

Real Time Video Arbitrary Style Transfer

Outline

1. Problem Overview
2. Related Work
3. Need To Extend
4. Proposed Approach
5. Progress
6. Next Steps
7. Future Work
8. References

1.

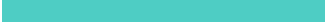
Problem Overview

Video Style
Transfer

Video Style Transfer

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





Each frame is processed and a new frame is synthesized with the content of the frame and the style of the image

2.

Related Word

Image Style Transfer, Gatys et al [1]

Gatys formalized the problem as Energy optimization problem

Trying to minimize a linear combination of Content Loss and Style Loss

Image Style Transfer, Gatys et al [1]

Content loss is defined to be the difference between the Feature Maps of the original image and the output image

$$\mathcal{L}_{content}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2 .$$

Image Style Transfer, Gatys et al [1]

Style loss is defined to be the difference between the Gram Matrices of the original image and the style image

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2$$

$$\mathcal{L}_{style}(\vec{a}, \vec{x}) = \sum_{l=0}^L w_l E_l$$

Image Style Transfer, Gatys et al [1]

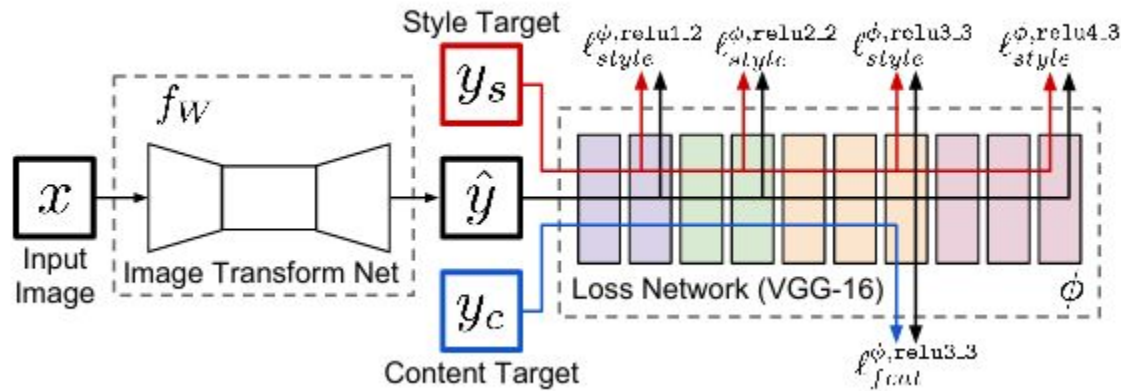
Impressive Results

Very Slow - Heavy Computations



Perceptual Losses, Johnson et al [2]

Instead of Computing results at layer l , use high level features on a pre-trained ImageNet Network



Perceptual Losses, Johnson et al [2]

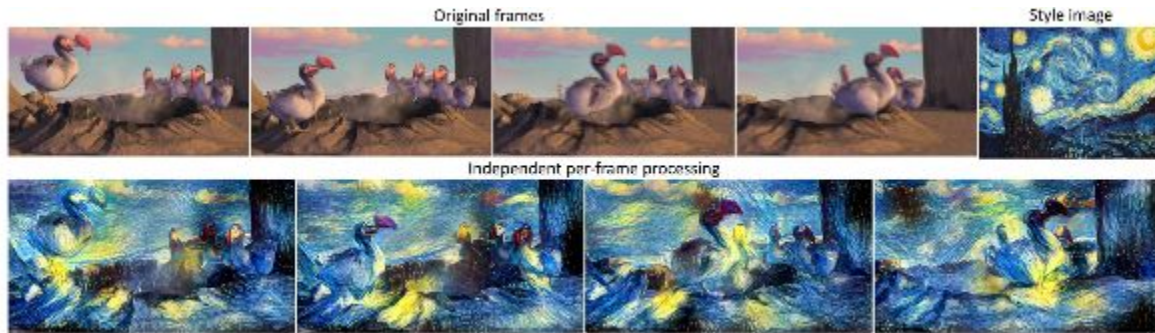
Comparable Qualitative Results
Three Order of Magnitude faster



Artistic Style Transfer, Ruder et al [3]

Stylizing frame by frame leads to flickering effect

Because each frame converges to a different local optima



Artistic Style Transfer, Ruder et al [3]

