Real Time Arbitrary Video Style Transfer

Outline

1. Problem Overview
2. Original Model Architecture
3. Proposed Model Architecture
4. Learning Process
5. Experiments
6. Future Work
7. References

1.

Problem Overview

Video Style Transfer

Video Style Transfer

Transferring a *Style* extracted from an image into a whole video sequence

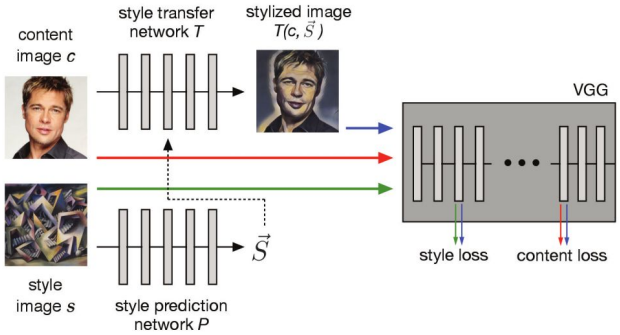
2.

Original Architecture

Arbitrary Style Transfer

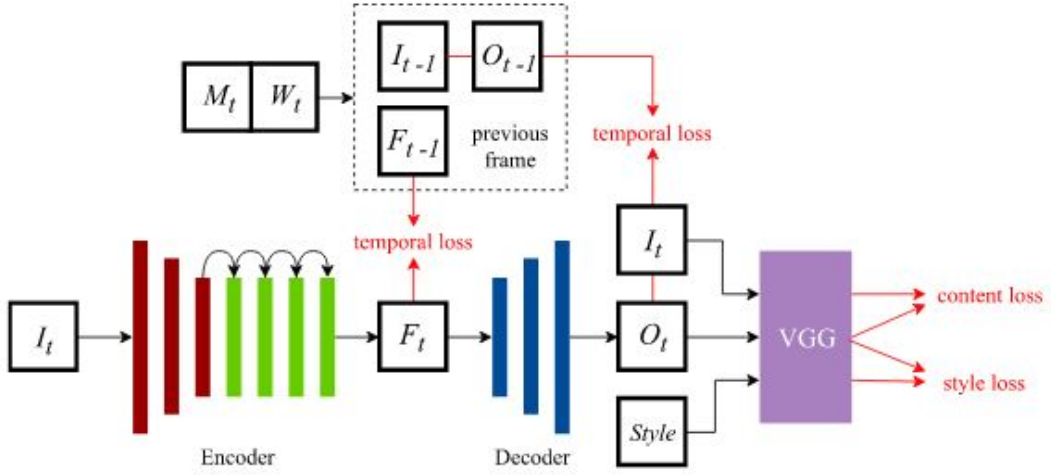
The architecture comprises of two sub networks

- The Style Prediction Network
- The Image Transformation Network



ReCoNet Video Style Transfer

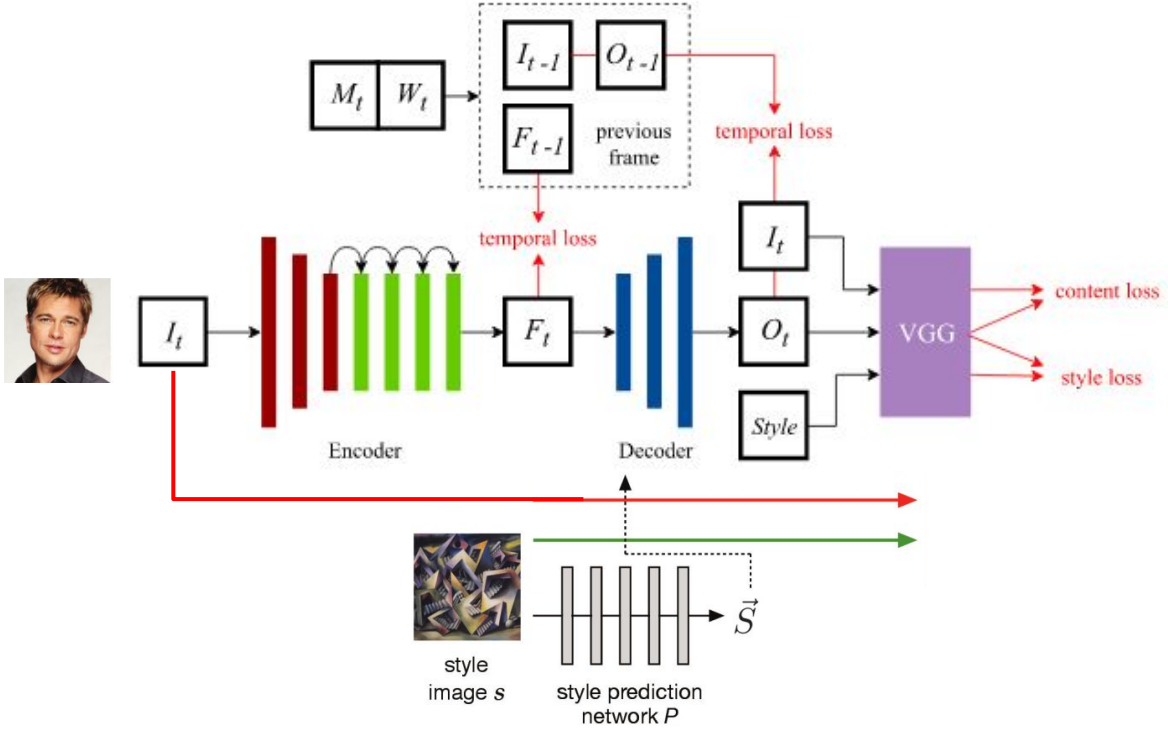
Uses multiple level temporal losses at the output and at the input



3.

