VAAD'AIR

Hue-Net: Intensity-based Image-to-Image Translation with Differentiable Histogram Loss Functions

Avi-Aharon, et al.

Joseph Wagane FAYE

INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES RENNES

IETR

VAADER IETR research team

Paper motivation

- New loss based on differentiable histogram construction for image-to-image translation.

- Application to color transfer.

- This presentation will focus on differentiable histogram computation and the earth mover's distance.

Image-to-Image Translation (I2I)

- I2I is a computer vision task.

- I2I aims to learn the mapping between an input image from one domain to an output image from another domain following the style or characteristics.

- I2I has a wide range of applications : image synthesis, segmentation, style transfer, restoration, and pose estimation, color transfer.

Image-to-Image Translation (I2I)

Style transfer

Source: StyleGAN

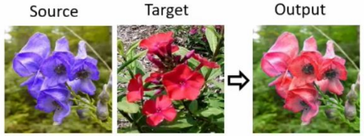
Image-to-Image Translation (I2I)

Color transfer

Source : HueNet

Key ideas

- Differentiable intensity based loss
  - Differentiable cyclic histogram construction
  - Cyclic Earth Mover's Distance (EMD)
- Differentiable mutual information loss
  - differentiable joint intensity histogram construction
  - statistical pixel-to-pixel similarity
- Conditionnal adversial loss

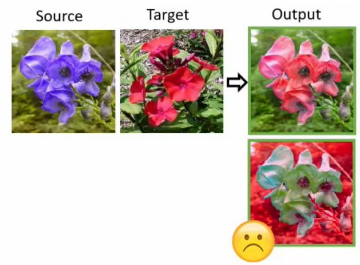


Key ideas

- Differentiable intensity based loss
  - Differentiable cyclic histogram construction
  - Cyclic Earth Mover's Distance (EMD)

- Differentiable mutual information loss
  - differentiable joint intensity histogram construction
  - statistical pixel-to-pixel similarity

- Conditionnal adversial loss



Key ideas

- Differentiable intensity based loss
  - Differentiable cyclic histogram construction
  - Cyclic Earth Mover's Distance (EMD)

- Differentiable mutual information loss
  - differentiable joint intensity histogram construction
  - statistical pixel-to-pixel similarity

- Conditionnal adversial loss

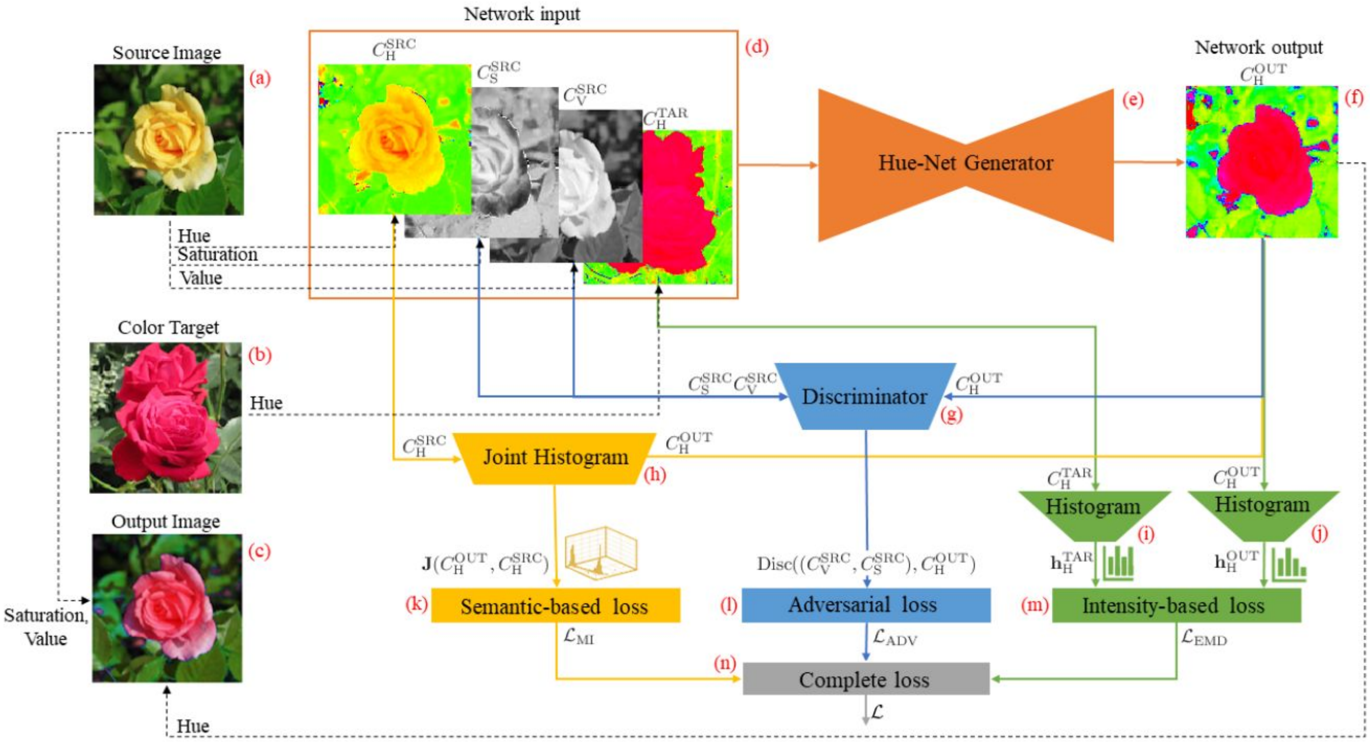
Key ideas

- Differentiable intensity based loss (Intensity based metric)
  - Differentiable cyclic histogram construction
  - Cyclic Earth Mover's Distance (EMD)

- Differentiable mutual information loss (Semantic based metric)
  - differentiable joint intensity histogram construction
  - statistical pixel-to-pixel similarity

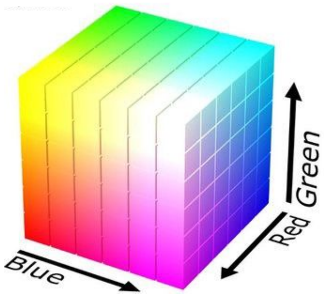
- Conditionnal adversial loss

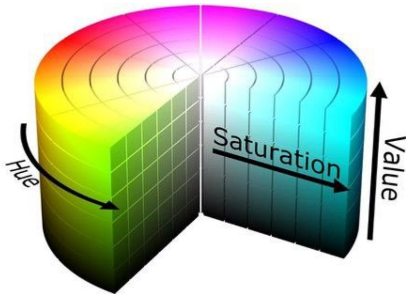
Hue-Net architecture



HSV color representation

- HSV color representation is used instead of RGB to get the cyclic aspect the loss.





