, A, et al. "Making a “completely blind” image quality analyzer." *IEEE Signal processing letters* 20.3 (2012).
[3] Chan, R, et al. "A psychovisually-based image quality evaluator for JPEG images." *Smc 2000 conference proceedings.* IEEE, 2000.

No – Reference Image Quality Metrics

- A Non-exhaustive List

- No reference (image or information) is necessary
- Commonly based on a certain dataset of images deemed of “good quality”

Blind Image Spatial Quality Evaluator (BRISQUE)[1]

Natural Image Quality Evaluator (NIQE)[2]

Perception based Image Quality Evaluator (PIQE)[3]

Natural Image Quality Evaluator (NIQE)

- Constructing a collection of ‘quality aware’ features and fitting them to a multivariate Gaussian (MVG) model.
- The quality aware features are derived from a simple but highly regular natural scene statistic (NSS) model from local image patches that effectively capture the essential low order statistics of natural images.
- *) Large variable: the image dataset to create the natural images (??)
  - Type of images
  - Amount of images

gambar-gambar ini pakai dataset outdoor dataset

berarti gambar-gambarnya gambar alam dengan cahaya yang bagus

Natural image patches selected using a local sharpness measure.

apabila kita mengambil foto bagus, nilainya pasti jelek karena fotonya itu dipanggung sandiwara.

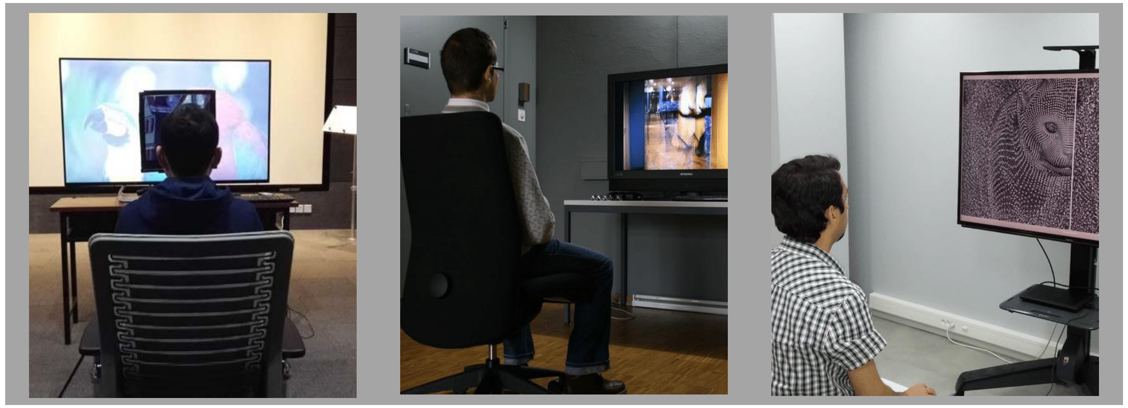
jadi ada masalah yang muncul:
- metric yang dianggap gambar bagus

- Chen, Yuanchun, et al. "Quality assessment for dual-view display system." *2016 Visual Communications and Image Processing (VCIP)*. IEEE, 2016.
- Ahar, Ayyoub, et al. "Subjective quality assessment of numerically reconstructed compressed holograms." *Applications of Digital Image Processing XXXVIII*. Vol. 9599. ISOP, 2015.
- Lutz Goldmann, et al: "A Comprehensive Database and Subjective EvaluationMethodology for Quality of Experience in Stereoscopic Video", Electronic Imaging (EI), 3D Image Processing (3DIP) and Applications, San Jose, USA, 2010.

Subjective Image Quality Metrics

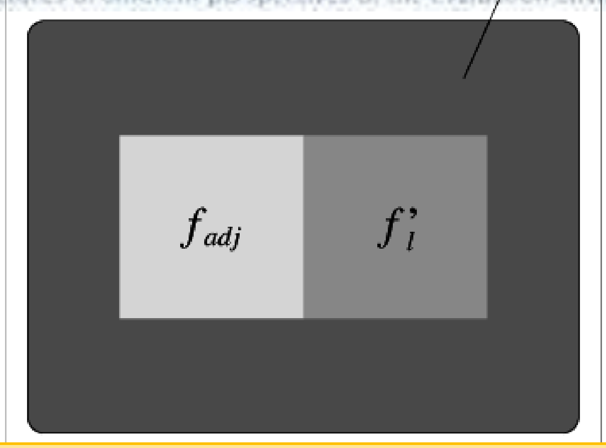
- We can also use human observers to judge the quality through psychovisual experiments.
- Psychovisual experiments attempt to extract a notion of image quality based on feedback from human observers
- This is conducted in a survey format – with a standard setup
- Often measured numerically using *Mean opinion score (MOS)*

$$MOS = \frac{\sum_{n=1}^{N} R_n}{N}$$



Standard Setup for Psychovisual Experiments

Fig.2 Pictures of different perspectives of the evaluation environment.