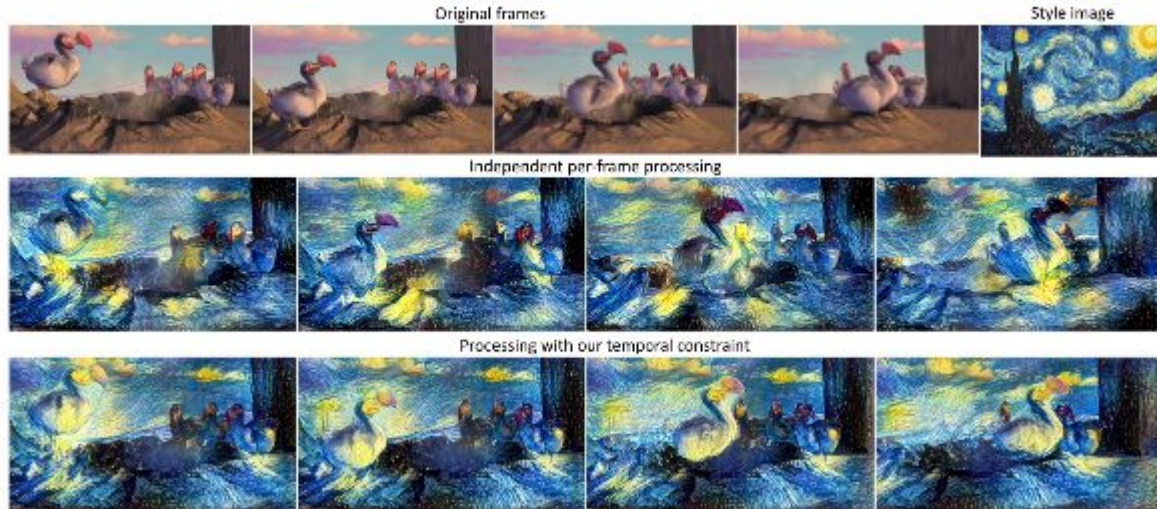
Solved by penalizing temporal inconsistencies

Utilizing optical flow warping to estimate each frame from the frame preceding it

Each frame is initialized to the warp of the previous frame

Artistic Style Transfer, Ruder et al [3]

As with Gatys, impressive results
Slow run time



Real time Neural Transfer, Huang et al [4]

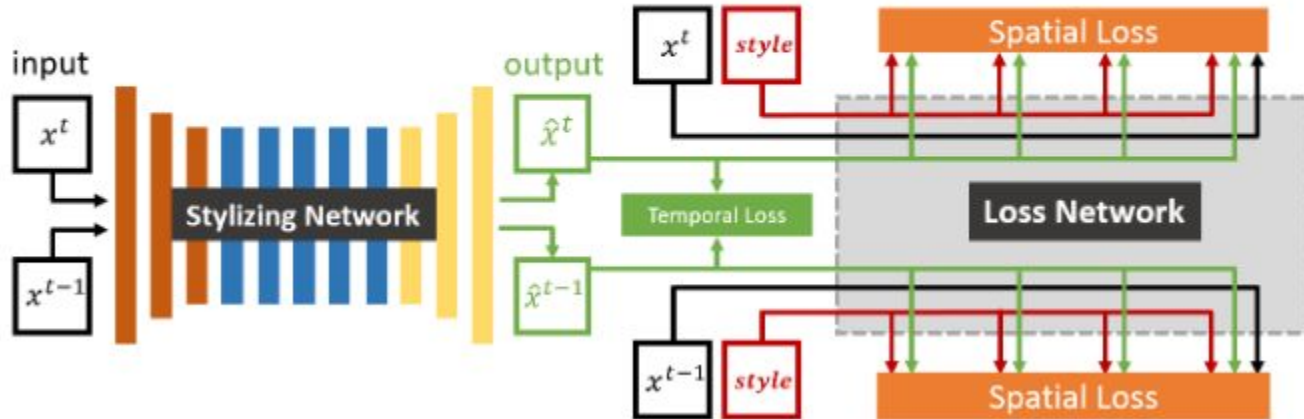
Uses a pre-trained style network for fast inference

The style network is an auto-encoder feed forward network

Minimizes temporal loss

Real time Neural Transfer, Huang et al [4]

