



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Multimodal Style Transfer via Graph Cuts Experiments

Clément BELIVEAU, Paul GODARD, Alvaro Scaramberg, Santiago Quinteros, Rogerio
Bombardelli

Summary

1. Style Transfer / Related works
2. Multimodal Style Transfer
3. Methodology
4. Ressources
5. Conclusion / Wrap up

Summary

- 1. Style Transfer / Related works**
- 2. Multimodal Style Transfer**
- 3. Methodology**
- 4. Resources**
- 5. Conclusion / Wrap up**

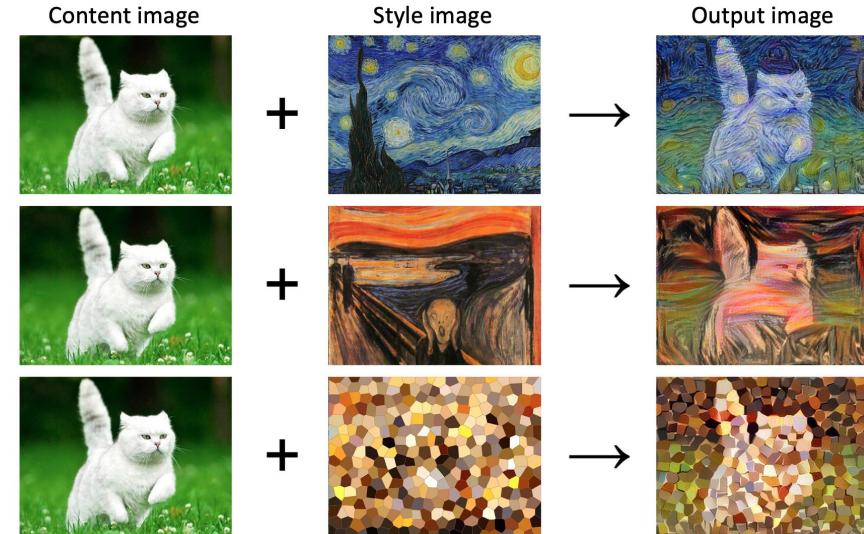
Style Transfer: The beginning

4

Started in 2015 with the paper A Neural Algorithm of Artistic Style, Gatys et al (2015).

Their method utilized Convolutional Neural Networks (CNNs) to separate and recombine the content and style of images.

- Content loss and style loss
- VGG network as a feature extractor
- Image is initialized as white noise and iteratively updated to minimize the total loss.
- Gradient descent algorithms.



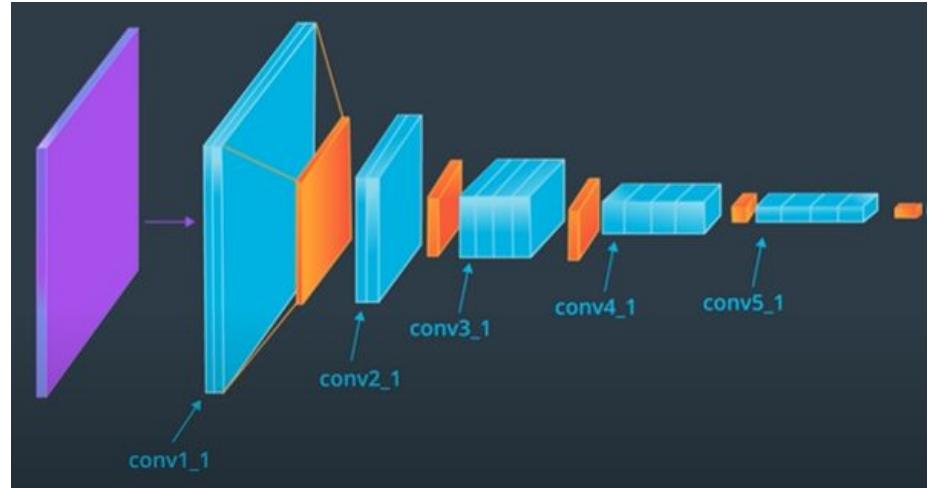
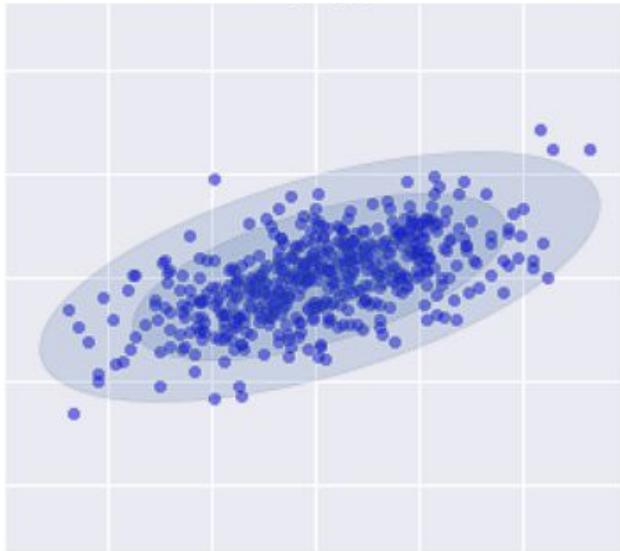
Summary

1. Style Transfer / Related works
2. **Multimodal Style Transfer**
3. Methodology
4. Resources
5. Conclusion / Wrap up

Multimodal Style Transfer:

6

Previous methods assume that the distribution associated to a style is unimodal



The style can be characterized by global statistics such as the Covariance or Gram matrices.

Multimodal Style Transfer:

7

New approach



A Style can be decomposed
into a set of *Substyles*

- Style of the Sky
- Style of the tree
- Style of the buildings

...