This is time and labor intensive.

Holistic Image Quality

- How *visually pleasing* is this image?
- Objective metrics are not yet able to describe *visually pleasing* aspect of the human eye

- A.Y. Azizah, L. Rahadianti, H. Deborah. An Introductory Study on Image Quality of Dehazed Images. Proceedings of ICACSIS, 2020.
- L. Rahadianti, A.Y. Azizah, and H. Deborah. Evaluation of the Quality Indicators in Dehazed Images: Color, Contrast, Naturalness, and Visual Pleasingness". Heliyon, 2021.

Image Quality of Dehazed Images

Collaboration with NTNU Norway

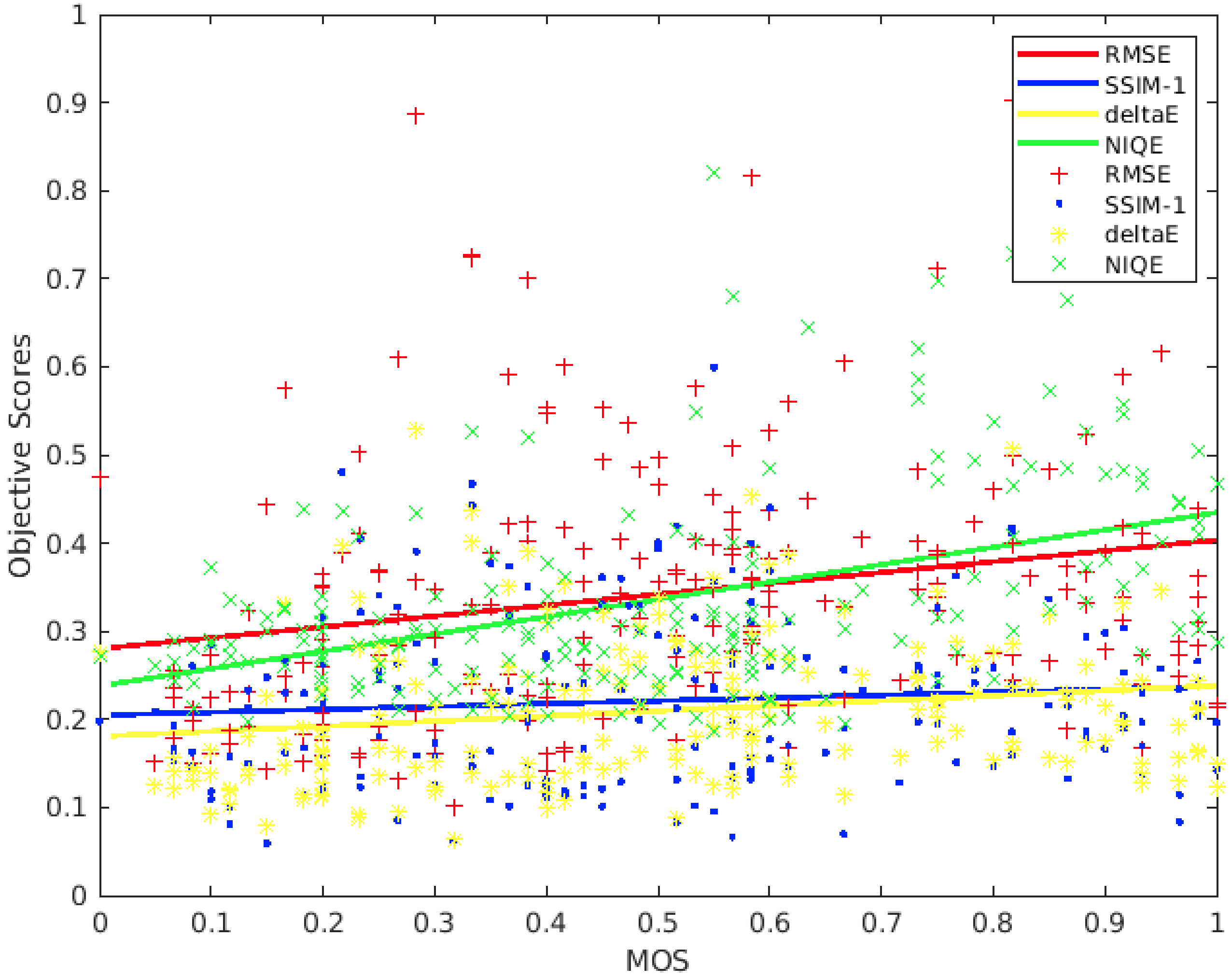
- DL approaches are overfit
- Non-DL approaches are more adaptable
- Low error does not automatically correlate with *visually pleasingness*