Differentiable intensity histogram formulation

- Kernel density estimation :
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen–Rosenblatt* method).

$$\hat{f}_I(x) = \frac{1}{NW} \sum_{x=1}^{N} K\left(\frac{x - x_i}{W}\right)$$

$x = [x_1, x_2, \ldots, x_N]$ a gray scale image – Sample of an unknown distribution.

N Number of pixels in the image.

W The bandwidth.

Differentiable intensity histogram formulation

- **Kernel density estimation :**
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen-Rosenblatt* method).

$$\hat{f}_I(x) = \frac{1}{NW} \sum_{x=1}^{N} K\left(\frac{x - x_i}{W}\right)$$

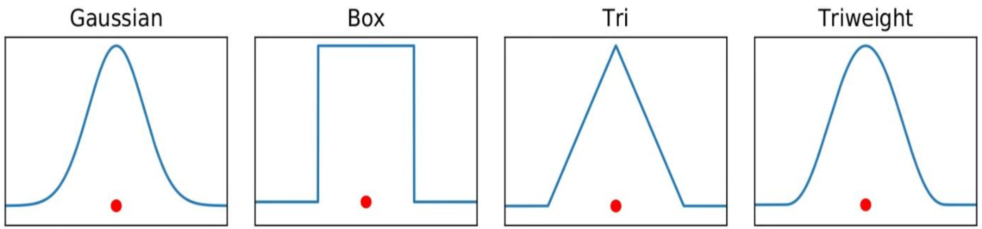
$$\hat{f}_I(x) = \sum_{observations} K\left(\frac{x - obersation}{bandwidth}\right)$$

Differentiable intensity histogram formulation

- Kernel density estimation :
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen-Rosenblatt* method).

The kernel function K is typically:

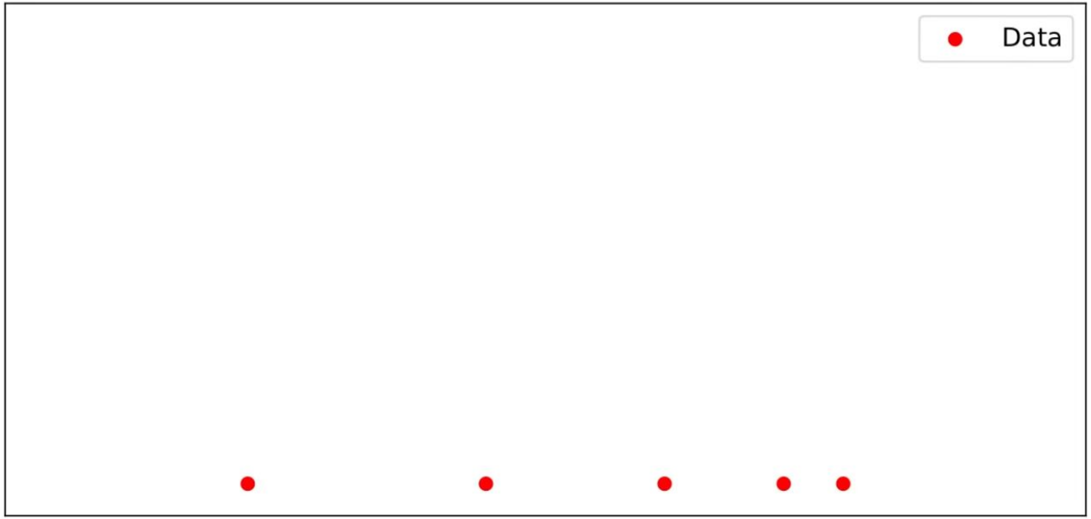
- everywhere non-negative : $K(x) \geq 0$, for every x.
- symmetric : $K(x) = K(-x)$ for every x.
- decreasing: $K'(x) \leq 0$ for every $x>0$



$$K(z) = \frac{d}{dz}\sigma(z) = \sigma(z)\sigma(-z)$$

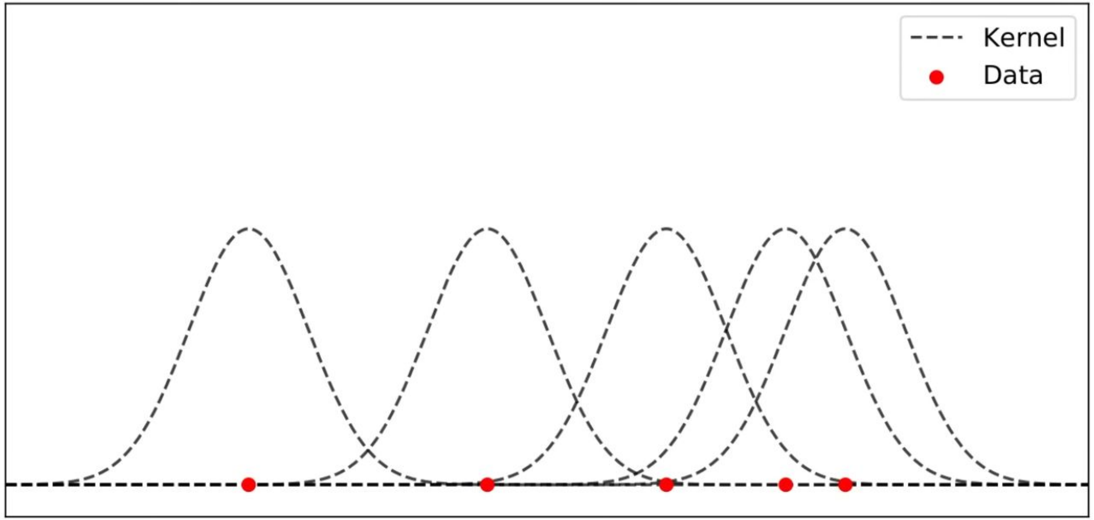
Differentiable intensity histogram formulation

- Kernel density estimation :
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen–Rosenblatt* method).



