Distinguishing Professional and Abstract Art Using CNNs

Sharman Tan

CS230: Deep Learning
Fall 2019

Stanford
Computer Science

Intro & Background

Distinguishing between professional and amateur artwork *is vital when evaluating its commercial value. In 2010, psychologists Hawley-Dolan and Winner found that even untrained adults can distinguish between professionals' and children's abstract works} [1]. Can a machine learn to do the same?*

I demonstrate that by applying deep learning techniques to address issues of high variance, even standard CNN architectures can successfully distinguish between most professional and amateur artwork.

Methodology

Datasets

DeviantArt (amateur online art community) 440
MART (pro Italian/European/American art collection) 500

Figure 1: deviantArt images

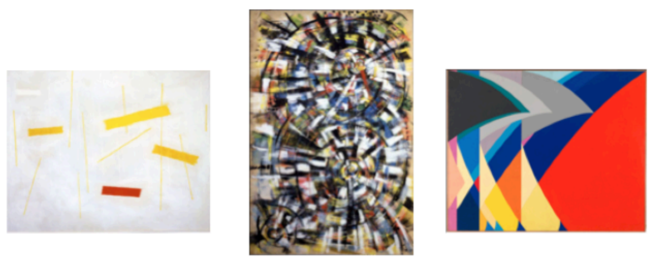

Figure 2: MART images

Table 1: Model Architectures

Baseline Model		Shallow Model	
Layer	Output	Layer	Output
Input	3, 512, 512	Input	3, 512, 512
Conv1 (f=3, p=1, s=1) * 8	8, 512, 512	Conv1 (f=3, p=1, s=1) * 8	8, 512, 512
ReLU	8, 512, 512	ReLU	8, 512, 512
MaxPool2D(2,2)	8, 256, 256	MaxPool2D(2,2)	8, 256, 256
Conv2 (f=3, p=1, s=1) * 16	16, 256, 256	Conv2 (f=3, p=1, s=1) * 8	8, 256, 256
ReLU	16, 256, 256	ReLU	8, 256, 256
MaxPool2D(2,2)	16, 128, 128	MaxPool2D(2,2)	8, 128, 128
Conv3 (f=3, p=1, s=1) * 32	32, 128, 128	Conv3 (f=3, p=1, s=1) * 16	16, 128, 128
ReLU	32, 128, 128	ReLU	16, 128, 128
MaxPool2D(2,2)	32, 64, 64	MaxPool2D(2,2)	16, 64, 64
Conv4 (f=3, p=1, s=1) * 32	64, 64, 64	Conv4 (f=3, p=1, s=1) * 16	16, 64, 64
ReLU	64, 64, 64	ReLU	16, 64, 64
MaxPool2D(2,2)	64, 32, 32	MaxPool2D(2,2)	16, 32, 32
FC1	256	FC1	256
ReLU	256	ReLU	256
FC2	84	FC2	84
ReLU	84	ReLU	84
Dropout	84	Dropout	84
Softmax	2	Softmax	2

Stanford University

Links to Code & Results

For code, please see my Github repo *sharmant/abstract-art-classification*.
All experimental results run on Google Colab are displayed in the Python3 notebook at *https://colab.research.google.com/drive/1B4rgRkUycxYdySqgZUfcNjnq6afr0FeA*.

Experiments, Results & Analysis

Addressing High Variance Tuned Dropout, Simplified Model & Data Augmentation

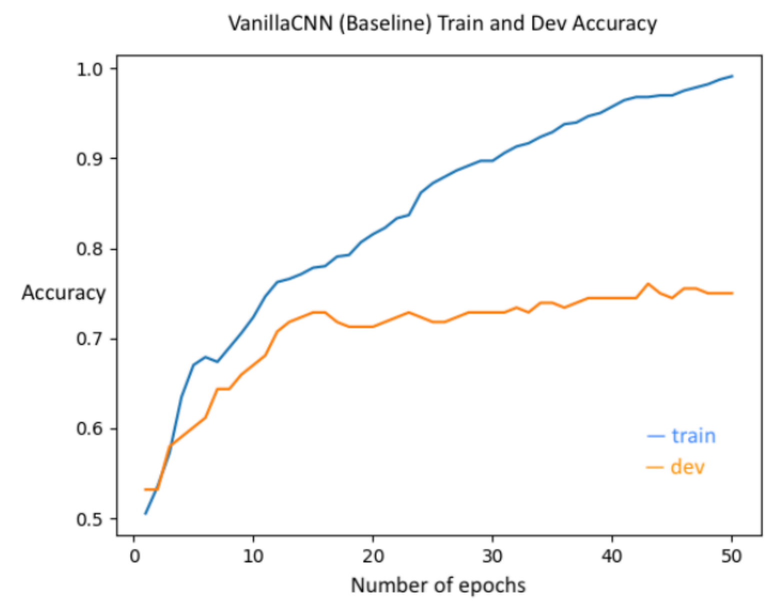

Figure 3: VanillaCNN (Baseline)