Hazy and Clear (GT) | Non-DL Restoration | DL Restoration

Image Quality of Dehazed Images (2)

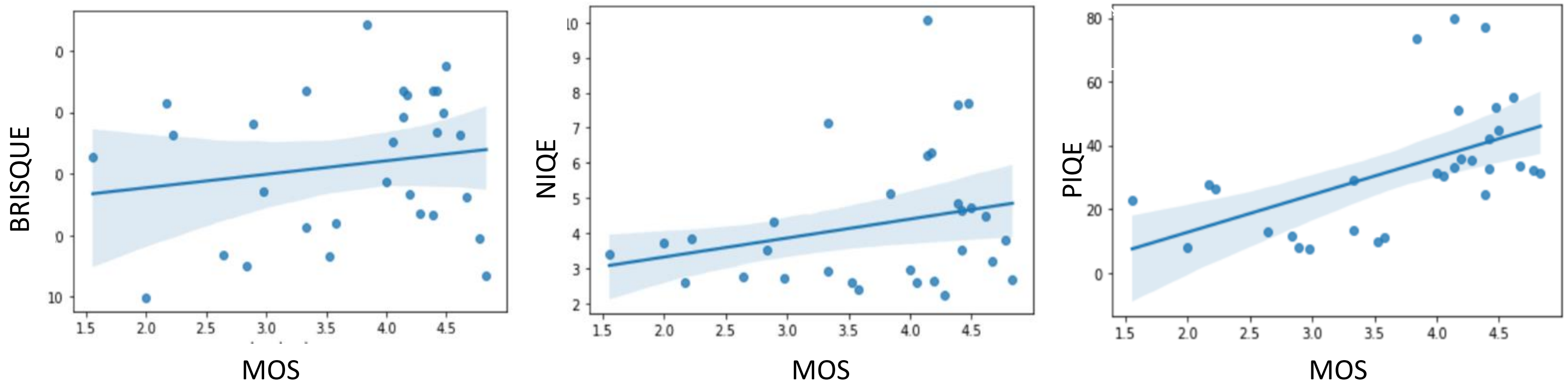
Collaboration with NTNU Norway

- The quantitative objective metrics correlate *poorly* with the subjective assessment
- When considering *visual pleasingness* for humans – *naturalness* is the largest factor.
- Deep learning approaches often introduce artifacts and unnatural colors.



Objective vs Subjective Image Quality

Carlo Tupa Indrauan, Skripsi, Fakultas Ilmu Komputer Universitas Indonesia, 2022.