Proposed Model Architecture

Proposed Model



Proposed Model

Loss Function for two-frame learning

$$\mathcal{L}(t-1,t) = \sum_{i\in\{t-1,t\}} (\alpha\mathcal{L}_{content}(i) + \beta\mathcal{L}_{style}(i) + \gamma\mathcal{L}_{tv}(i)) + \lambda_f\mathcal{L}_{temp,f}(t-1,t) + \lambda_o\mathcal{L}_{temp,o}(t-1,t)$$

The Style Network

To produce the vector *S*, the network is fed the style image and produces the vector.

This takes into account the shared features and textures between different styles.

4.

Experiments

Experiments

To solve the real-time arbitrary video style transfer, two approaches are proposed

1. Frame by Frame inference using Arbitrary image
2. Enforcing Temporal Consistency on Image arbitrary style transfer with Low learning rate
3. Enforcing Temporal Consistency on Image arbitrary style transfer with High learning rate

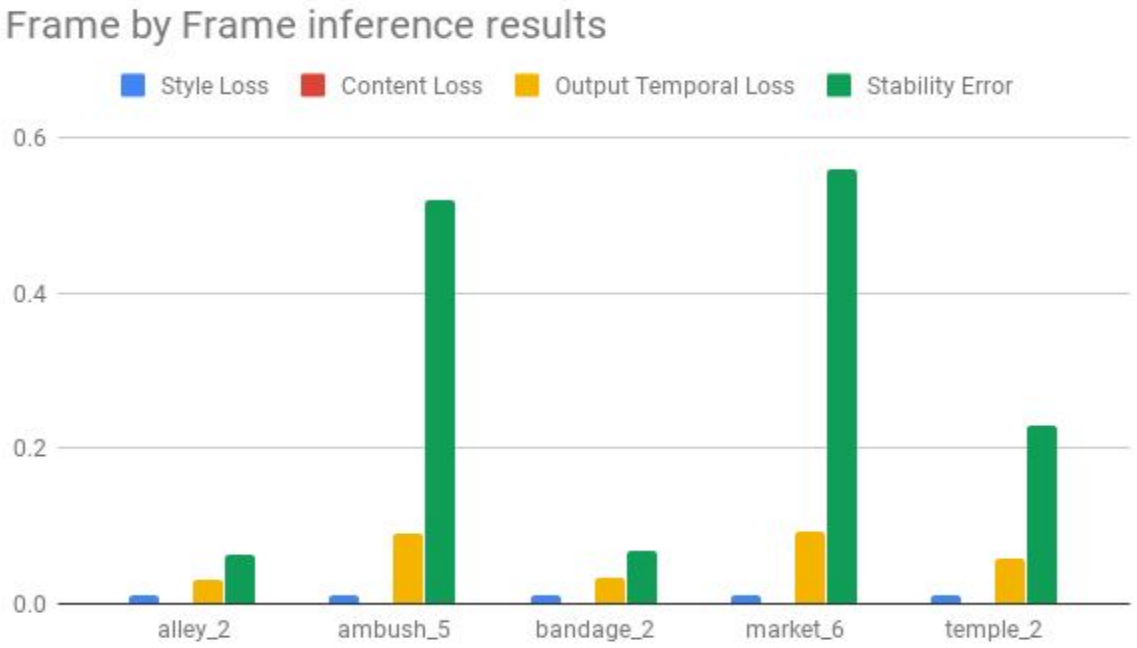
Frame by Frame inference

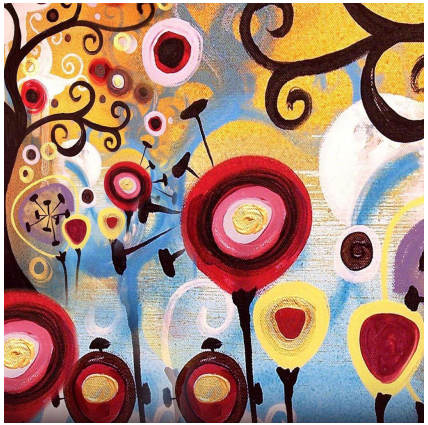
We conjectured that since the network perform styling using forward pass using high-level features

Objects in consecutive frames will have same representation in the output as opposed to methods used by Ruder[1]

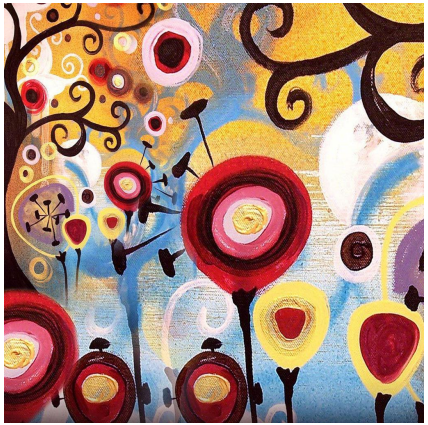
1. A similar concept was discussed by Huang et al [4]

Results

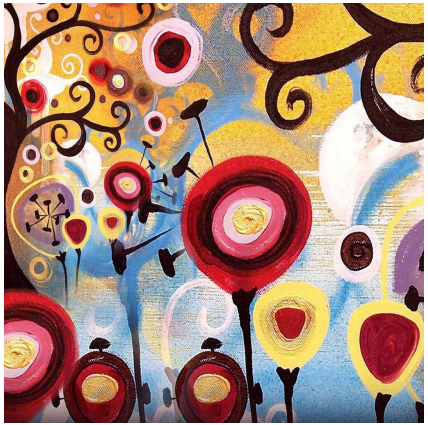

Frame by Frame inference results
Style Loss
Content Loss
Output Temporal Loss
Stability Error
0.6
0.4
0.2
0.0
alley_2
ambush_5
bandage_2
market_6
temple_2



