

Leistungsanalyse der Funktion von Middleboxes

Simon Sternsdorf

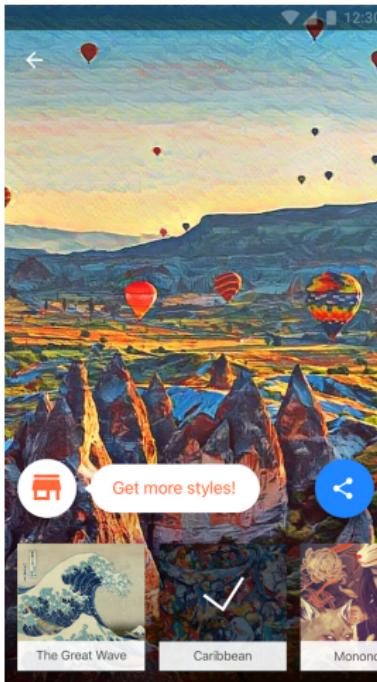
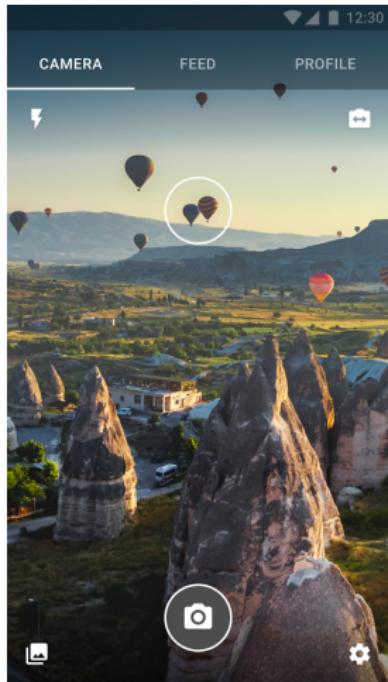
November 30, 2017

Introduction



Figure: Prisma, the App for AI Powered Art Styles

Introduction



Outline

1 Introduction

2 Picture recognition

- Biological vision for Neural Networks
- Content and Style
- Encoding of one image

3 Video recognition

- Differences in video recognition
- Specific problems
- Examples

4 Conclusion

5 References

Outline

1 Introduction

2 Picture recognition

- Biological vision for Neural Networks
- Content and Style
- Encoding of one image

3 Video recognition

- Differences in video recognition
- Specific problems
- Examples

4 Conclusion

5 References

Biological vision for Neural Networks

Biological vision for Neural Networks

We use Convolutional Neural Networks

Biological vision for Neural Networks

Biological vision for Neural Networks

We use Convolutional Neural Networks

- Many layers of processing units

Biological vision for Neural Networks

Biological vision for Neural Networks

We use Convolutional Neural Networks

- Many layers of processing units
- Each unit applies image filters

Biological vision for Neural Networks

Biological vision for Neural Networks

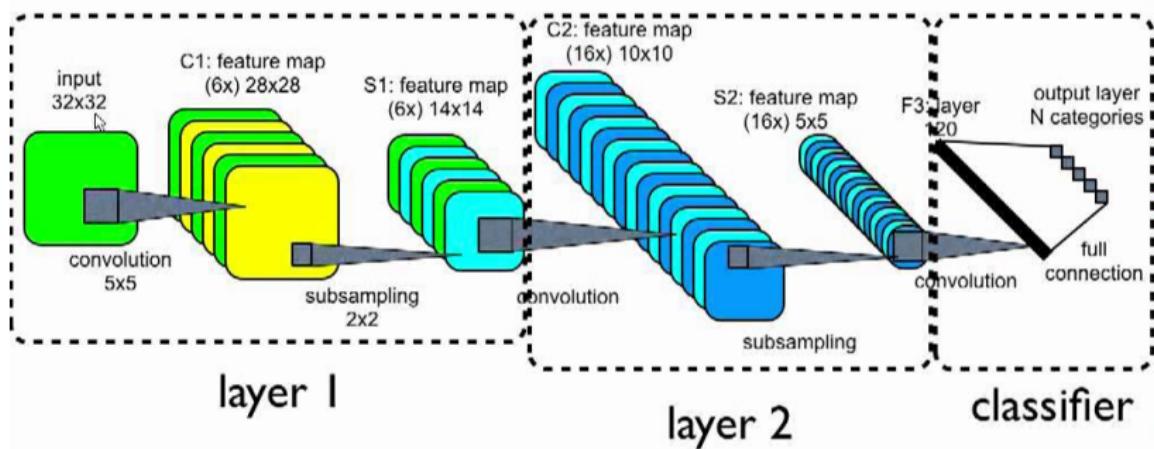
We use Convolutional Neural Networks

- Many layers of processing units
- Each unit applies image filters
- Output: Feature map