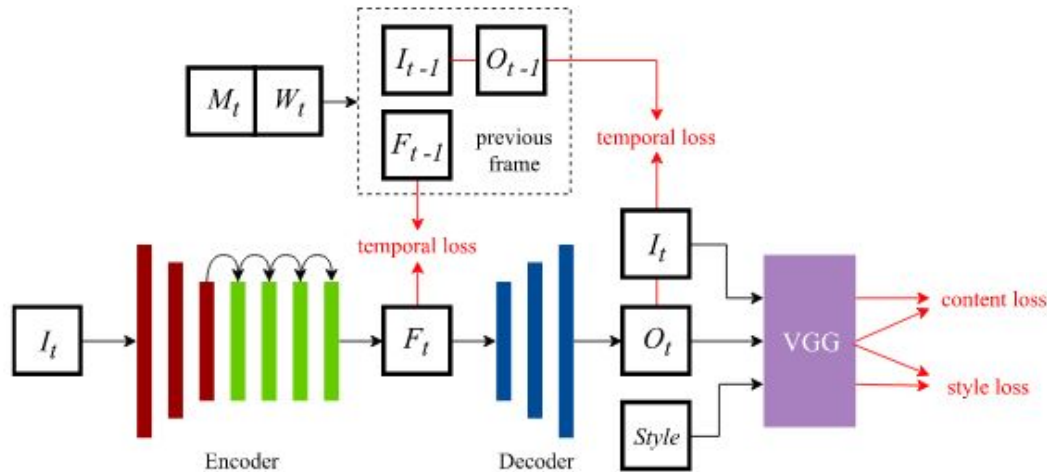
Same as Johnson, with the temporal loss added



Real time Neural Transfer, Gao et al [5]

Use ground truth optical flows

Takes into account illuminance instability



Real time Neural Transfer, Gao et al [5]

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

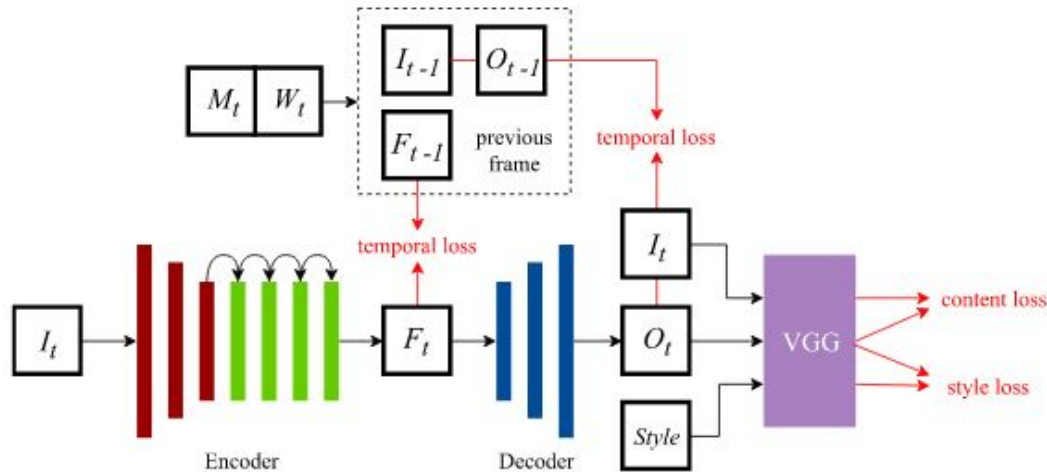


Image Arbitrary Style Transfer, Ghiasi [6]

Real time styling for an arbitrary styles

Utilizes a style prediction network (style embedding)

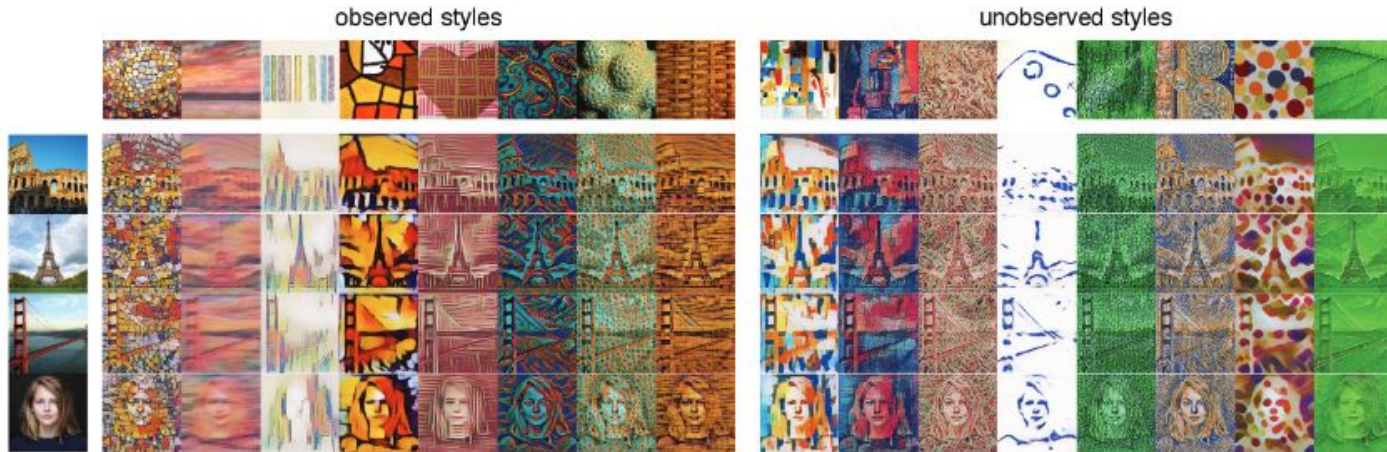
- Given an image style it embeds in vector space