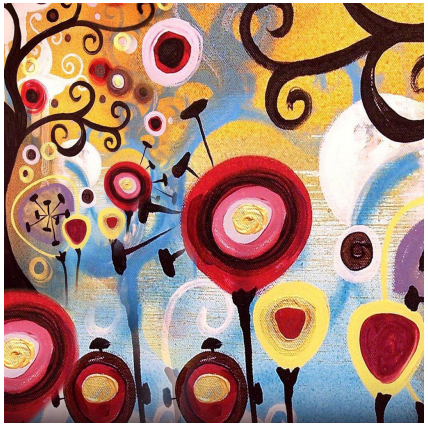
Bandage 2 + Candy



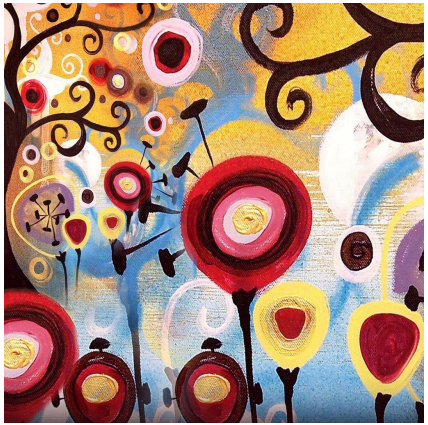
Alley 2 + Candy



Ambush 5 + Candy



Temple 2 + Candy



Market 6 + Candy

Enforcing Temporal Consistency with Low learning rate

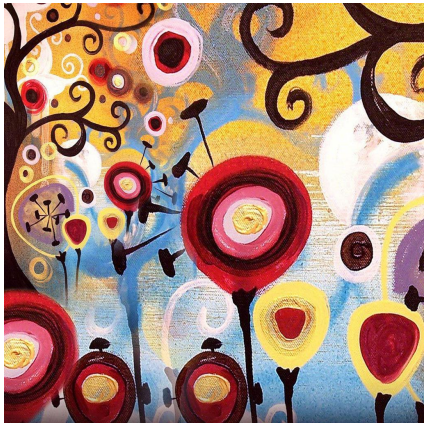
Since the content and style loss are optimized, and temporal loses incurs large losses,

The learning rate was set to a small value 1e-12, as to avoid overshooting

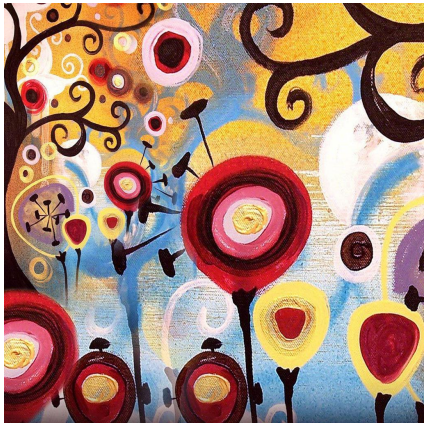
1. A similar concept was discussed by Huang et al [4]

Results

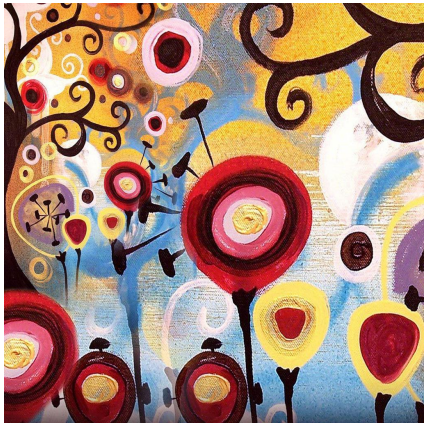

Temporal Consistency with low learning rate
Style Loss
Content Loss
Output Temporal Loss
Stability Error
0.6
0.4
0.2
0.0
alley_2
ambush_5
bandage_2
market_6
temple_2