Biological vision for Neural Networks

Biological vision for Neural Networks

Convolutional Neural Networks



Biological vision for Neural Networks

Biological vision for Neural Networks

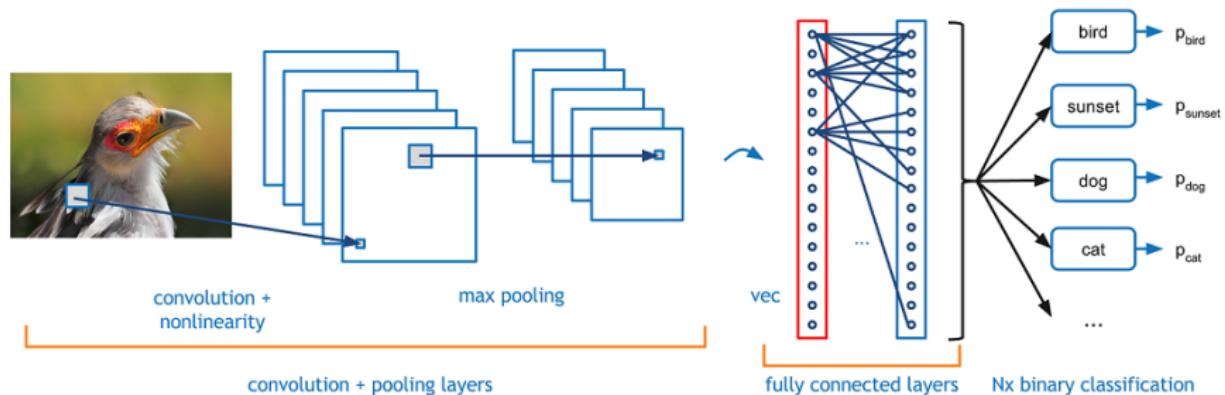
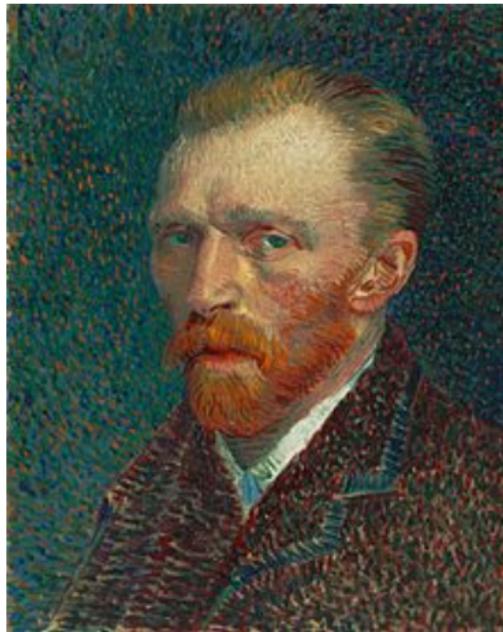


Figure: Work of a CNN

Content and Style

Content and Style



Difference between content
and style?