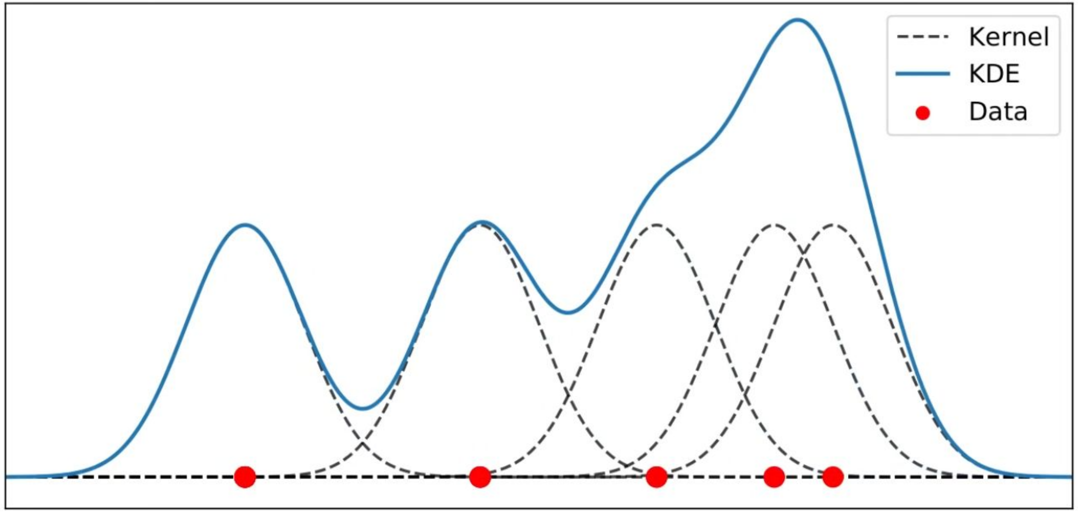
Differentiable intensity histogram formulation

- **Kernel density estimation :**
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen–Rosenblatt* method).



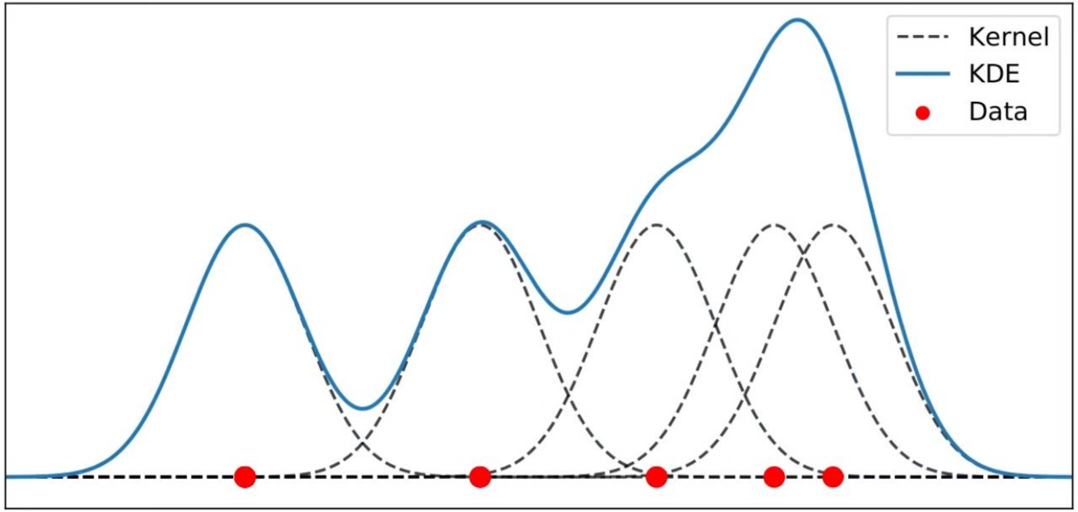
Differentiable intensity histogram formulation

- Kernel density estimation :
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen–Rosenblatt* method).



Differentiable intensity histogram formulation

- Kernel density estimation :
  - Kernel density estimation is the process of estimating an unknown probability density function using a kernel function (also known as *Parzen-Rosenblatt* method).

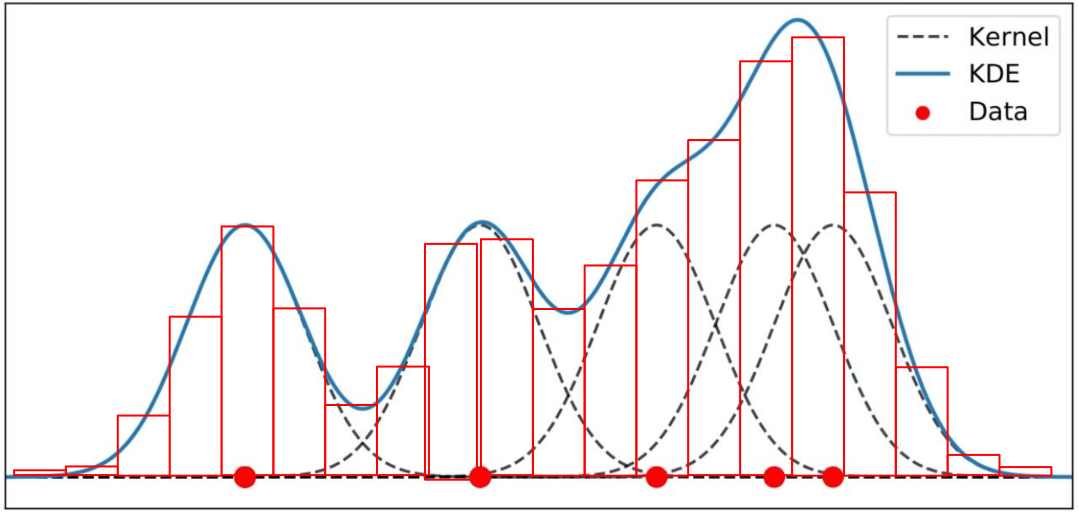


Differentiable intensity histogram formulation

- Differentiable histogram construction

  * The total range is partitioned into K subintervals B_k of length $L = \frac{1}{K}$ and center $\mu_k = L\left(k + \frac{1}{2}\right)$.