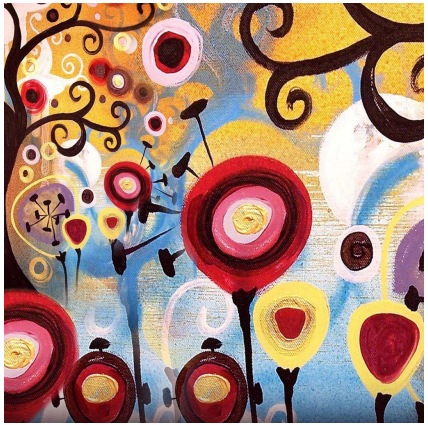
Bandage 2 + Candy



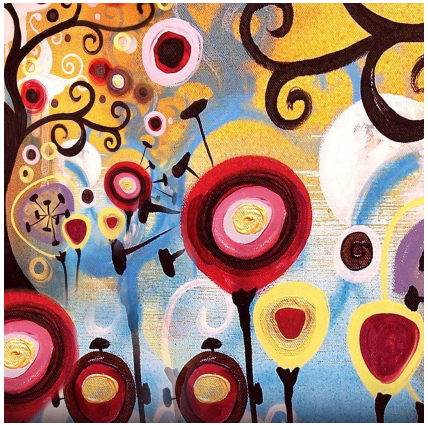
Alley 2 + Candy



Ambush 5 + Candy



Temple 2 + Candy



Market 6 + Candy

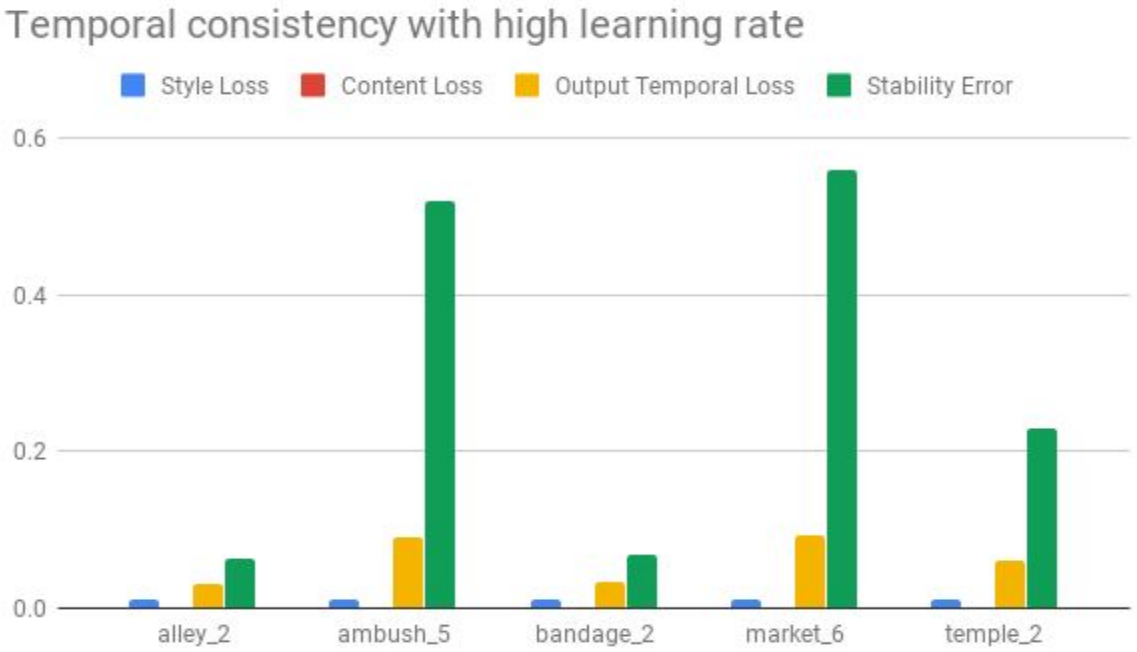
Enforcing Temporal Consistency with high learning rate

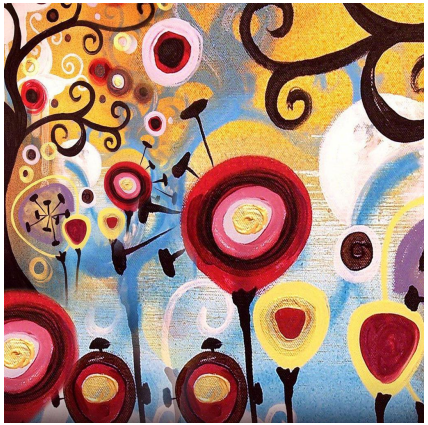
The weights didn’t change much with low learning rate,

Learning rate was set to 1e-5 and relative weights in loss function were adjusted

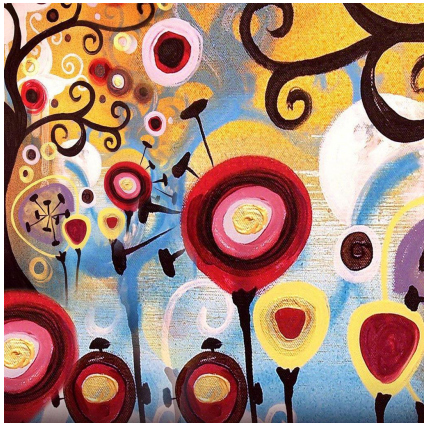
1. A similar concept was discussed by Huang et al [4]

Results

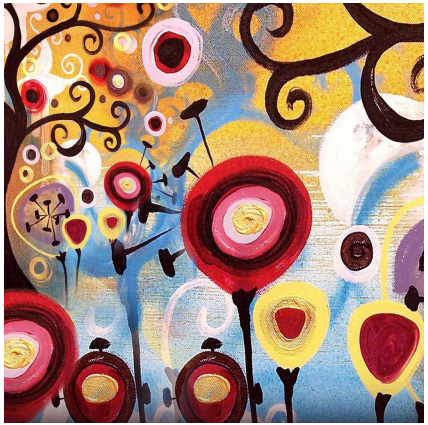

Temporal consistency with high learning rate
Style Loss
Content Loss
Output Temporal Loss
Stability Error
0.6
0.4
0.2
0.0
alley_2
ambush_5
bandage_2
market_6
temple_2