Multimodal Style Transfer:

8

New approach



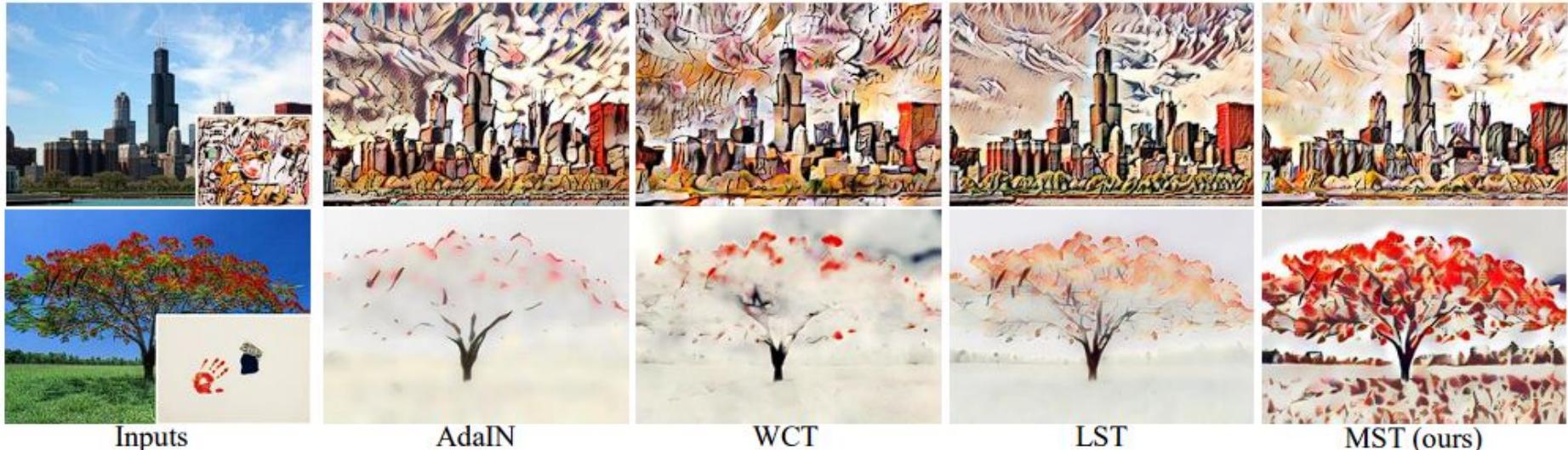
A Style can be decomposed into a set of *Substyles*



Only the style of the sky was captured.

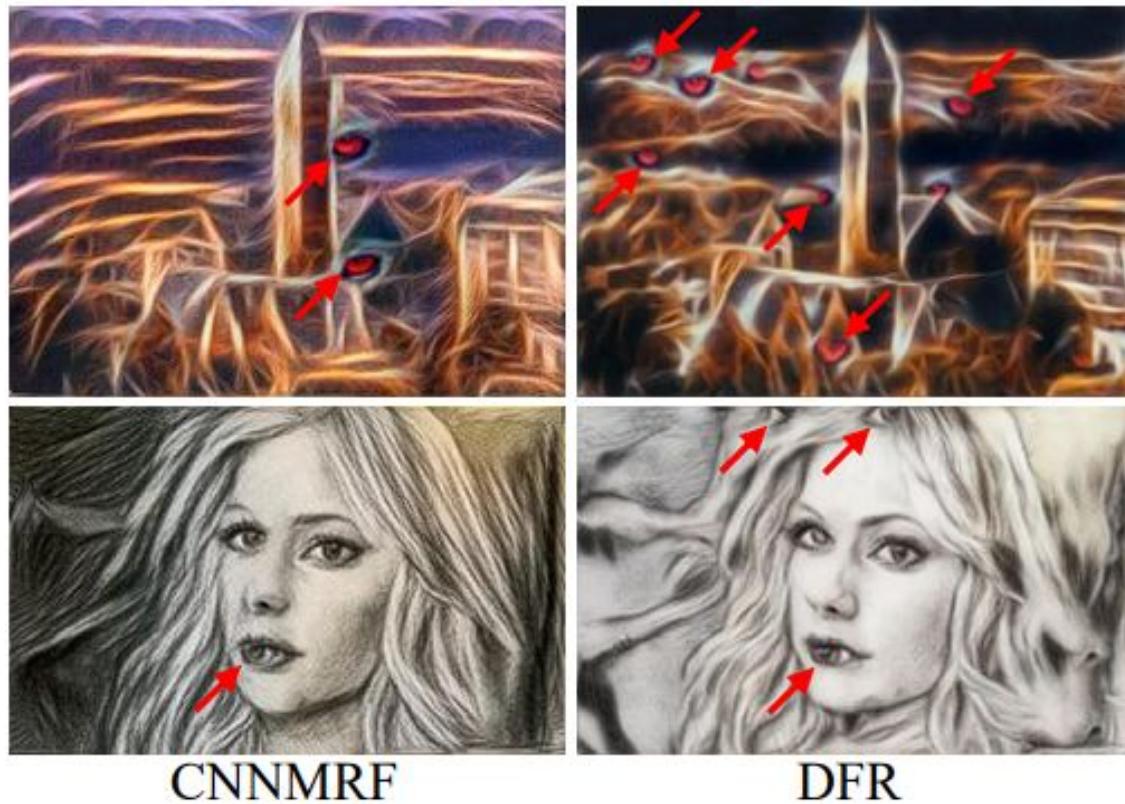
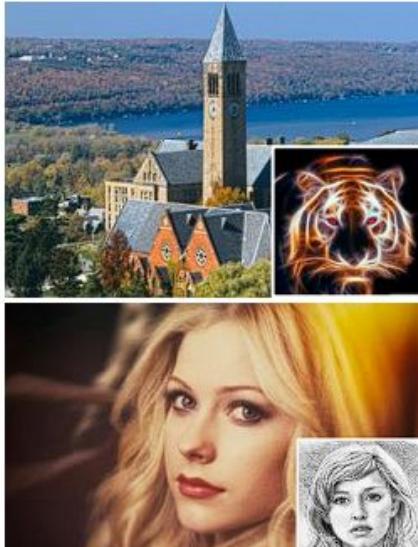
Multimodal Style Transfer:

9



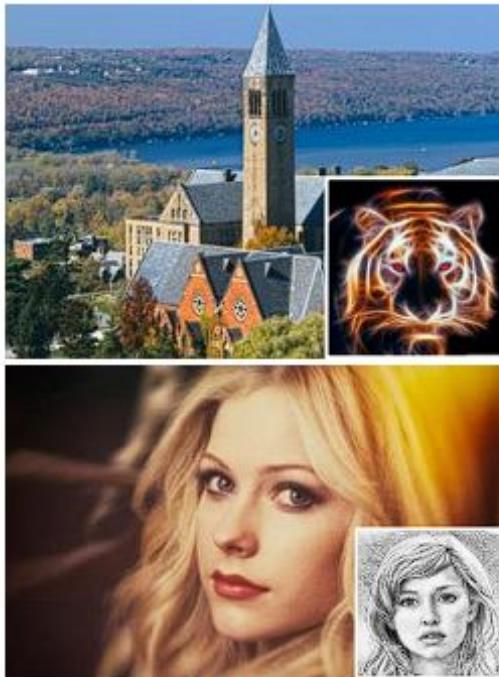
Multimodal Style Transfer:

10



Multimodal Style Transfer:

11



Inputs



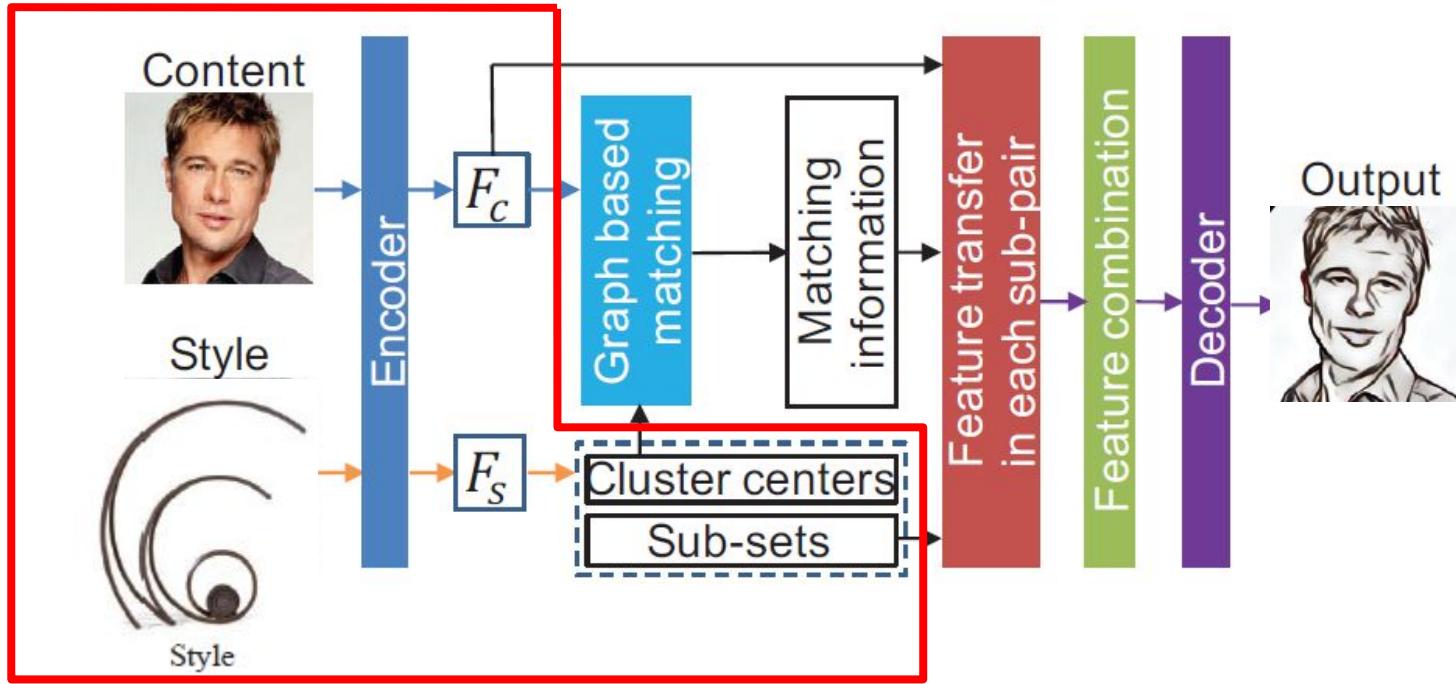
MST (ours)

Summary

1. Style Transfer / Related works
2. Multimodal Style Transfer
3. **Methodology**
4. Resources
5. Test and results

Methodology:

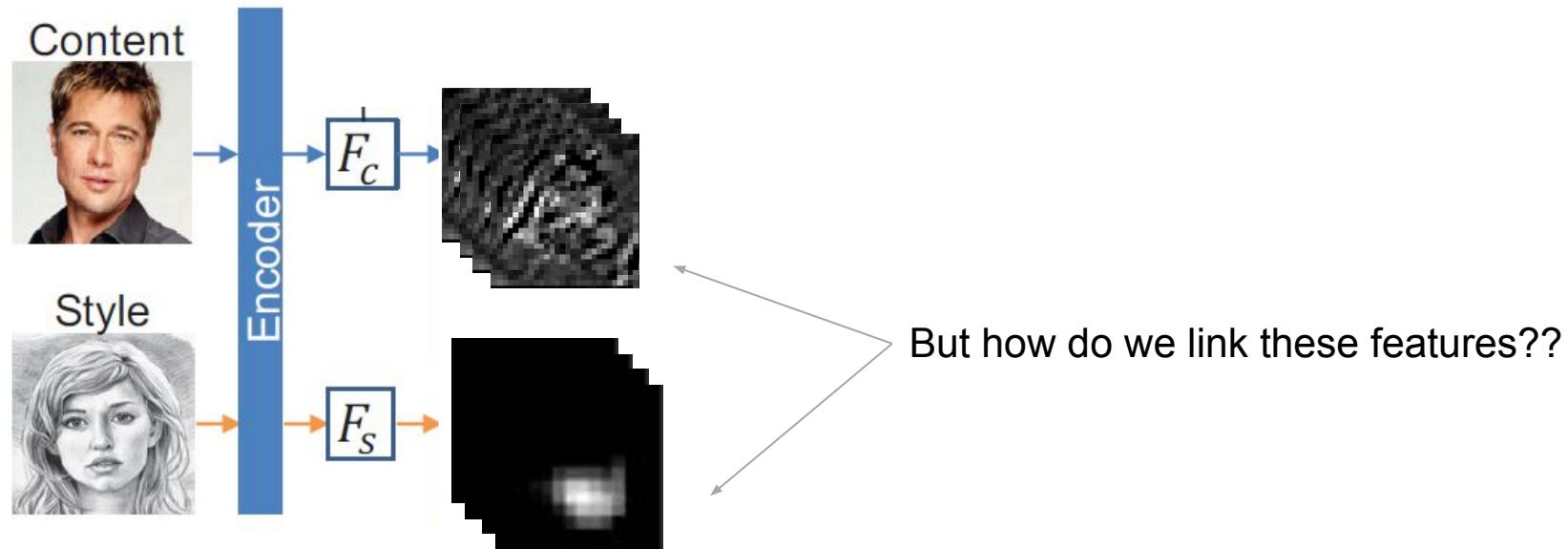
13



Multimodal Style Transfer: Visualizing the substYLES

14

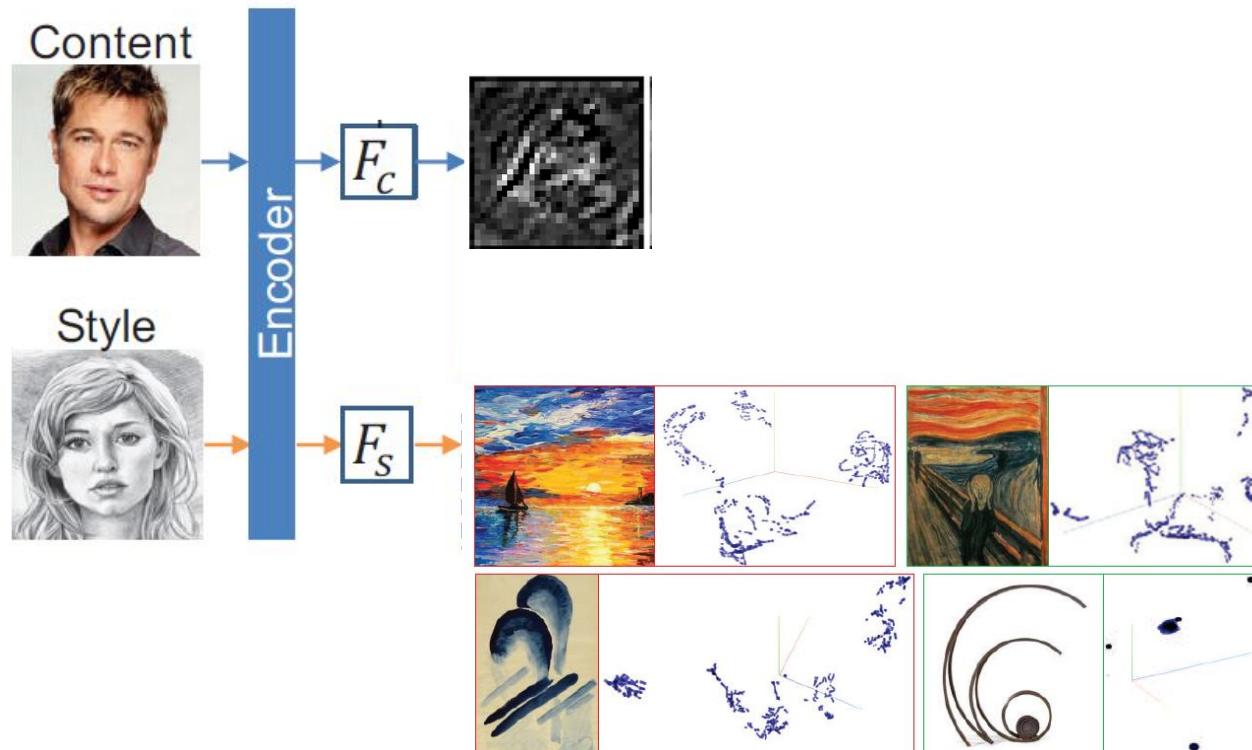
For a given style and content image, we can extract its deep features via a pre-trained encoder.
In this case VGG-19.



- Distinguish each different style.
- Assign a style for each feature of the content image.
- Apply the style transfer for each part of the content image.