  *Then $B_k = [kL,\, (k+1)L]$

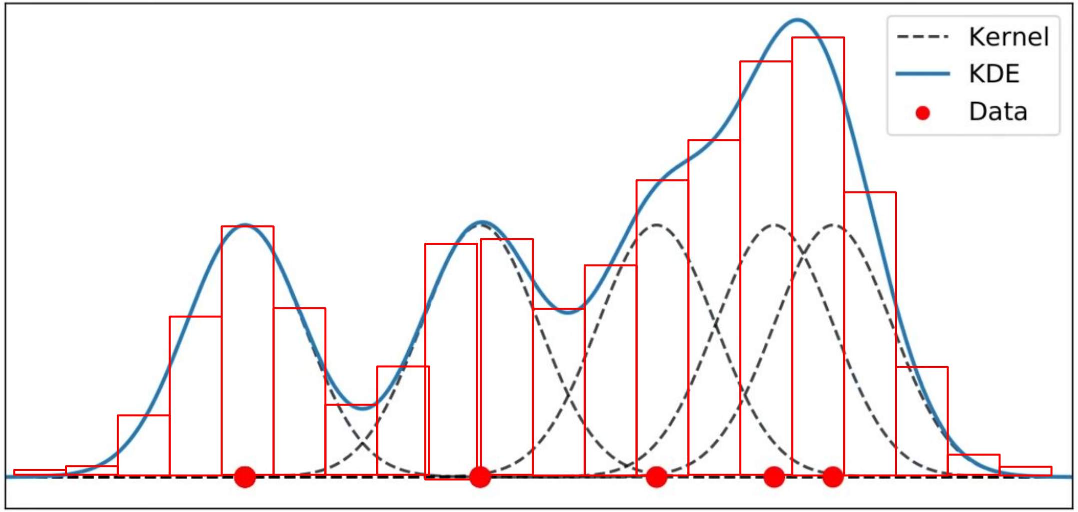


Differentiable intensity histogram formulation

- **Differentiable histogram construction**

  * The total range is partitioned into K subintervals B_k of length $L = \frac{1}{K}$ and center $\mu_k = L\left(k + \frac{1}{2}\right)$.

  *Then $B_k = [kL,\ (k+1)L]$



$$P_I(k) \cong P_r(g \in B_k) = \int_{B_k} \hat{f}_I(g)dg$$

$$P_I(k) = \frac{1}{N}\sum_{x=1}^{N} \pi_k(I(x))$$

$$where\,,$$

$$\pi_k(z) \cong \sigma\left(\frac{z-\mu_k + L/2}{W}\right) - \sigma\left(\frac{z-\mu_k - L/2}{W}\right)$$

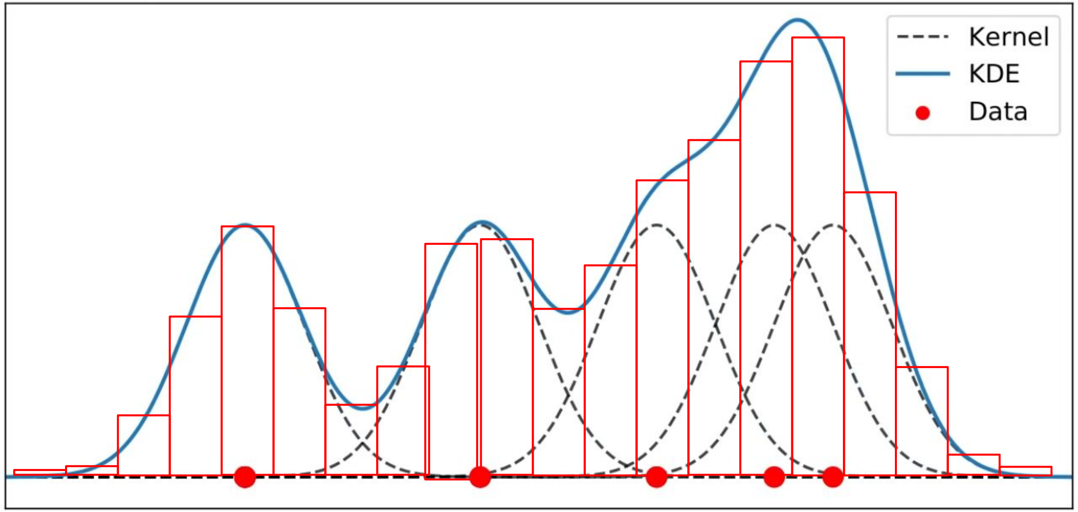
differentiable approximation of Rect function.

Differentiable intensity histogram formulation

- **Differentiable histogram construction**

* The total range is partitioned into K subintervals B_k of length $L = \frac{1}{K}$ and center $\mu_k = L\left(k + \frac{1}{2}\right)$.