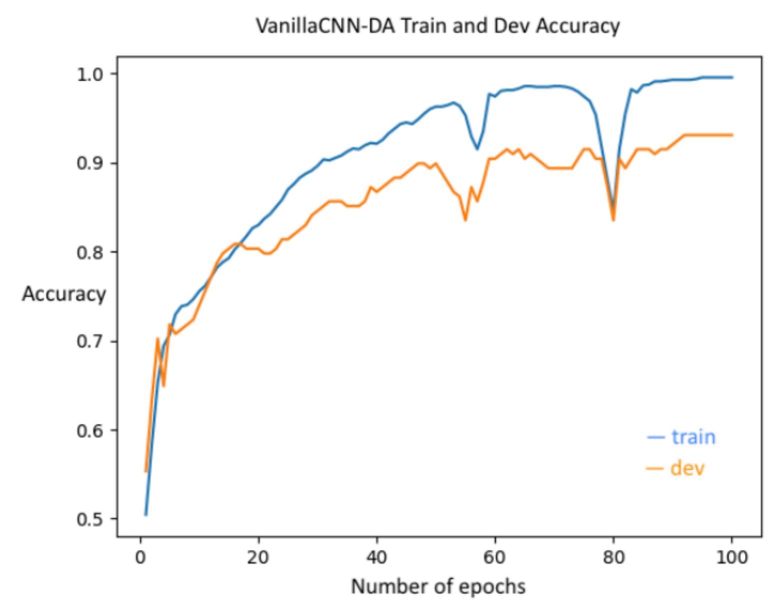

Figure 4: VanillaCNN-DA

Table 2: VanillaCNN Training and Dev Accuracy (Tuning Dropout Probability)

Dropout Prob	Learning Rate	Epochs	Train Acc (%)	Dev Acc (%)
0.1	1e-4	50	97.16	73.40
0.2	1e-4	50	92.91	73.94
0.3	1e-4	50	99.29	69.68
0.4	1e-4	50	97.34	73.94
0.5	1e-4	50	98.94	69.68
0.6	1e-4	50	93.97	74.47
0.7	1e-4	50	99.11	75.00
0.8	1e-4	50	99.11	70.74

Table 6: Final Tuned Model Train and Test Accuracies

Model	Learning Rate	Dropout	Epochs	Train Acc (%)	Test Acc (%)
VanillaCNN (Baseline)	1e-4	0.7	50	99.11	77.13
ShallowCNN	1e-4	0.2	50	84.57	78.19
VanillaCNN-DA	1e-4	0.1	100	99.56	93.09
ShallowCNN-DA	1e-4	0.4	100	99.65	92.55

Figure 5: All Incorrectly Classified Images for VanillaCNN-DA and ShallowCNN-DA

1 VanillaCNN 50 epochs
train **99.11** || dev **75.00** || test **77.13**

Tuned Dropout Prob
0.1, 0.2, 0.3, 0.4, 0.5, 0.6, **0.7**, 0.8
Tuned Learning Rate
1e-5, 3e-5, **1e-4**, 3e-4, 1e-3
Problem High variance
Large gap b/t train and dev loss (Fig. 3)

2 ShallowCNN 50 epochs
train **84.57** || dev **76.60** || test **78.19**

Fewer Convolutional Filters
Simplify CNN architecture (Table 1)
Tuned Dropout Prob
0.1, **0.2**, 0.3, 0.4, 0.5
Problem High variance
Still large gap b/t train and dev loss, < 80% acc

3 VanillaCNN-DA 100 epochs
train **99.56** || dev **93.09** || test **93.09**

4 ShallowCNN-DA 100 epochs
train **84.57** || dev **92.02** || test **92.55**

Data Augmentation
Vertically flip all images (double count)
Randomly split 60/20/20 train/dev/test
Use *new* train and *old* dev/test sets
train **1128** || dev **188** || test **188**
Tuned Dropout Prob
Vanilla **0.1**, 0.2, 0.3, 0.4, 0.5
Shallow 0.1, 0.2, 0.3, **0.4**, 0.5

Takeaways

Main problem High variance
Solutions
Tune dropout probability
Simplify model architecture
[KEY] Increase training data using data augmentation
Analyzing Misclassified Examples
Amateur (predicted Pro)
highly complex & colorful
simple and monochrome
Professional (predicted Amateur)
geometric
2 main colors, large blobs

Conclusion & Next Steps

Standard CNN architectures work!

By tackling the issue of high variance*, even standard CNN architectures are successful at distinguishing professional and amateur abstract artwork.*

More Advanced CNNs

- Can more advanced CNNs correctly label easy-to-misclassify images? (e.g. when simple monochrome pieces are professional vs. amateur)

More Data

- Data augmentation by vertical flips drastically boosted accuracy!
- Can performing more data augmentation (e.g. crops, horizontal flips) further improve accuracy?
- Increasing data overall (train/dev/test)

What are humans missing?

- What exactly are the CNNs learning that humans find difficult to learn?
- Which features contribute most to their predictions? Color? Types of edges?

Content vs. Style

- Can focusing on content vs. style of images change how images are classified?
- Can attention mechanisms be applied to focus on the most distinguishing aspects?

Acknowledgements

I would like to thank Huizi Mao for mentoring me throughout this project and the CS230 teaching staff for the knowledge I've gained through this class.

References

1. Andreza Sartori. http://disi.unitn.it/~sartori/. Accessed: 2019-10-10.
2. Deviantart. https://www.deviantart.com. Accessed: 2019-10-10.
3. Mart. http://www.mart.trento.it. Accessed: 2019-10-10.
4. Angelina Hawley-Dolan and Ellen Winner. Seeing the mind behind the art: People can distinguish abstract expressionist paintings from highly similar paintings by children, chimps, monkeys, and elephants. Association for Psychological Science, 2011.
5. Wei-Lin Hsiao and Kristen Grauman. Learning the latent "look": Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing image style. arXiv preprint arXiv:1311.3715, 2013.