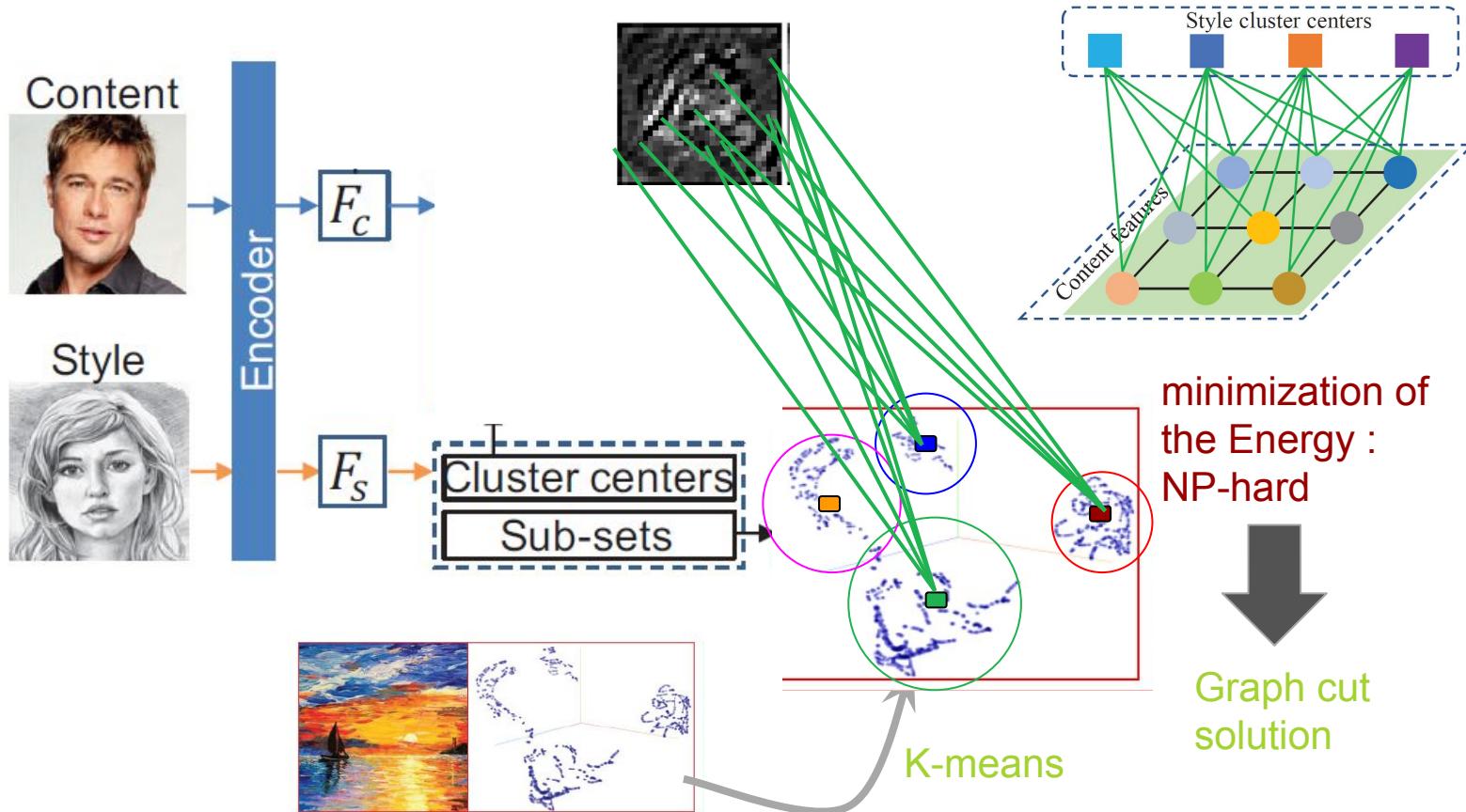
Methodology: feature extraction

15



Methodology: graph cut

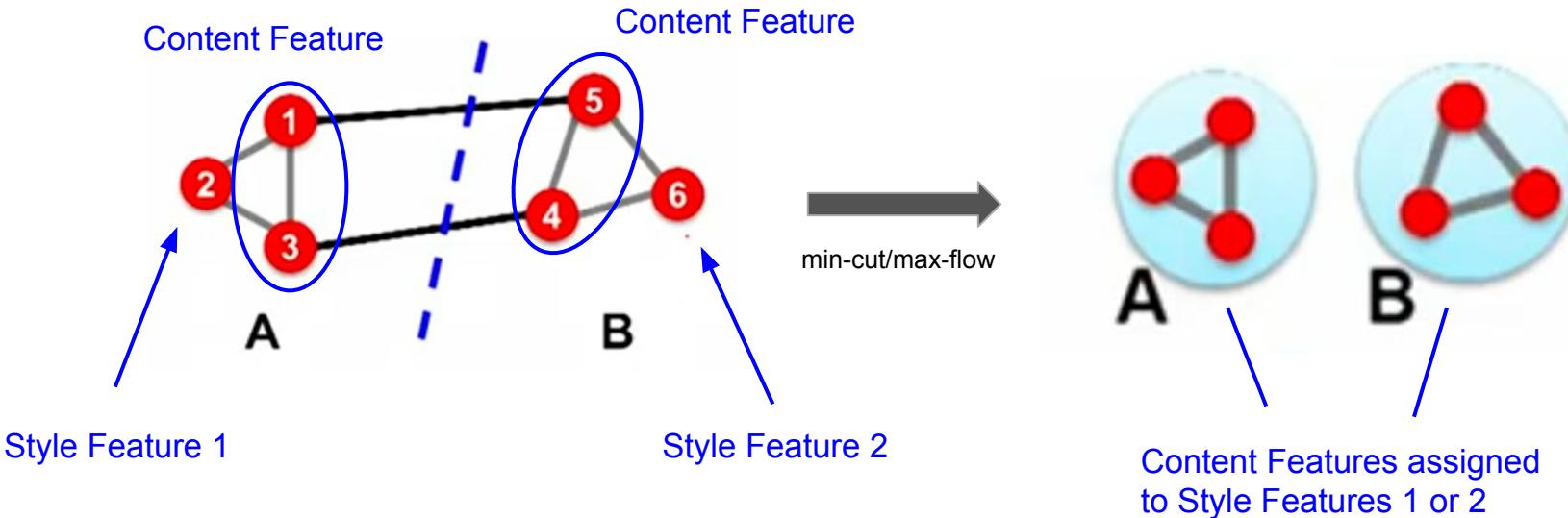
16



Methodology: graph cut

17

Basically, what graph cut does is it uses a min-cut/max-flow to find the optimal cut in the graph



Summary

1. Style Transfer / Related works
2. Multimodal Style Transfer
3. Methodology
- 4. Resources**
5. Test and results

What was available : The paper code

19

Code fully furnished by the authors

BUT :

- In tensorflow : old version
- Not maintained since 5 years

The screenshot shows a GitHub repository named 'MST'. The repository is public, has 1 branch, and 0 tags. It contains 11 commits from user 'yulunzhang' made 5 years ago. The commits include updates to README.md, creation of 'data', 'models', and 'src' folders, and scripts like 'scripts_MST.sh' and 'test_MST.py'. The repository description is 'TensorFlow code for our ICCV 2019 paper "Multimodal Style Transfer via Graph Cuts"' and the title of the main file is 'Multimodal Style Transfer via Graph Cuts'.

MST · Public

Watch 8

master · 1 Branch · 0 Tags

Go to file Add file Code

yulunzhang · Update README.md · 4b1c6c3 · 5 years ago · 11 Commits

data · Add files via upload · 5 years ago

models · TensorFlow code for ICCV-2019-MST · 5 years ago

src · TensorFlow code for ICCV-2019-MST · 5 years ago

README.md · Update README.md · 5 years ago

scripts_MST.sh · TensorFlow code for ICCV-2019-MST · 5 years ago

test_MST.py · TensorFlow code for ICCV-2019-MST · 5 years ago

README

MST

TensorFlow code for our ICCV 2019 paper "Multimodal Style Transfer via Graph Cuts"

Multimodal Style Transfer via Graph Cuts

What was available : A newer version

20

Screenshot of a GitHub repository page for "Multimodal-Style-Transfer-via-Graph-Cuts". The repository is public and has 1 branch and 0 tags. The master branch was last updated 4 years ago by Rakshit-Shetty. The commit history shows 75 commits, all of which are 4 years old. The commits are listed below:

| File / Action | Description | Date |
|--|----------------------------|-------------|
| Notebook_for_training_and_testing_on_Google... | Created using Colaboratory | 5 years ago |
| README.md | Cleanup | 5 years ago |
| VGGNormalized.py | resolving typos and bugs | 4 years ago |
| dataset.py | Cleanup | 5 years ago |
| decoder.py | resolving typos and bugs | 4 years ago |
| feature_transfer.py | resolving typos and bugs | 4 years ago |
| model.py | resolving typos | 4 years ago |
| model_state.pth | Create model_state.pth | 4 years ago |
| test.py | resolving typos and bugs | 4 years ago |
| train.py | resolving typos and bugs | 4 years ago |

A new version in pytorch

BUT :

- Only possible to run it on collab
- No handling of file tree
- No possibility to launch large scale automatic experiments
- Not maintained since 4 years

What we have done :

21

- Code which can run locally
- Able to launch large scale experiments
- OutputDisplay
- Automatic Result savings

 **Multimodal-Style-Transfer-via-Graph-Cuts** Public

forked from [Rakshit-Shetty/Multimodal-Style-Transfer-via-Graph-Cuts](#)

 master  2 Branches  0 Tags Go to file Add file

This branch is 28 commits ahead of [Rakshit-Shetty/Multimodal-Style-Transfer-via-Graph-Cuts:master](#).