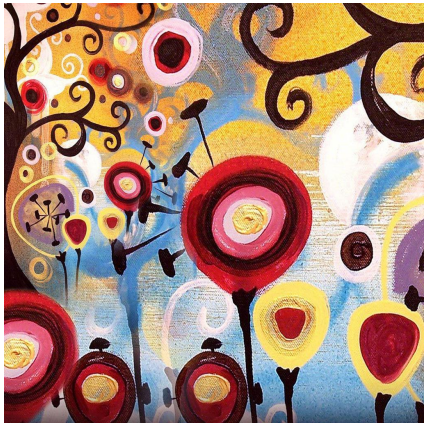
Bandage 2 + Candy



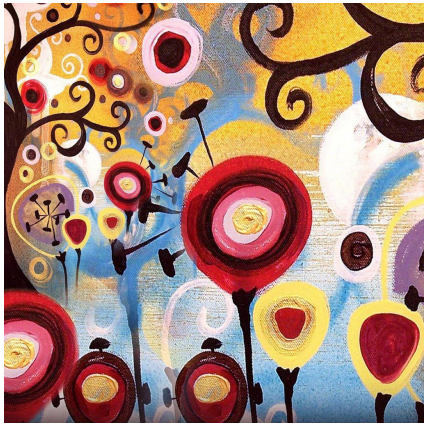
Alley 2 + Candy



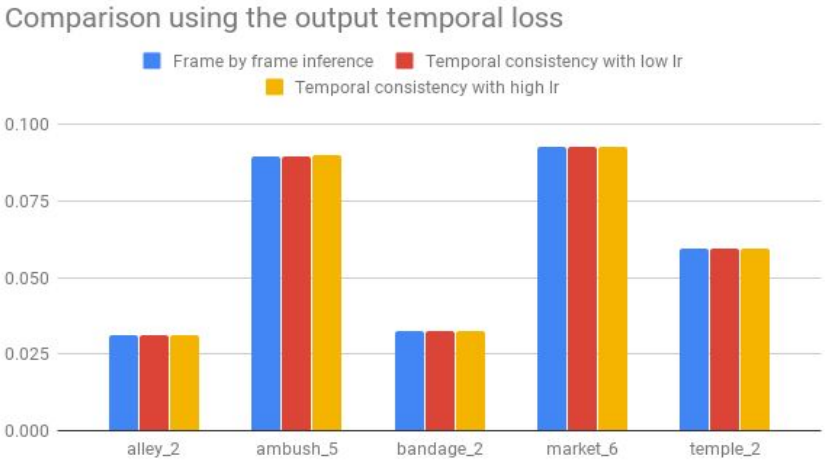
Ambush 5 + Candy

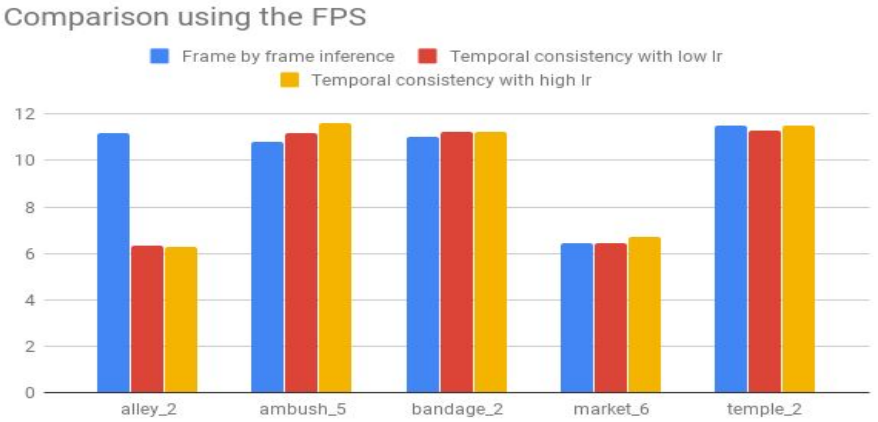


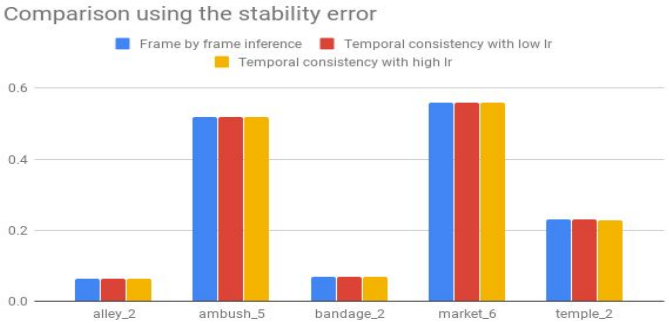
Temple 2 + Candy



Market 6 + Candy







Comparison between the different experiments.

5.

Observations and Conclusions

Validity of Estab

While the Estab is used extensively in literature, it didn’t produce accurate results

The objects were styled in the same manner across frames, same textures and colors. However, color strokes may get smaller or larger.

While this gives the video a painting feeling, it was penalized by estab

Decolorization and Smudging

To optimize the temporal loss across frames,

The model began to remove color, and smudges content to be fixed across frames

This could be fixed probably with proper hyper-tuning, but was not done due to lack of computational resources

Conclusion

- The model has low style & content loss - as expected
- Although the temporal loss is low, it’s still one order of magnitude higher.
- Estab was not improved by much, but the results are visually appealing.

6.

Future Work

Future Work

- Hyper-parameter tuning
- Export to TF-lite
- Find a new temporal stability metric
- Using a long-term, short-term temporal loss

7.

References

References

- [1] L. A. Gatys, et al., “A neural algorithm of artistic style,” 2015.
- [2] J. Johnson, A. Alahi, et al., “Perceptual losses for real-time style transfer and super-resolution,” 2016.
- [3] M. Ruder, et al., “Artistic style transfer for videos,”
- [4] H. Huang, et al., “Real-time neural style transfer for videos,”, July 2017.
- [5] C. Gao, et al., “Reconet: Real-time coherent video style transfer network,” 2018.
- [6] G. Ghiasi, et al., “Exploring the structure of a real-time, arbitrary neural artistic stylization network,” 2017

Thanks!

Any questions?

Quotations are commonly printed as a means of inspiration and to invoke philosophical thoughts from the reader.

This is a slide title

- Here you have a list of items
- And some text
- But remember not to overload your slides with content

Your audience will listen to you or read the content, but won’t do both.

Big concept

Bring the attention of your audience over a key concept using icons or illustrations

You can also split your content

White

Is the color of milk and fresh snow, the color produced by the combination of all the colors of the visible spectrum.

Black

Is the color of coal, ebony, and of outer space. It is the darkest color, the result of the absence of or complete absorption of light.

In two or three columns

Yellow

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

Blue

Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

Red

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.