- Uses a transformer model which uses the vector to transform its activation layers to produce different styles

At inference, one pass at feed forward network to produce stylized output

Image Arbitrary Style Transfer, Ghiasi [6]

Using 80,000 painting to train the style network
Comparable results on unobserved styles



3.

Need to extend

Need to extend

Current real time models use pre-trained networks for each style, to perform fast computations

Arbitrary Style networks take long run times

No current model can produce real-time results for video style transfer for an arbitrary styles



New Milestone

This represents a new milestone in the problem of Artistic video style transfer.

4.

Proposed Approach

Two approaches

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

1. Frame by Frame inference
2. Enforcing Temporal Consistency on Image arbitrary style transfer

Frame by Frame inference



This approach simply styles each frame independently using the arbitrary style image transfer network

The intuition is the network is already trained on high level features. Then, using forward passes will yield similar results for the same object across different frames.¹

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

Enforcing Temporal Consistency on Image arbitrary style transfer

Using knowledge transfer, the image arbitrary style network is trained to minimize a new objective function initialized to published weights

The training process is adopted from that of the ReCoNet in [5]

Enforcing Temporal Consistency on Image arbitrary style transfer

This forces the feed forward network to learn temporal consistencies using the **two frame training synergic**

Pair of consecutive frames are passed to the network and temporal losses are computed after using ground truth optical flow.

Enforcing Temporal Consistency on Image arbitrary style transfer

The network starts off while it's already minimized the style and content loss

It then tries to minimize temporal loss while keeping content and style loss minimal

The style embedding network is frozen while the transformer model is trained

Enforcing Temporal Consistency on Image arbitrary style transfer

Care must be taken with training examples to train variety of painting styles as was recommended by Ghiasi in [6]

5.

Progress

Approach #1

- Image Arbitrary Style model code import
- Preliminary tests on frame-by-frame independant Styles
- Qualitative analysis shows it's promising
- Colab notebook [here](#)

Approach #1

- Styling Sintel example with Starry Night Style



6.

Next Steps

Approach #1

- ▣ Frame size and resolution are yet to be configured
- ▣ Quantitative analysis is yet to be done

Approach #2

- Proper fork of code
- Implement losses as described in [5]
- Evaluate over the same datasets using the same benchmarks

7.

Future work

Future work

- Enforcing long term consistency as described in [3]

8.

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

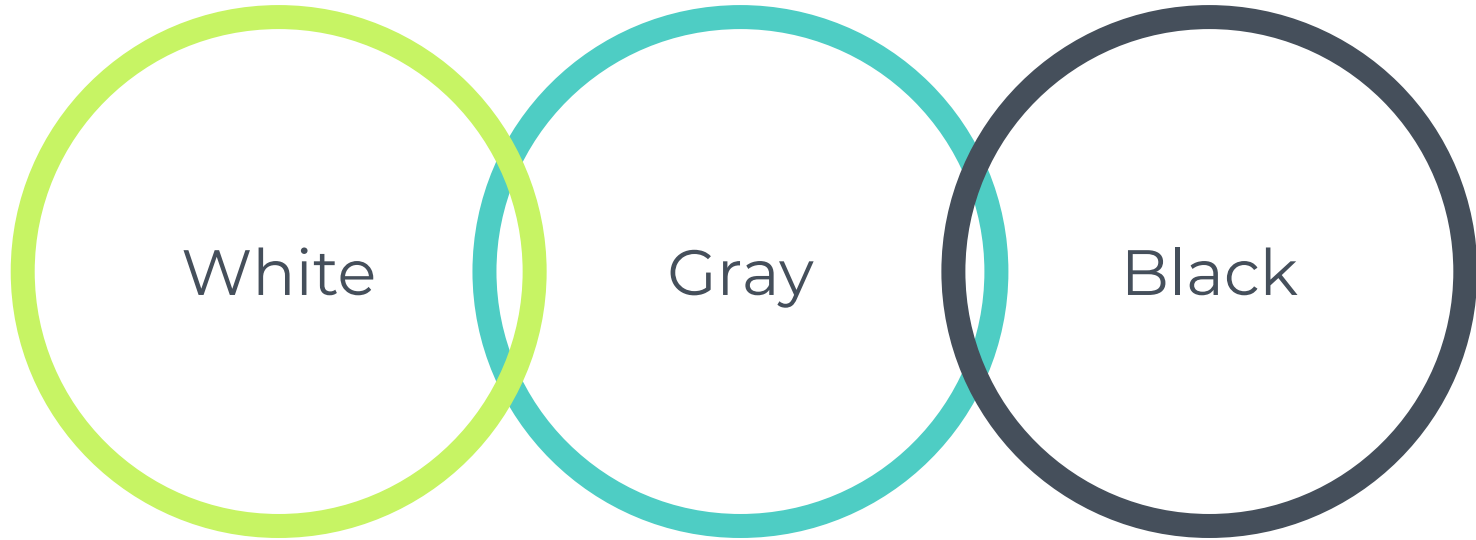
A complex idea can be conveyed with just a single still image, namely making it possible to absorb large amounts of data quickly.