*Then $B_k = [kL, \,(k+1)L]$



$$P_I(k) \cong P_r(g \in B_k) = \int_{B_k} \hat{f}_I(g)dg$$

$$P_I(k) = \frac{1}{N}\sum_{x=1}^{N} \pi_k(I(x))$$

where ,

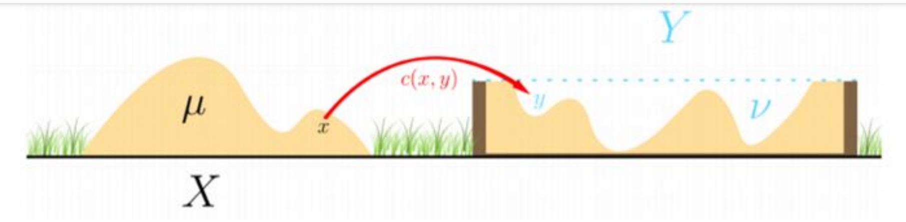
$$\pi_k(z) \cong \sigma\left(\frac{z - \mu_k + L/2}{W}\right) - \sigma\left(\frac{z - \mu_k - L/2}{W}\right)$$

The histogram h of an image I is defined as follows :

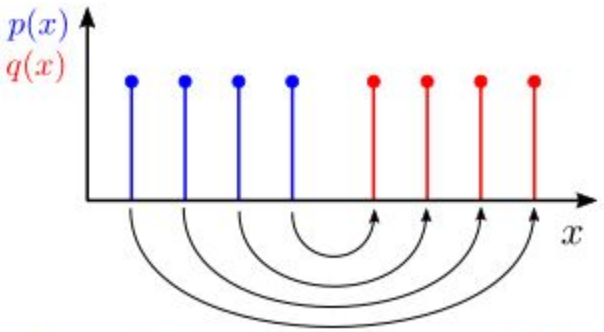
$$h = \{\mu_k, \, P_I(k)\}_{k=0}^{K-1}$$

Earth Mover's Distance (EMD).

We assumed previously that we don't know the distribution of the images and we estimate them by KDE.
We can now compute a loss based on these distributions based on Optimal Transport.

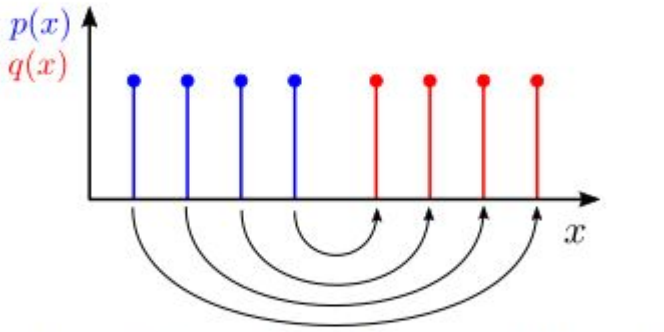


Earth Mover's Distance (EMD).



Let's define two uniform distributions p(x) and q(x).

Earth Mover's Distance (EMD).

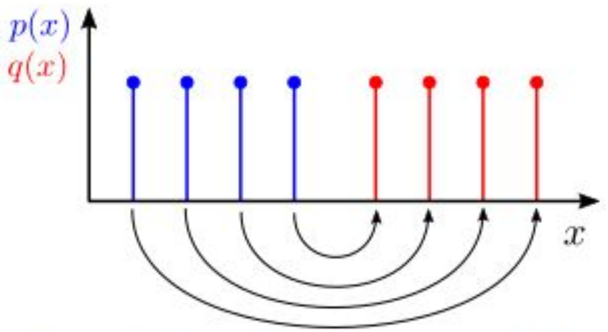


Let's define two uniform distributions p(x) and q(x).

We can define a coupling matrix. That represents how much probability mass from one point in the support of *p(x)* is assigned to a point in the support of *q(x)*.

$$P = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1/4 & 0 \\ 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \end{pmatrix}$$

Earth Mover's Distance (EMD).



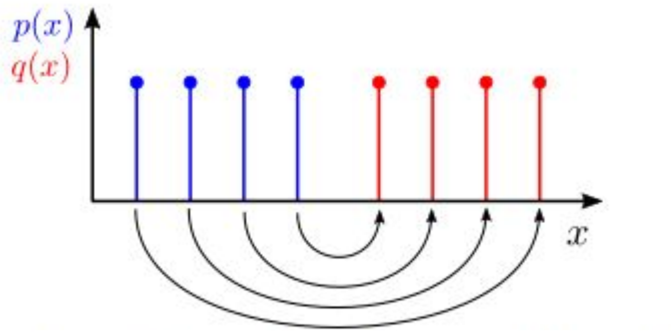
If we assume that the supports for ρ(x) and q (x) are {1, 2, 3, 4} and {5,6,7,8}, respectively, the matrix cost is :

$$C = \begin{pmatrix} 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

We can define a coupling matrix. That represents how much probability mass from one point in the support of *ρ(x)* is assigned to a point in the support of *q(x)*.

$$P = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1/4 & 0 \\ 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \end{pmatrix}$$

Earth Mover's Distance (EMD).



With these definitions, the total cost can be calculated as the Frobenius inner product between P and C:

$$\left\langle C, P \right\rangle = \sum_{ij} C_{ij} P_{ij}$$

Earth Mover's Distance (EMD).

More generally :

$$h_j = \{\mu_k,\ P_{I_j}(k_j)\}_{k_j=0}^{K-1} \quad ;\ j = 1, 2$$

We aim to find a flow that minimizes the overall cost :

$$W(h_1,\ h_2,\ P) = \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} c_{k_1 k_2} p_{k_1 k_2}$$

For one-dimensional histograms with equal areas, EMD has been shown to be equivalent to Mallows distance which has the following closed-form solution :

$$\mathcal{D}_{\mathrm{EMD}}(\mathbf{h_1}, \mathbf{h_2}) = \left(\frac{1}{K}\right)^{\frac{1}{l}} \|\mathrm{CDF}(\mathbf{h_1}) - \mathrm{CDF}(\mathbf{h_2})\|_l,$$

Semantic-based metric

The mutual information between two images is defined as follows :

$$\mathcal{I}(I_1, I_2) = \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} P_{I_1,I_2}(k_1, k_2) \log \frac{P_{I_1,I_2}(k_1, k_2)}{P_{I_1}(k_1) P_{I_2}(k2)},$$

Joint entropy

$$\mathcal{I}(I_1, I_2) = \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} P_{I_1,I_2}(k_1, k_2) \log \frac{P_{I_1,I_2}(k_1, k_2)}{P_{I_1}(k_1) P_{I_2}(k2)},$$

Semantic-based metric

Differentiable joint intensity histogram :

- Multivariate KDE is used for the estimation of the joint density.

$$\hat{f}_{I_1,I_2}(g_1,g_2) = \frac{1}{N}|\mathbf{W}|^{-1/2}\sum_{x\in\Omega}\mathcal{K}\left(\mathbf{W}^{-1/2}(\mathbf{I}(x)-\mathbf{g})\right)$$