 Contribute  Sync fork

 **ClementBLV** Data 7992717 · 37 minutes ago

| File | Type | Last Commit |
|--|-------------------------------|-------------|
| __pycache__ | Data | 3 |
| data | Data | 4 |
| .gitignore | Feature: ignore useless files | 4 |
| Notebook_for_training_and_testing_on_Google... | Data | |
| README.md | Features : Update README.md | |
| VGGNormalized.py | resolving typos and bugs | |

Summary

1. Style Transfer / Related works
2. Multimodal Style Transfer
3. Methodology
4. Resources
5. **Test and results**

Observations

23



Content



Style

Observations

24



k=1



k=2



k=3



k=4



k=5



k=6



k=7



k=8



k=9



k=10

Optimal Cluster Number :

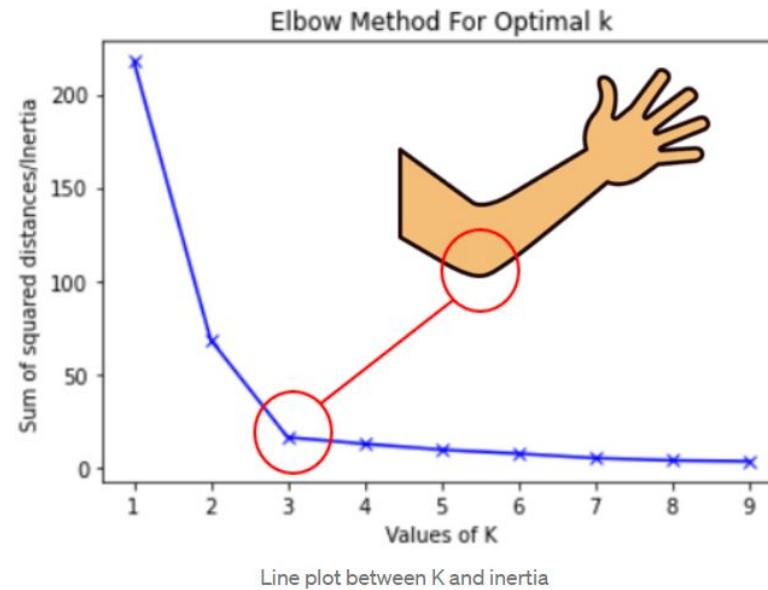
25

Silhouette Critérium

- measure the consistency within a cluster
- find the cluster with the highest silhouette score

The silhouette score measures how similar an object is to its own cluster (cohesion) compared to other clusters (separation)

Elbow Curve



Optimal Cluster Number :

26



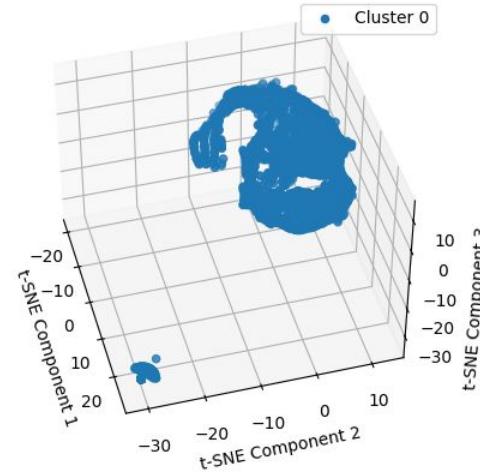
Style

Optimal Cluster Number :

27



t-SNE of Style Feature (3D). Clustering using Kmeans.

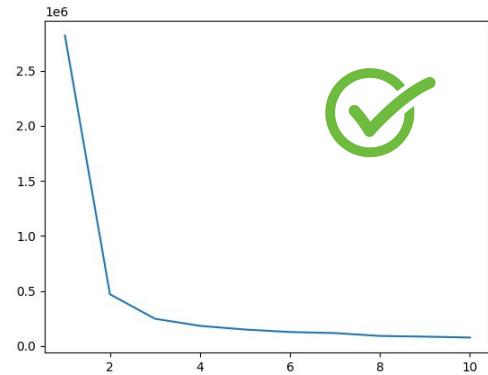
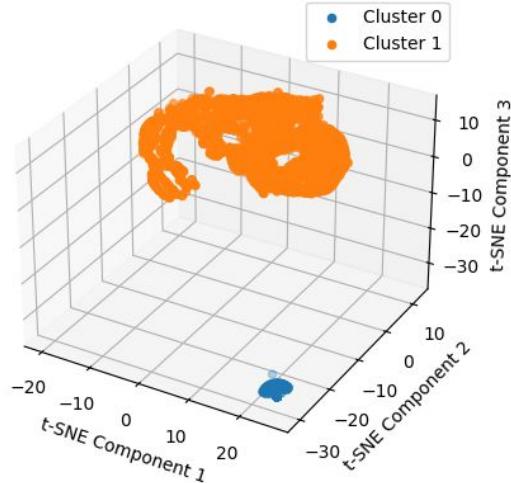


Optimal Cluster Number :

28



t-SNE of Style Feature (3D). Clustering using Kmeans.