Want big impact?
Use big image.

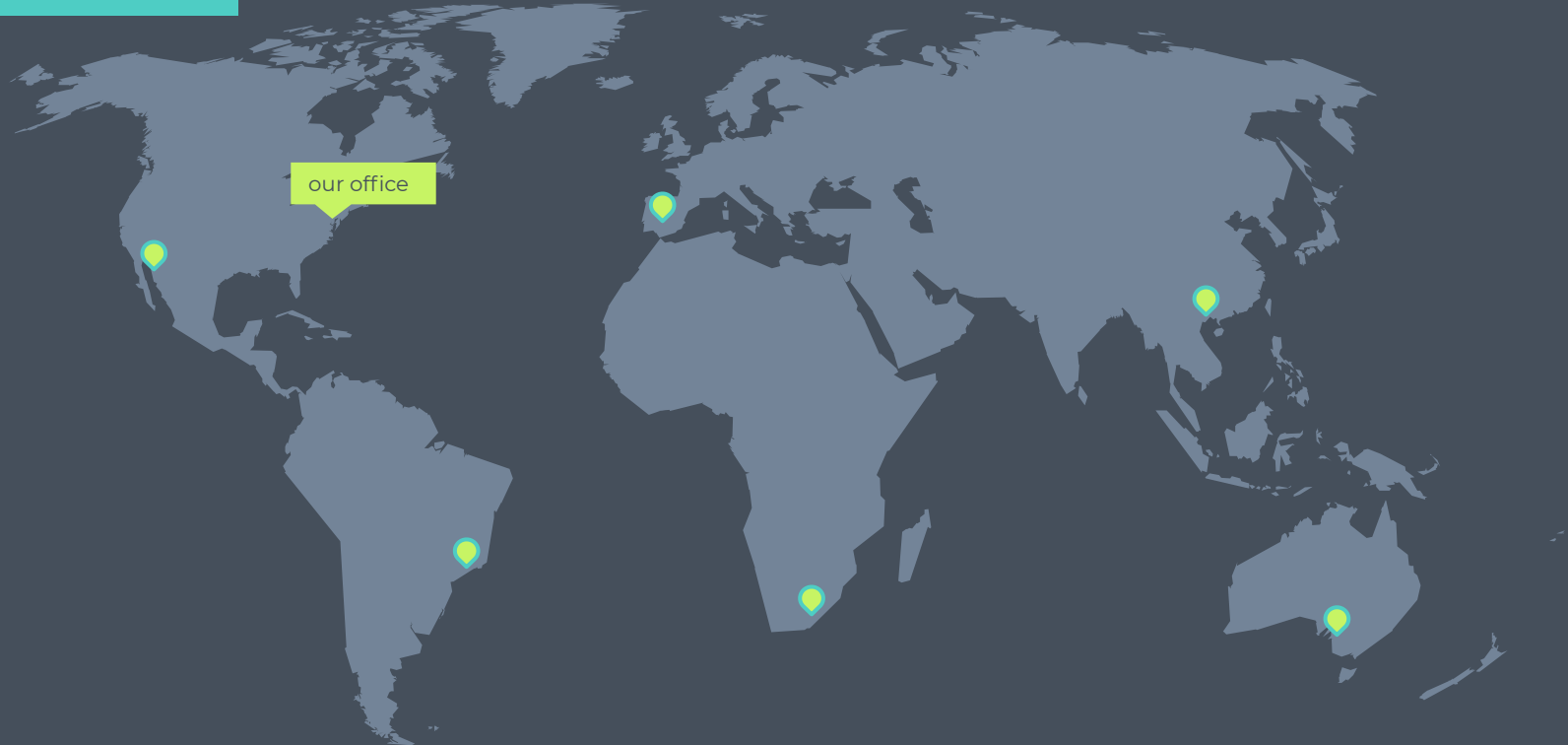
Use charts to explain your ideas



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

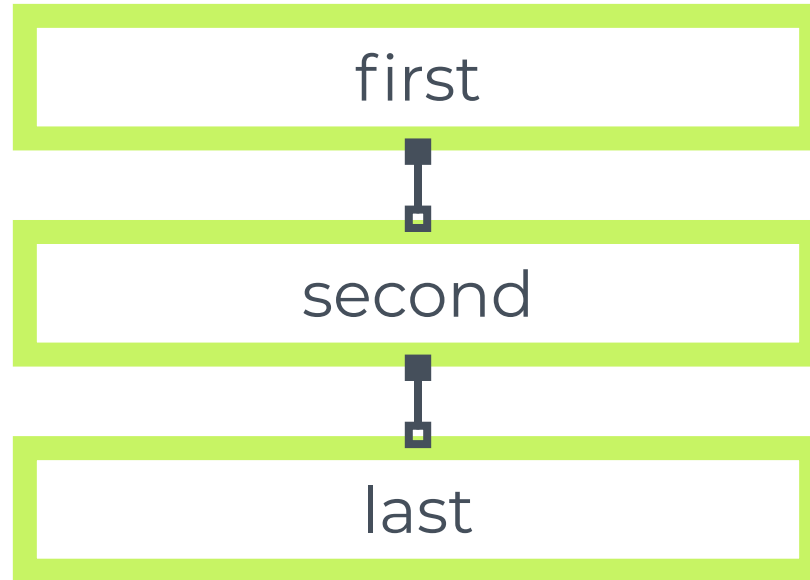
185,244 users

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.



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.