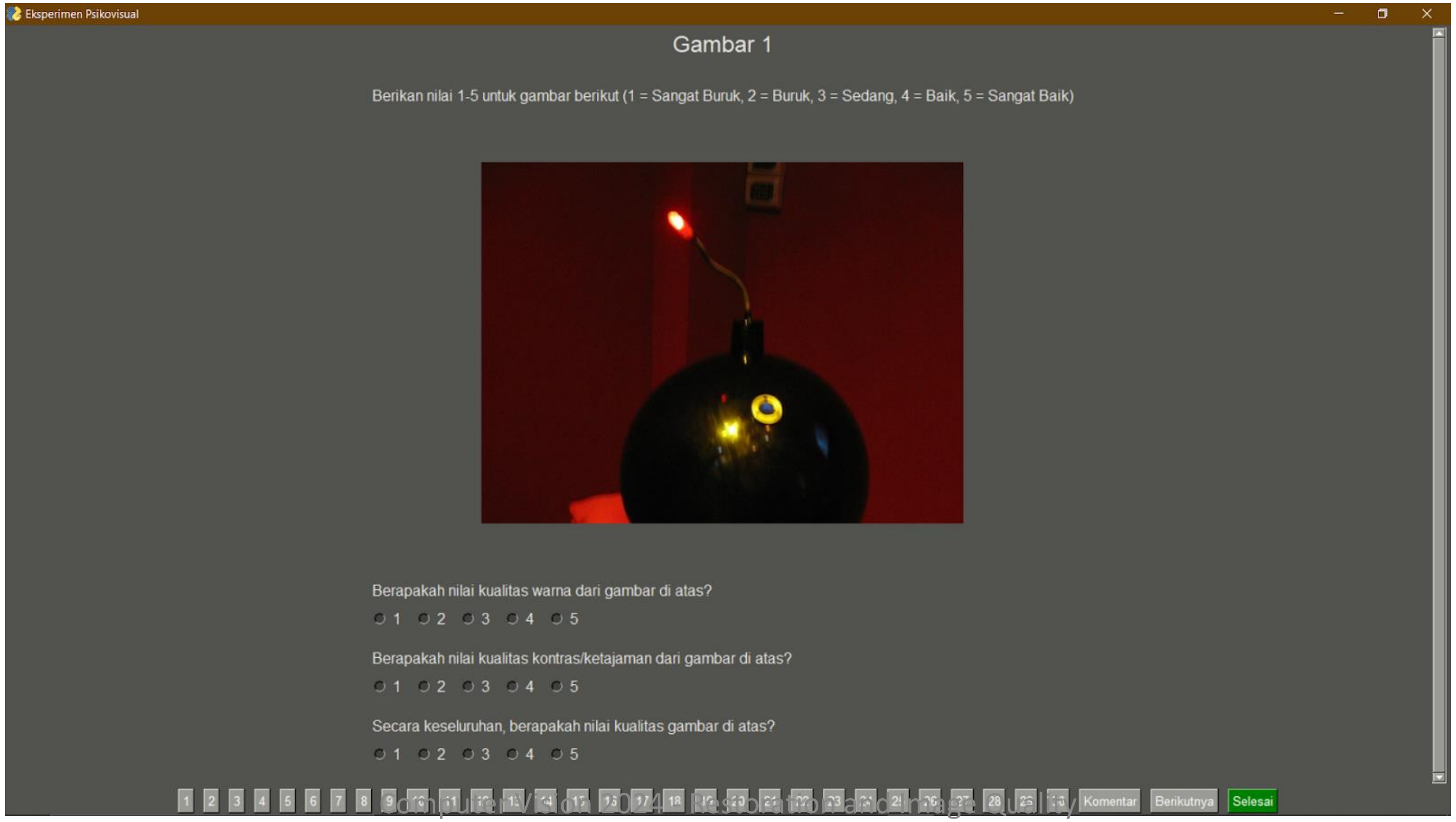


- BRISQUE has the best correlation with the subjective IQA scores, both using a large and small dataset
- We have confirmed that all measures: BRISQUE, NIQE, dan PIQE are still unable to measure image quality based on human perception.
- The larger the dataset, the correlation gets better.