Figure: Van Gogh
self-portrait

Content and Style

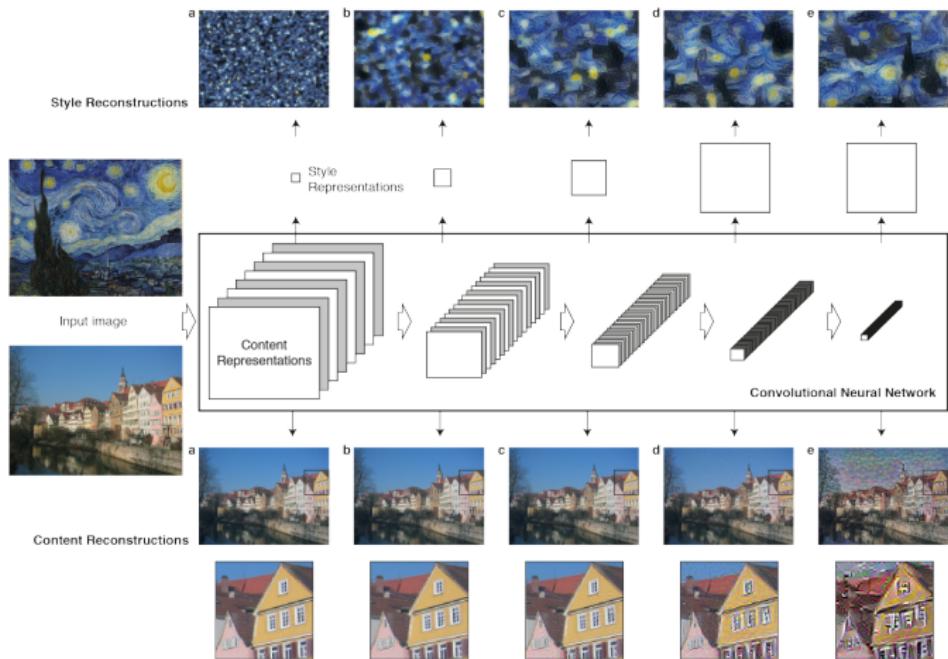


Figure: Model for the network [1]

Content and Style

Content and Style

Result:

- Representations of content and style in CNNs are separable

Content and Style

Content and Style

Result:

- Representations of content and style in CNNs are separable
- Both representations can be changed independently.

Content and Style

Content and Style

Result:

- Representations of content and style in CNNs are separable
- Both representations can be changed independently.
- → Mix style and content of two source images

Content and Style



Encoding of one image

Content reconstruction

Notation:

- \vec{p} = Original Picture, \vec{x} = initial random image

Content reconstruction: Encode \vec{x} until it generates the same response in some layers of the CNN as \vec{p}

Encoding of one image

Content reconstruction

Notation:

- \vec{p} = Original Picture, \vec{x} = initial random image
- Set \vec{x} to a white noise image

Encoding of one image

Content reconstruction

Notation:

- \vec{p} = Original Picture, \vec{x} = initial random image
- Set \vec{x} to a white noise image
- Perform gradient descent on the image

Encoding of one image

Content reconstruction

Notation:

- \vec{p} = Original Picture, \vec{x} = initial random image
- Set \vec{x} to a white noise image
- Perform gradient descent on the image
- → find \vec{x} that matches feature responses of \vec{p}

Encoding of one image

Style reconstruction

- For style reconstruction we basically do the same
- Perform gradient descent on the image
- find \vec{x} that matches style representation of original image