Using the definition of π_k we can expressed the value of joint histogram k_1, k_2 bin as :

$$P_{I_1 I_2}(k_1,\, k_2) \;=\; \frac{1}{N}\sum_{1}^{N}\pi_{k_1} I_1(x))\pi_{k2}\,(I_2(x))$$

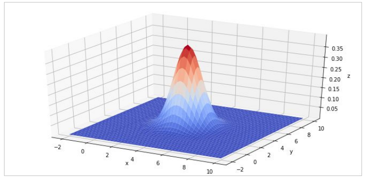
$$J(I_1,\, I_2) \;=\; \frac{1}{N}P_1 P_2^T$$

Semantic-based metric

Differentiable joint intensity histogram :

- Multivariate KDE is used for the estimation of the joint density.

$$\hat{f}_{I_1,I_2}(g_1, g_2) = \frac{1}{N}|\mathbf{W}|^{-1/2}\sum_{x\in\Omega}\mathcal{K}\left(\mathbf{W}^{-1/2}(\mathbf{I}(x)-\mathbf{g})\right)$$



Using the definition of π_k we can expressed the value of joint histogram k_1, k_2 bin as :

$$P_{I_1 I_2}(k_1,\, k_2) \;=\; \frac{1}{N}\sum_{1}^{N}\pi_{k_1} I_1(x))\pi_{k2}\,(I_2(x))$$

$$J(I_1,\, I_2) \;=\; \frac{1}{N}P_1 P_2^T$$

Semantic-based metric

Differentiable joint intensity histogram :

- Multivariate KDE is used for the estimation of the joint density.

$$\hat{f}_{I_1,I_2}(g_1, g_2) = \frac{1}{N}|\mathbf{W}|^{-1/2} \sum_{x\in\Omega} \mathcal{K}\left(\mathbf{W}^{-1/2}(\mathbf{I}(x) - \mathbf{g})\right)$$

Using the definition of π_k we can expressed the value of joint histogram k_1, k_2 bin as :

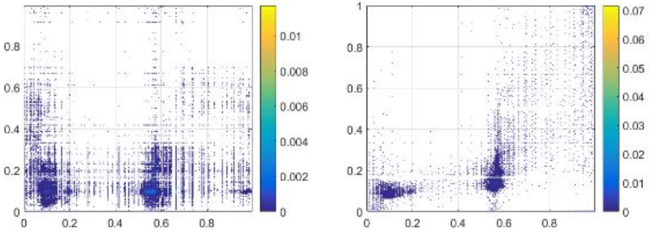
$$P_{I_1 I_2}(k_1,\, k_2) \;=\; \frac{1}{N}\sum_{1}^{N} \pi_{k_1} I_1(x))\pi_{k2}\,(I_2(x))$$

$$J(I_1,\, I_2) \;=\; \frac{1}{N} P_1 P_2^T$$

Semantic-based metric

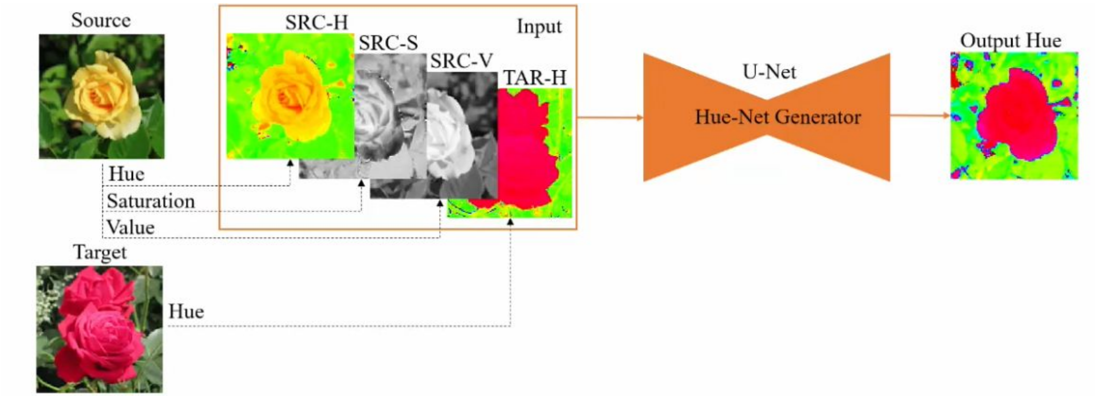
The semantic based metric is defined as following :

$$\mathcal{D}_{\mathrm{MI}}(I_1, I_2) = 1 - \frac{\mathcal{I}(I_1, I_2)}{\mathcal{H}(I_1, I_2)}$$



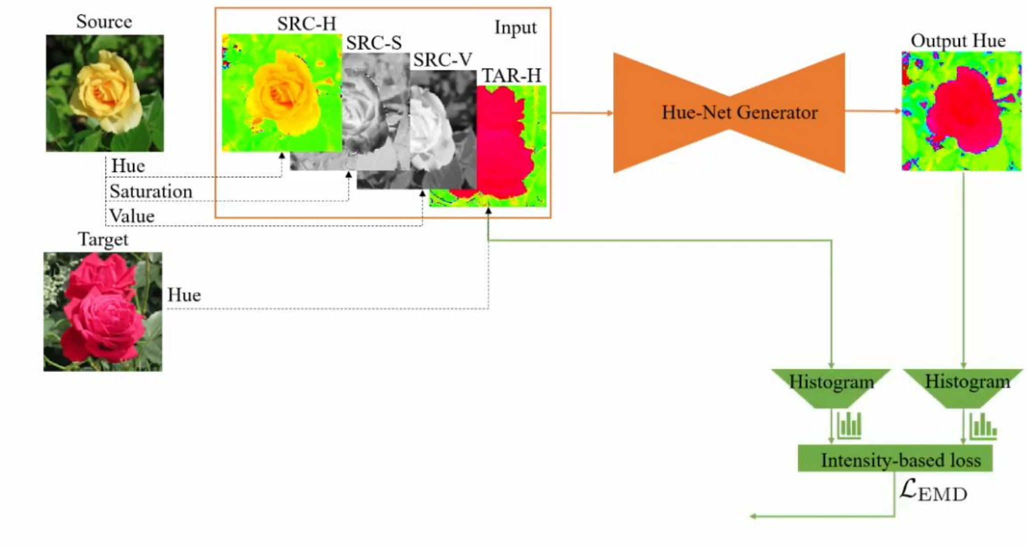
2D joint histograms of source and target