Blue

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



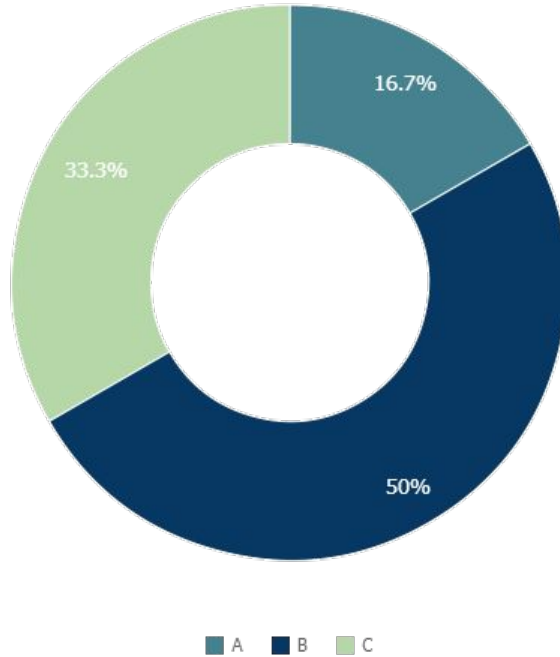
Red

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



Red

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

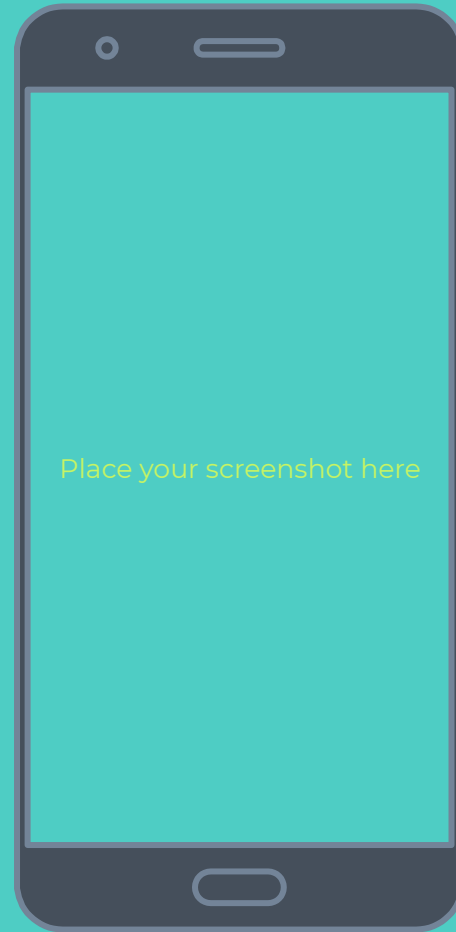


You can copy&paste graphs from [Google Sheets](#)