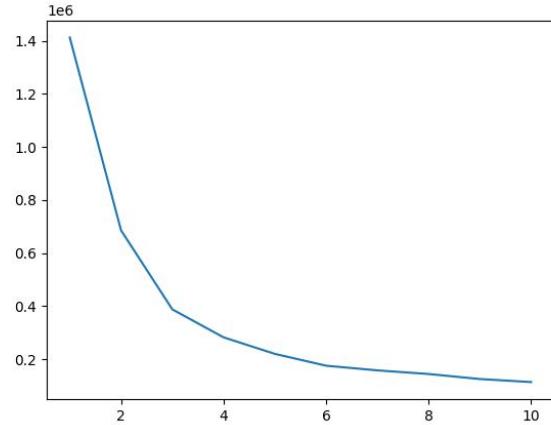
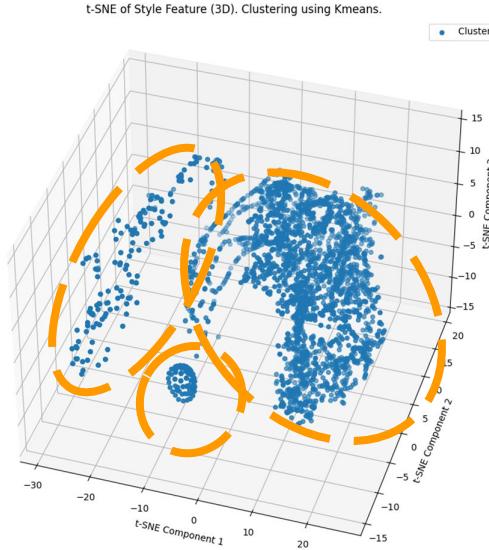
**silhouette score [0.77956665,
0.56250405, 0.51108336, 0.50775415,
0.4993118, 0.4822937, 0.50364006,
0.50337696, 0.48477888]
cluster number 2**



Optimal Cluster Number : Why use Kmeans?

29

With the current code from the paper, the user has to choose the number of clusters. But which is best?



silhouette score [0.53704184, **0.620371**,
0.58846825, 0.5824427, 0.5868706,
0.5799979, 0.57591444, 0.5759357,
0.5797304]
cluster number 3

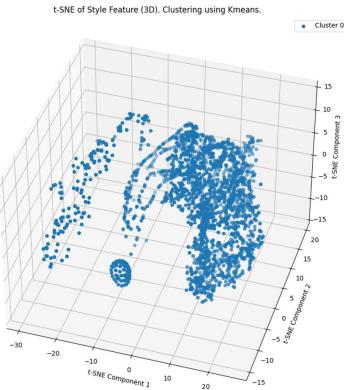


One could think this image has $n=2$ but it's not true.
A visual verification of the t-SNE and both criteria used reveal
we have 3 features!

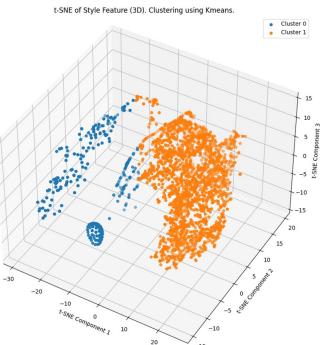
Optimal Cluster Number : Why use Kmeans?

Results for k means and different number of clusters

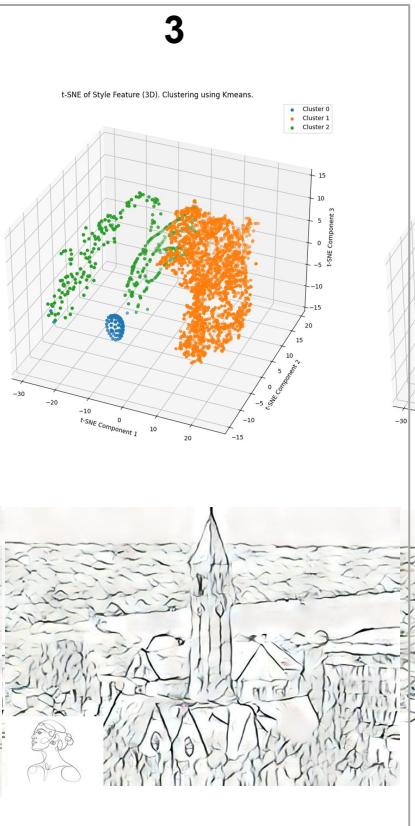
1



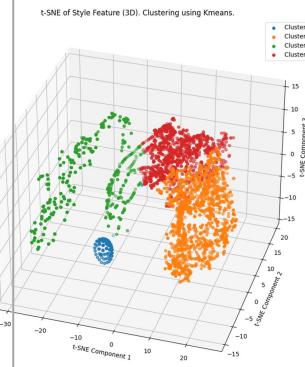
2



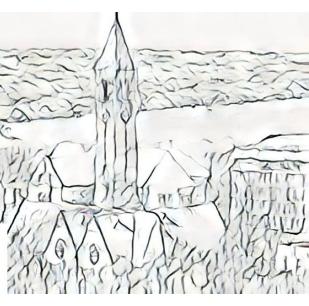
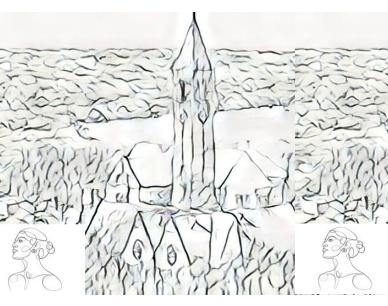
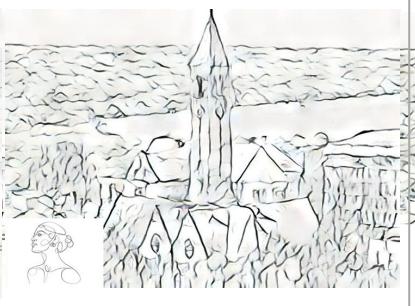
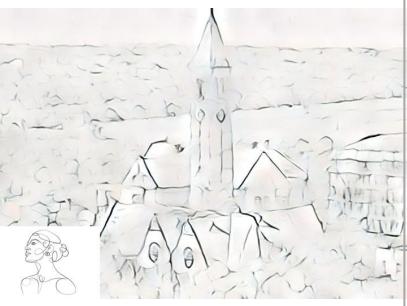
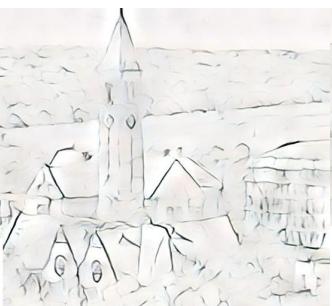
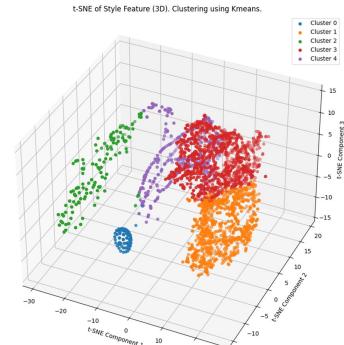
3



4



5

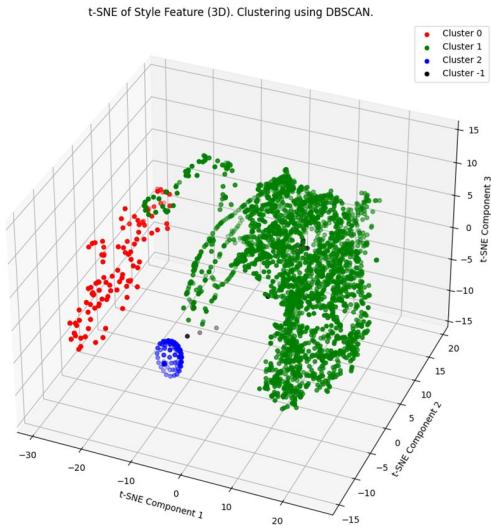


Optimal Cluster Number : DBSCAN

31

With the current code from the paper, the user has to choose the number of clusters. But which is best? We have to visually compare and intuitively guess...