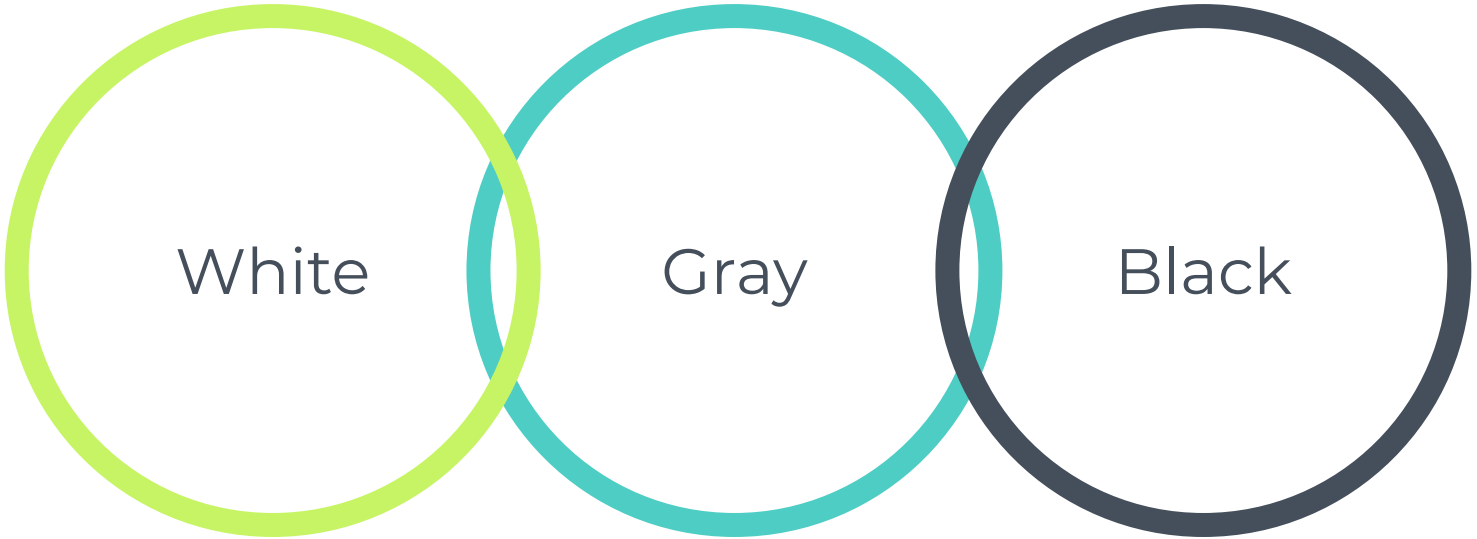
A picture is worth a thousand words

A complex idea can be conveyed with just a single still image, namely making it possible to absorb large amounts of data quickly.

Want big impact?

Use big image.

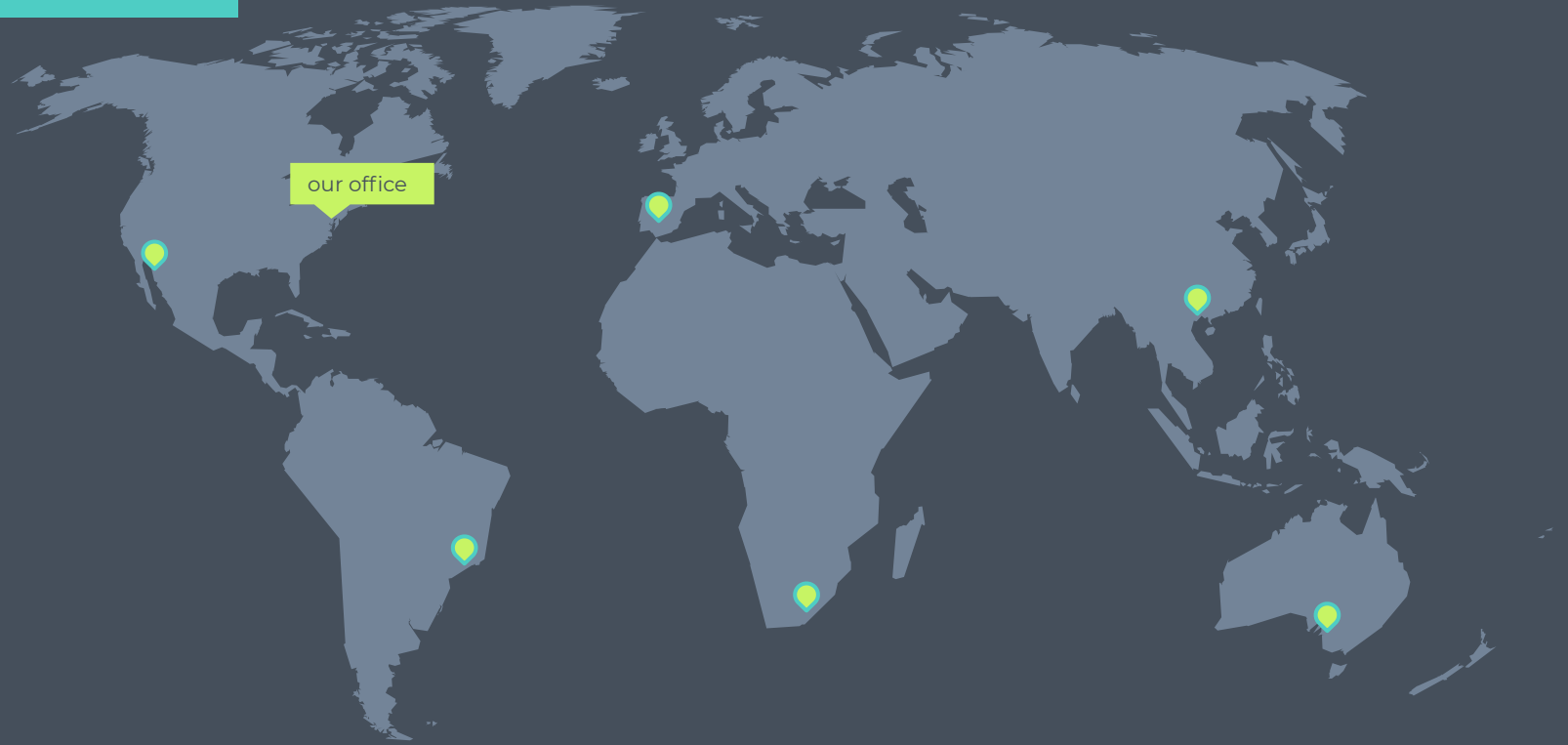
Use charts to explain your ideas



And tables to compare data

	A	B	C
Yellow	**10**	**20**	**7**
Blue	**30**	**15**	**10**
Orange	**5**	**24**	**16**