Mobile project

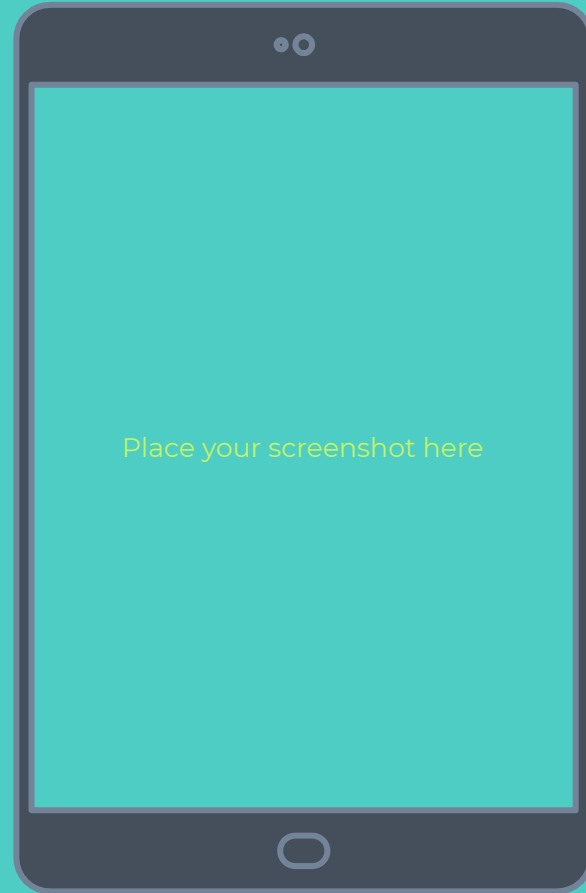
Show and explain your web, app or software projects using these gadget templates.





Tablet project

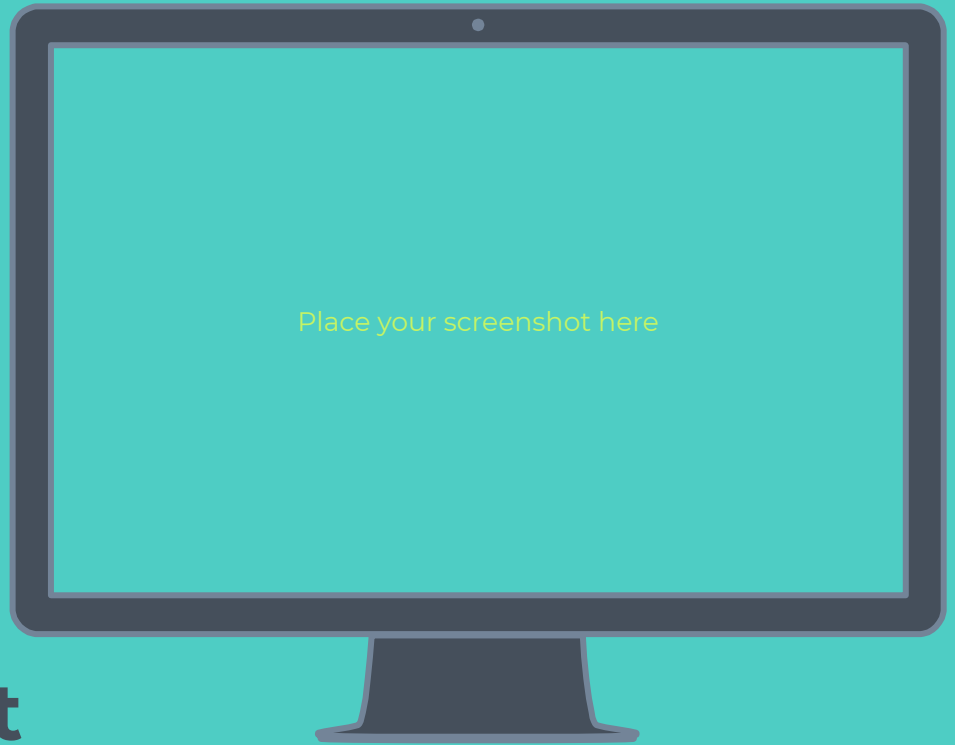
Show and explain your web, app or software projects using these gadget templates.