Solution! Use DBSCAN and this will be done automatically!



Automatically found the 3 intuitive clusters!

Optimal Cluster Number :

32

Custom Criterion:

Content Image



Style Image



Model Output



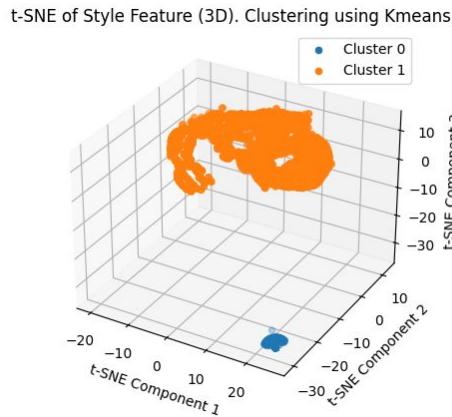
What happens if we use the same style and content image ?

Optimal Cluster Number :

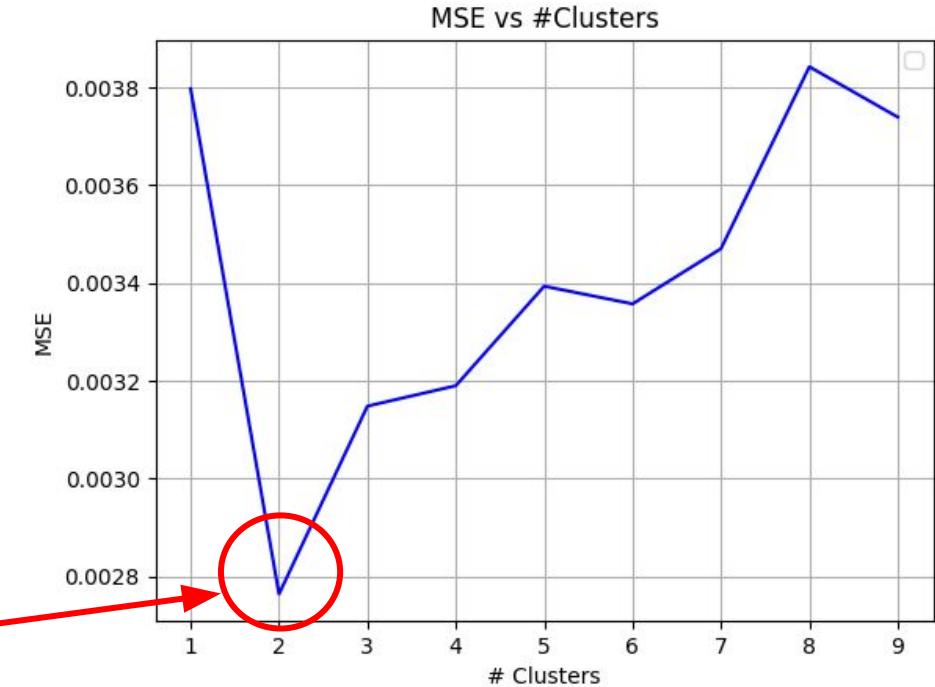
33

Custom Criterion:

We use the style as content and style for different cluster numbers.



We chose the number of clusters that minimizes the mse.



Optimal Cluster Number :

34

Content image



Style image



Optimal Cluster Number :

35



k=1



k=2



k=3



k=4



k=5



k=6



k=7



k=8



k=9



k=10

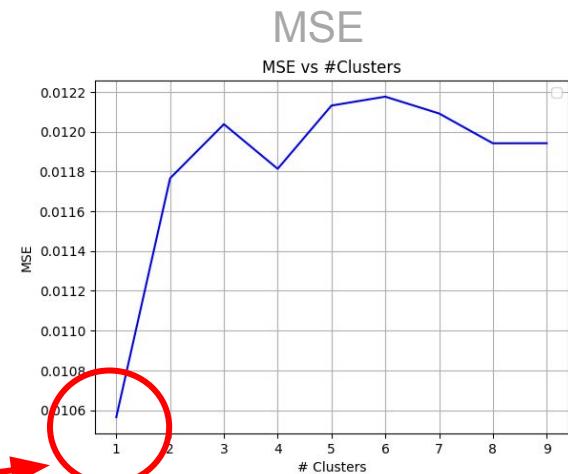
Optimal Cluster Number :

36

Content image



Style image



Optimal Cluster Number :

37



k=1

k=2

k=3

k=4

k=5



k=6

k=7

k=8

k=9

k=10

Optimal Cluster Number :

38

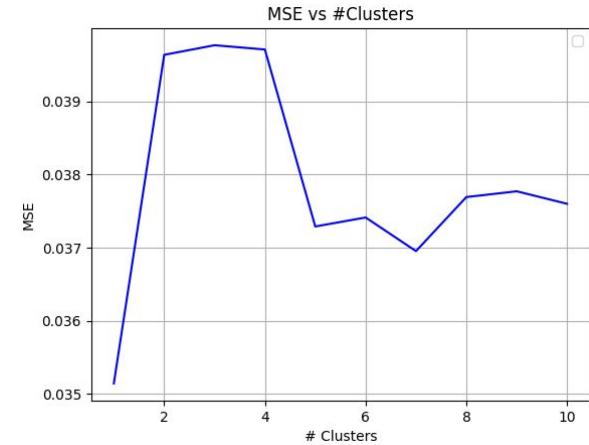
Content image



Style image



MSE



Optimal Cluster Number :

39



k=1

k=2

k=3

k=4

k=5



k=6

k=7

k=8

k=9

k=10

Thank you for your attention