Summary of the architecture



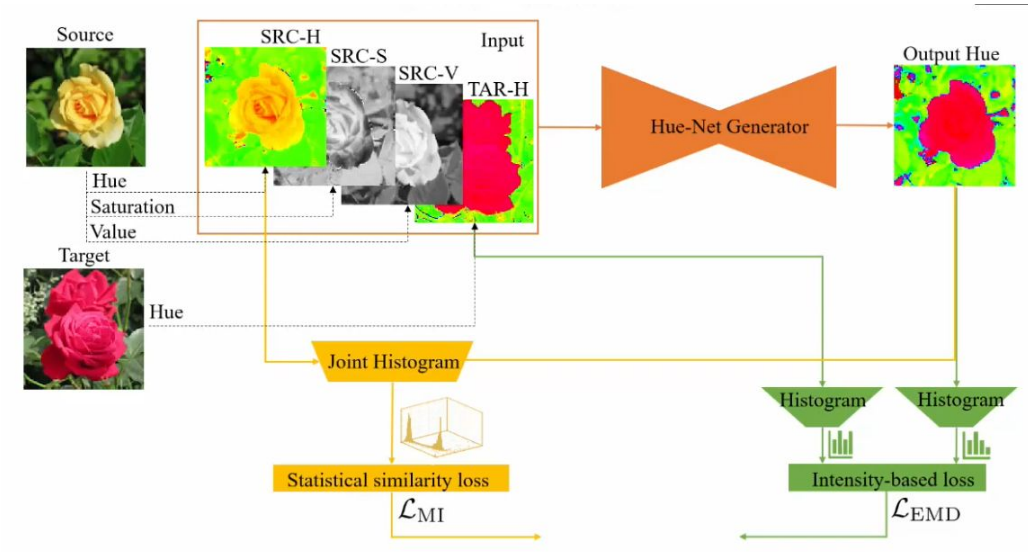
https://forms.gle/wAmRSChHpa65d1So7

Summary of the architecture

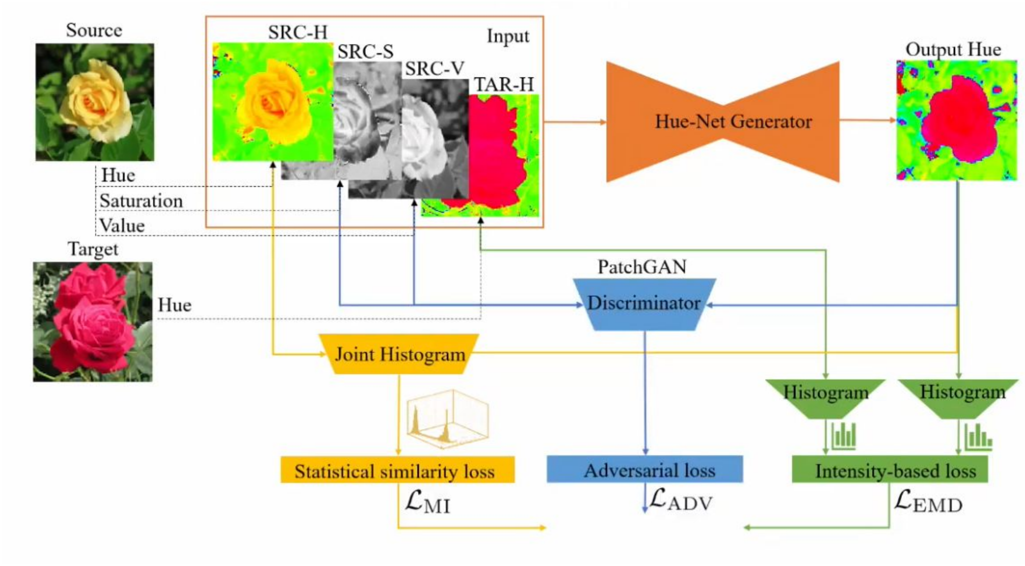


https://forms.gle/wAmRSChHpa65d1So7

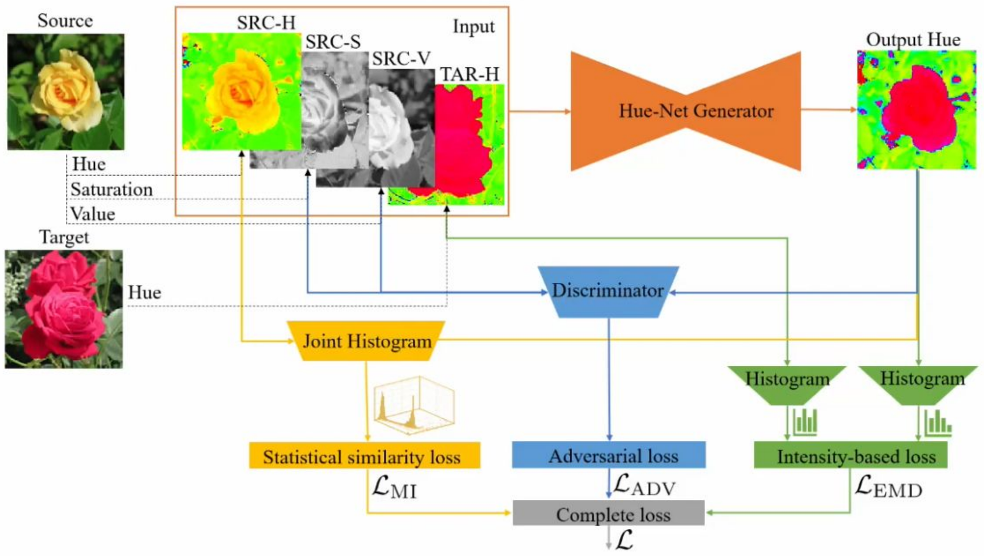
Summary of the architecture



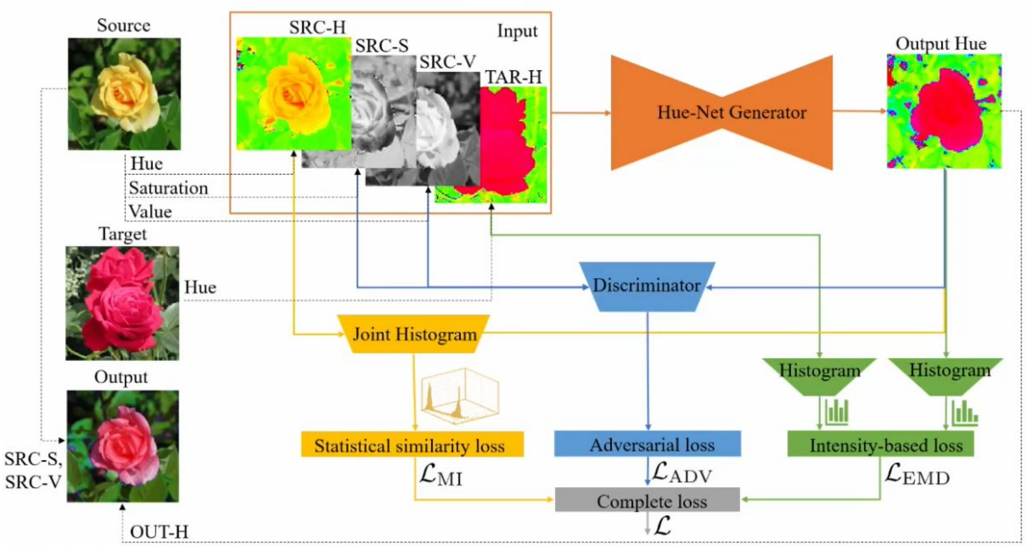
Summary of the architecture



Summary of the architecture



Summary of the architecture



Results

Datasets:

Flower Color Images – (210 images (128 x128x3))
Kaggle web site:
https://www.kaggle.com/olgabelitskaya/flower-color-images

Oxford 102 Category Flower Dataset (8189 images)

Cars Dataset - Stanford AI Lab (16,185 images)