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:

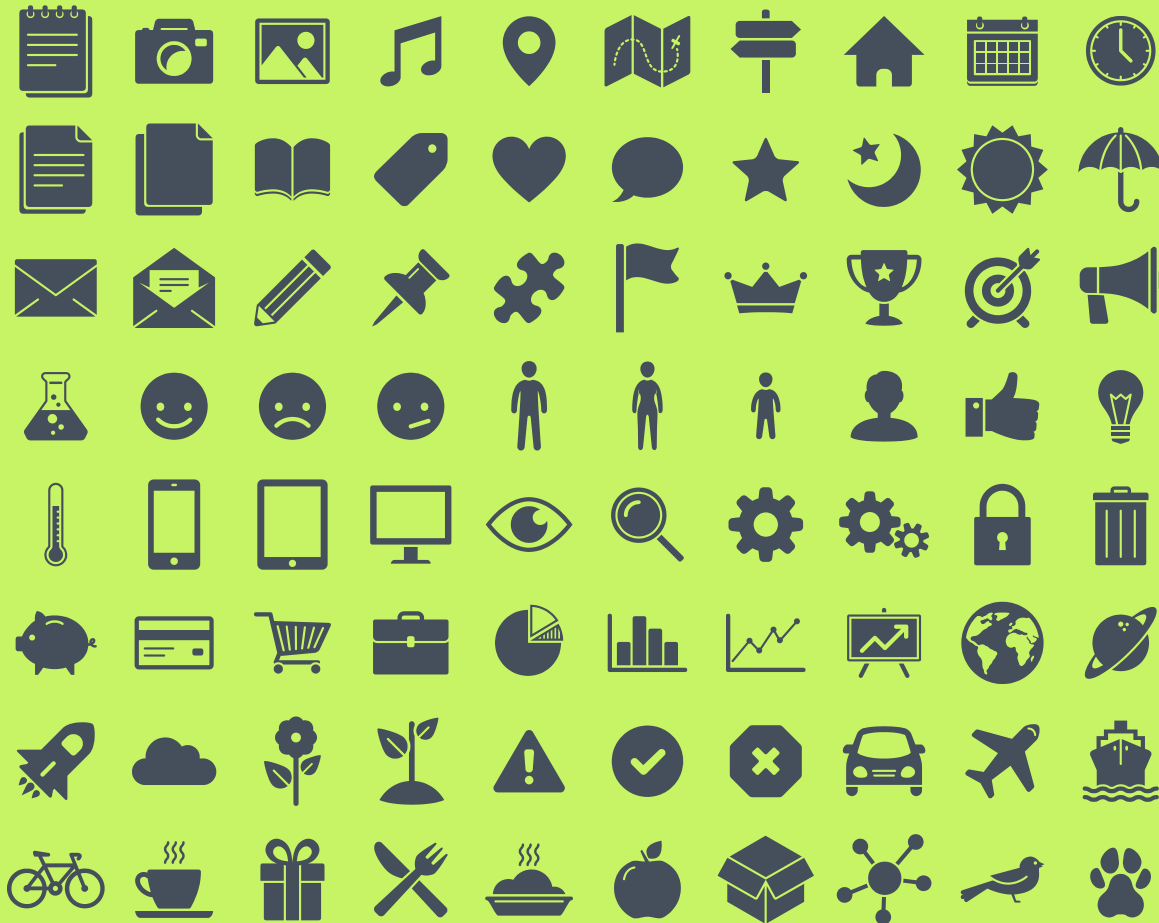
- ▣ Titles & Body copy: **Montserrat**

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®



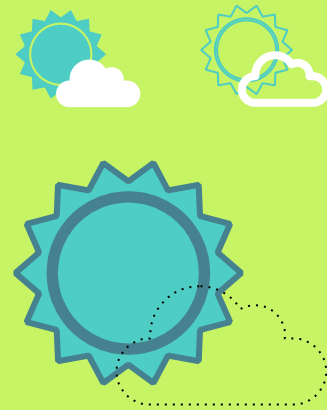
SlidesCarnival icons are editable shapes.

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...



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

This template is free to use under [Creative Commons Attribution license](#). You can keep the Credits slide or mention SlidesCarnival and other resources used in a slide footer.

Hello!

I am Jayden Smith

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