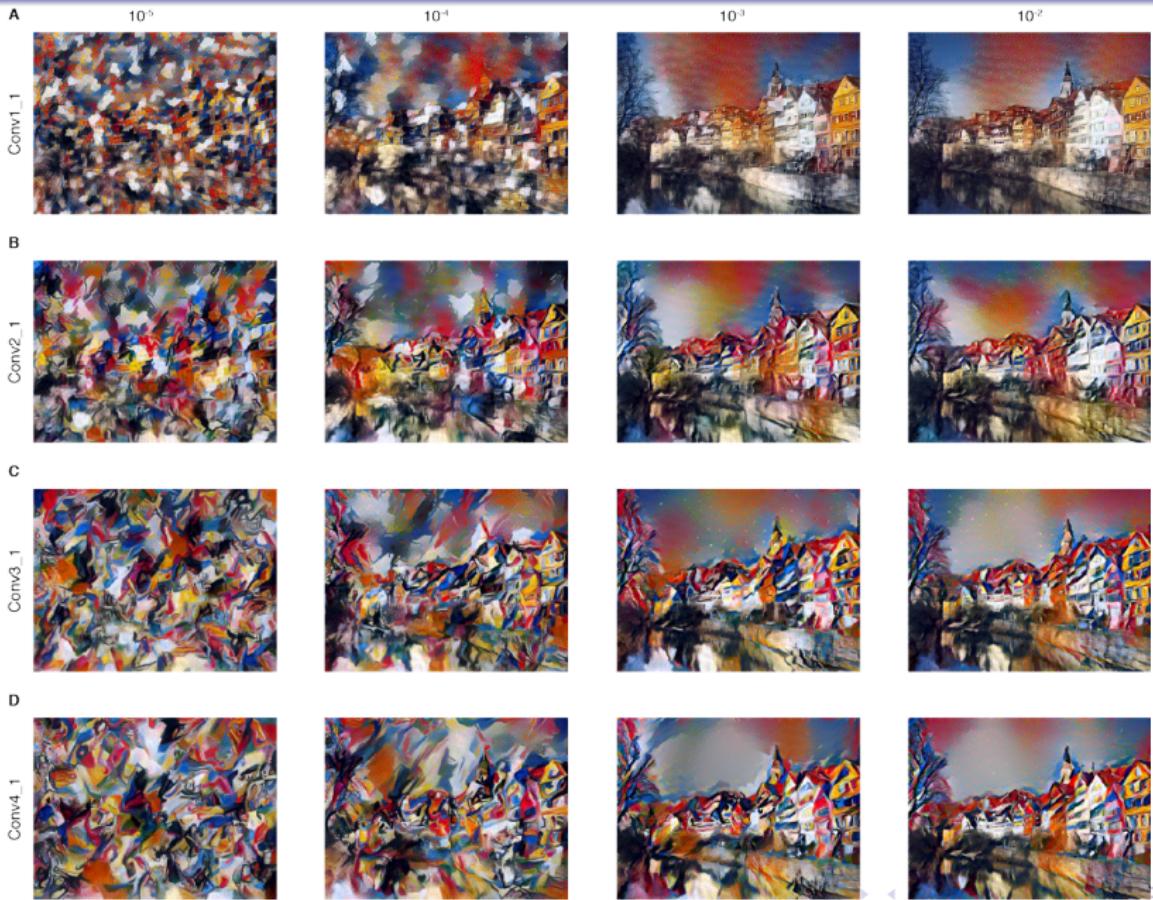
Encoding of one image

Content and style

In total: When \vec{p} is the photograph and \vec{a} is the artwork, we try to minimise the loss function:

$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x})$$

Encoding of one image



Outline

1 Introduction

2 Picture recognition

- Biological vision for Neural Networks
- Content and Style
- Encoding of one image

3 Video recognition

- Differences in video recognition
- Specific problems
- Examples

4 Conclusion

5 References

Differences in video recognition

Differences in video recognition

- First approach: Process each frame individually

Differences in video recognition

Differences in video recognition

- First approach: Process each frame individually
- → flickering and false continuities

Differences in video recognition

Video placeholder

Specific problems

Specific Problems

- Every frame is initialized with an independent white noise picture

Specific problems

Specific Problems

- Every frame is initialized with an independent white noise picture
 - every frame converges differently

Specific problems

Specific Problems

- Every frame is initialized with an independent white noise picture
→ every frame converges differently
- Idea: Initialize the optimization for the frame $i + 1$ with the stylized frame i
- Result: Problem with movement of objects

Specific problems

Short-term consistency

- Estimate optical flow
- → Warp previous frame: $x'^{(i+1)} = \omega_i^{i+1}(x^{(i)})$

Specific problems

Temporal consistency

- Stronger consistency between frames needed
- In some areas we can estimate optical flow with high confidence
- → extra Loss-function
- Calculate flow forward and backward → should be approximately the same in areas without occlusion

Specific problems

Long-term consistency

Temporally occluded areas should get the same style as before they got occluded.

→ Punish deviations from older frames

Specific problems

Multi-pass algorithm

Problem: low contrast, image boundaries less diverse

→ Process sequence multiple times and with different directions

Examples

Video placeholder

Outline

1 Introduction

2 Picture recognition

- Biological vision for Neural Networks
- Content and Style
- Encoding of one image

3 Video recognition

- Differences in video recognition
- Specific problems
- Examples

4 Conclusion

5 References

Conclusion

Style transfer in Videos:

- suitable initialization
- shortterm consistency
- longterm consistency
- → Production of stable and stylized videos possible

Outline

1 Introduction

2 Picture recognition

- Biological vision for Neural Networks
- Content and Style
- Encoding of one image

3 Video recognition

- Differences in video recognition
- Specific problems
- Examples

4 Conclusion

5 References

References

-  L. A. G. et al., “A neural algorithm of artistic style.”
-  [Online]. Available:
[https://en.wikipedia.org/wiki/Vincent_van_Gogh#/media/File:Vincent_van_Gogh_-_Self-Portrait_-_Google_Art_Project_\(454045\).jpg](https://en.wikipedia.org/wiki/Vincent_van_Gogh#/media/File:Vincent_van_Gogh_-_Self-Portrait_-_Google_Art_Project_(454045).jpg)
-  [Online]. Available: <https://play.google.com/store/apps/details?id=com.neuralprisma&hl=en>
-  M. R. et al., “Artistic style transfer for videos.”
-  [Online]. Available: <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/>