https://forms.gle/wAmRSChHpa65d1So7

Results

https://forms.gle/wAmRSChHpa65d1So7

Results

https://forms.gle/wAmRSChHpa65d1So7

Results

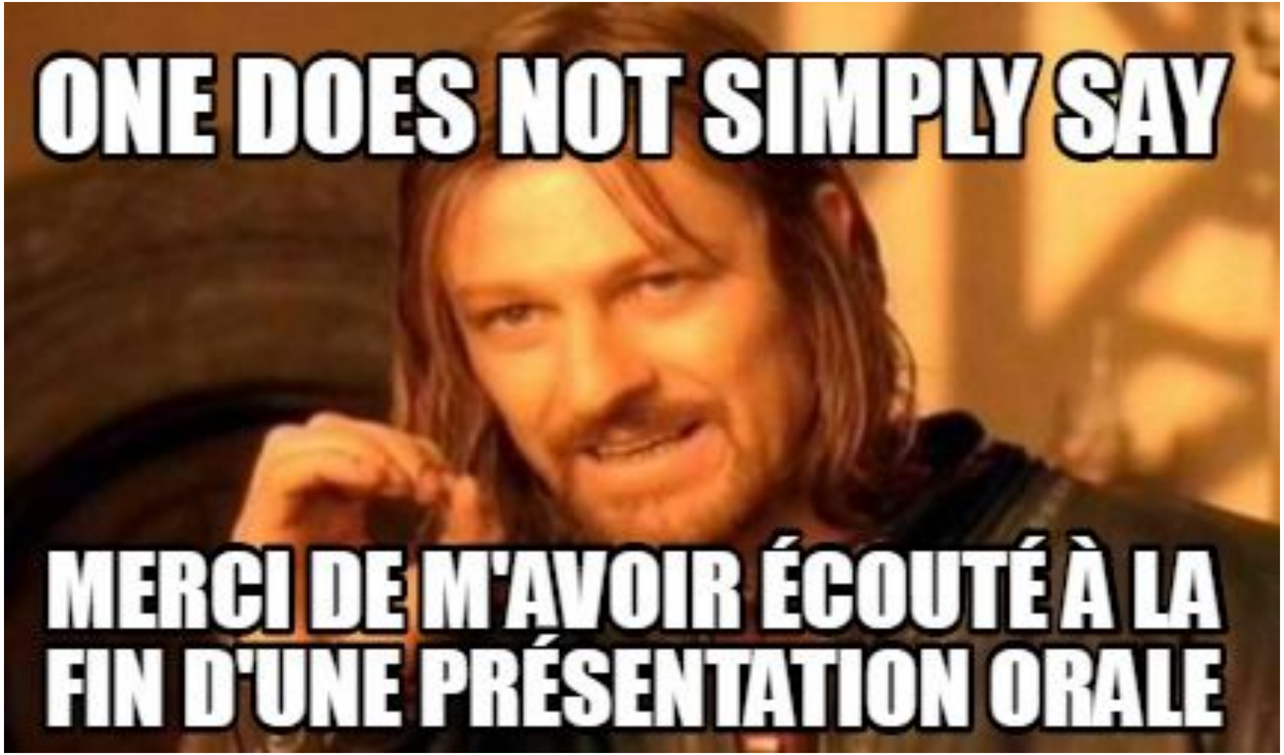
Perceptual realism

	Real	Fake
Target (real image)	68	32
Output (painted image)	50.5	49.5

https://forms.gle/wAmRSChHpa65d1So7

Summary and key contributions

- Differentiable cyclic and joint intensity histogram construction.
- Intensity-based loss with EMD
- Mutual information loss for statistical pixel-to-pixel similarity.
- Unified deep learning framework.


ONE DOES NOT SIMPLY SAY
MERCI DE M'AVOIR ÉCOUTÉ À LA
FIN D'UNE PRÉSENTATION ORALE