Measuring Subjective Image Quality of Images

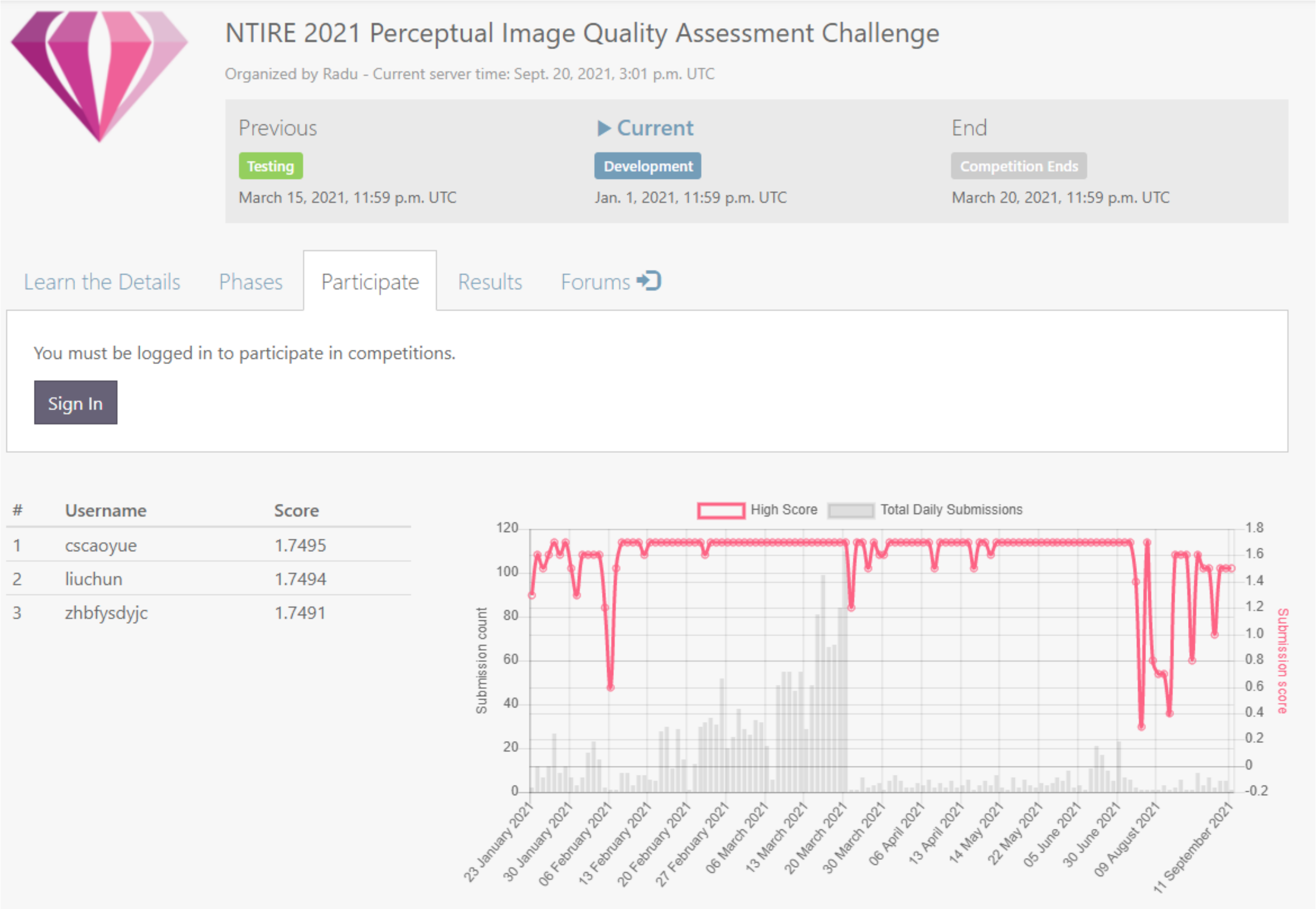
Carlo Tupa Indrauan, Skripsi, Fakultas Ilmu Komputer Universitas Indonesia, 2022.

- Psychovisual experiments are time and labor intensive… and difficult to do in person due to the pandemic. So, let’s make a standardized program



NTIRE 2021: Perceptual Image Quality Assessment Challenge

- NTIRE: New Trends in Image Restoration and Enhancement workshop (in conjunction with CVPR 2021).
- NTIRE challenge on perceptual image quality assessment: the task of predicting the perceptual quality of an image based on a set of prior examples of images and their perceptual quality labels.
- The aim is to obtain a network design/solution capable to produce high-quality results with the best correlation to the reference ground truth MOS score.



- https://data.vision.ee.ethz.ch/cvl/ntire22/
- https://codalab.lisn.upsaclay.fr/competitions/1567
- https://codalab.lisn.upsacla y.fr/competitions/1568

NTIRE 2022: Perceptual Image Quality Assessment (Full and No-Reference)

- NTIRE: New Trends in Image Restoration and Enhancement workshop and challenges on image and video processing (CVPR 2022).
- NTIRE challenge on perceptual image quality assessment: The aim is to obtain a network design/solution capable to produce high-quality results with the best correlation to the reference ground truth MOS score.
    - Full-Reference: algorithms that take reference images and distorted images as input.
    - No-Reference: algorithms that take only distorted images as input.