Maps



89,526,124

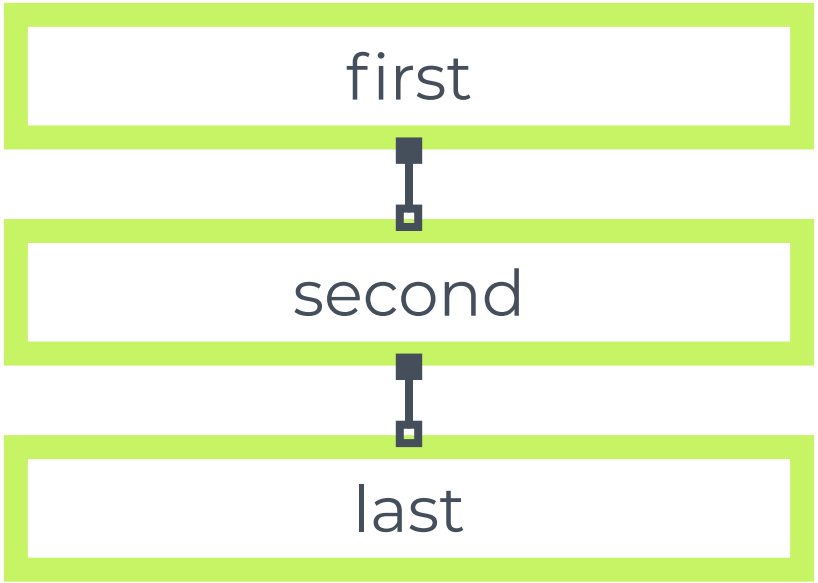
Whoa! That’s a big number, aren’t you proud?

89,526,124$
That’s a lot of money

185,244 users
And a lot of users

100%
Total success!

Our process is easy



Let’s review some concepts

Yellow
Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

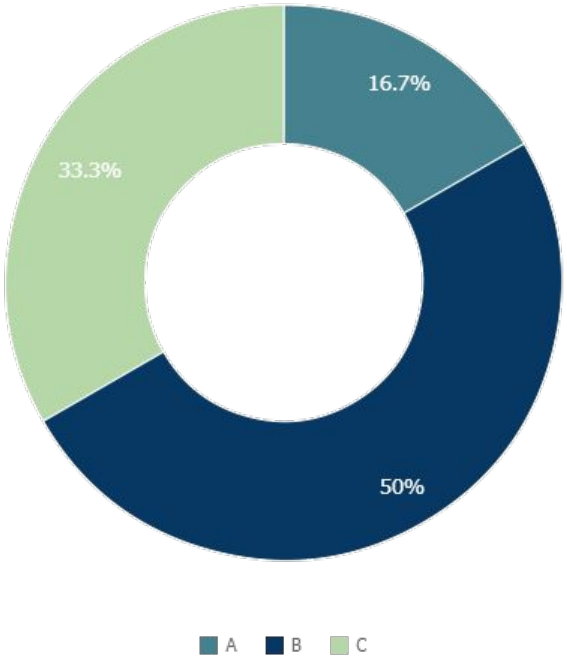
Blue
Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

Red
Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.

Yellow
Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

Blue
Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

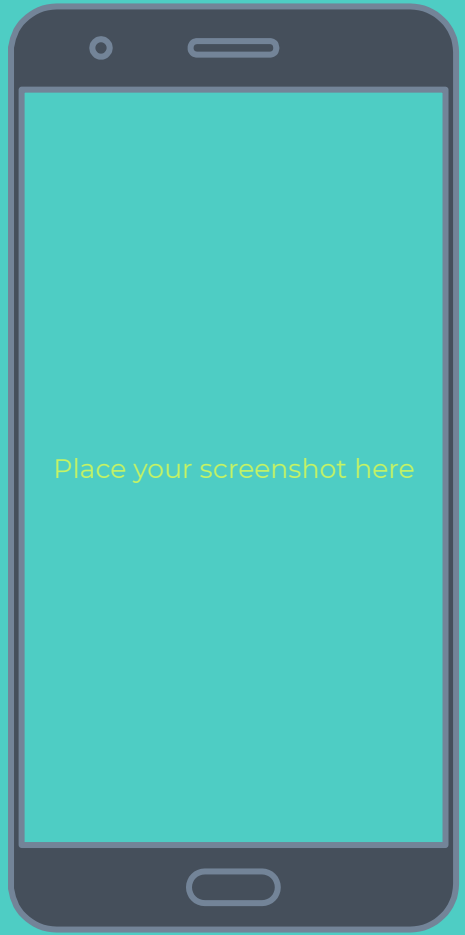
Red
Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.



You can copy&paste graphs from Google Sheets

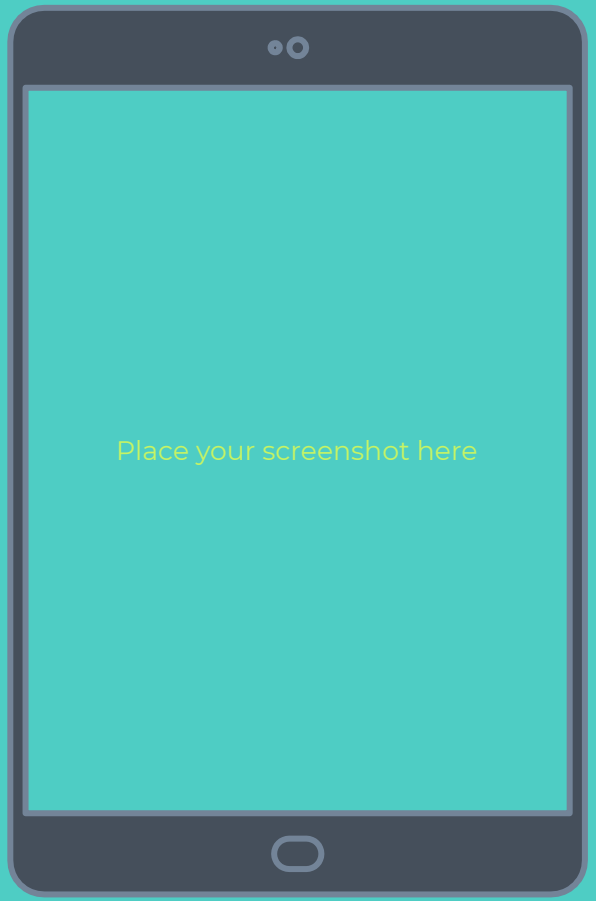
Mobile project

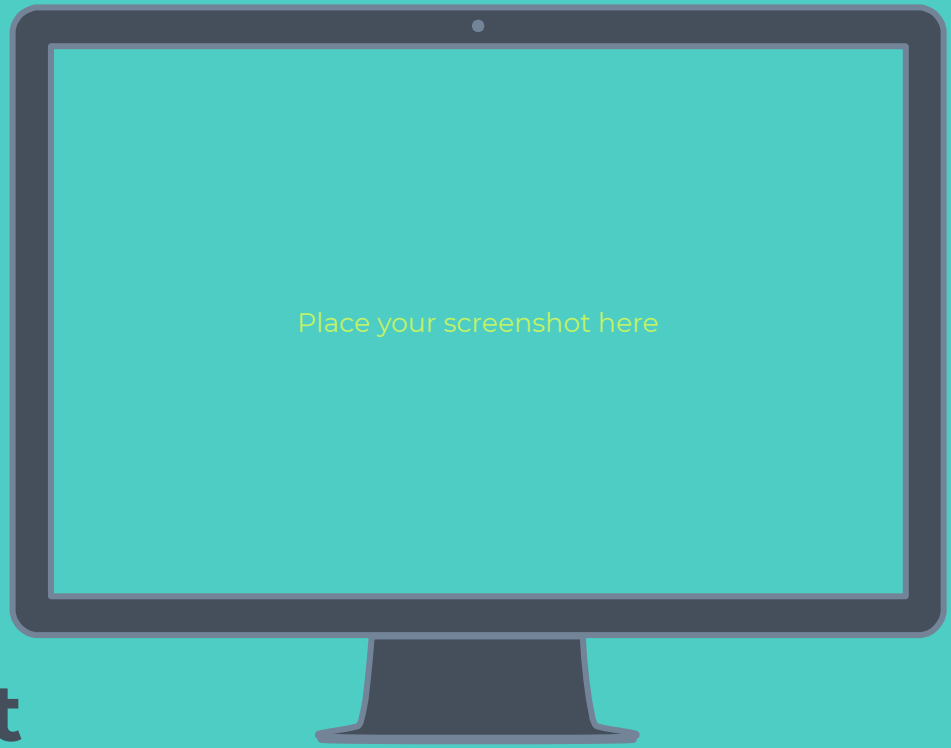
Show and explain your web, app or software projects using these gadget templates.



Tablet project

Show and explain your web, app or software projects using these gadget templates.





Desktop project

Show and explain your web, app or software projects using these gadget templates.

Presentation design

This presentations uses the following typographies and colors:

- Titles & Body copy: **Montserrat**

You can download the fonts on this page:

https://www.fontsquirrel.com/fonts/montserrat

- Grey **#454f5b**
- Light grey **#738498**
- Neon green **#c7f464**
- Aqua **#4ecdc4**

You don’t need to keep this slide in your presentation. It’s only here to serve you as a design guide if you need to create new slides or download the fonts to edit the presentation in PowerPoint®

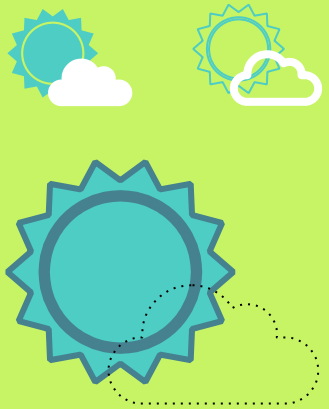
SlidesCarnival icons are editable shapes.

This means that you can:

- Resize them without losing quality.
- Change fill color and opacity.
- Change line color, width and style.

Isn't that nice? :)

Examples:



Now you can use any emoji as an icon!
And of course it resizes without losing quality and you can change the color.

How? Follow Google instructions
https://twitter.com/googledocs/status/730087240156643328

and many more...

Slides Carnival

Free templates for all your presentation needs

For PowerPoint and Google Slides

100% free for personal or commercial use

Ready to use, professional and customizable

Blow your audience away with attractive visuals

Instructions for use

EDIT IN GOOGLE SLIDES

Click on the button under the presentation preview that says "Use as Google Slides Theme".

You will get a copy of this document on your Google Drive and will be able to edit, add or delete slides.

You have to be signed in to your Google account.

EDIT IN POWERPOINT®

Click on the button under the presentation preview that says "Download as PowerPoint template". You will get a .pptx file that you can edit in PowerPoint.

Remember to download and install the fonts used in this presentation (you'll find the links to the font files needed in the Presentation design slide)

More info on how to use this template at www.slidescarnival.com/help-use-presentation-template

Hello!

I am Jayden Smith

I am here because I love to give presentations.
You can find me at @username