

Art Appraisal Using Convolutional Neural Networks

Cedric Orban, Rafi Ayub, Vidush Mukund
{cedrico,rafiayub,vmukund}@stanford.edu

Abstract

In order to provide a standardized method of appraising artwork, we suggest the use of a convolutional neural network (CNN). We constructed a CNN using the framework of TensorFlow that takes in as inputs the three RGB channels of a given artwork. We gathered the images of about 100,000 artworks along with associated metadata (artist name, sell price, medium, etc.). After training various CNN architectures, it is apparent that a CNN without the name of the artist is unable to achieve any meaningful accuracy.

Data

- 100,000 400x400 RGB images of artwork scraped from findartinfo.com
- Each piece was paired with metadata (artist name, dating, sell date, sell price, medium, signed)

Features

CNN

- 3D tensor of RGB values (float32) for each artwork

Logistic Regression

- Artist Name
- Age of painting
- Medium
- Signed/Not signed

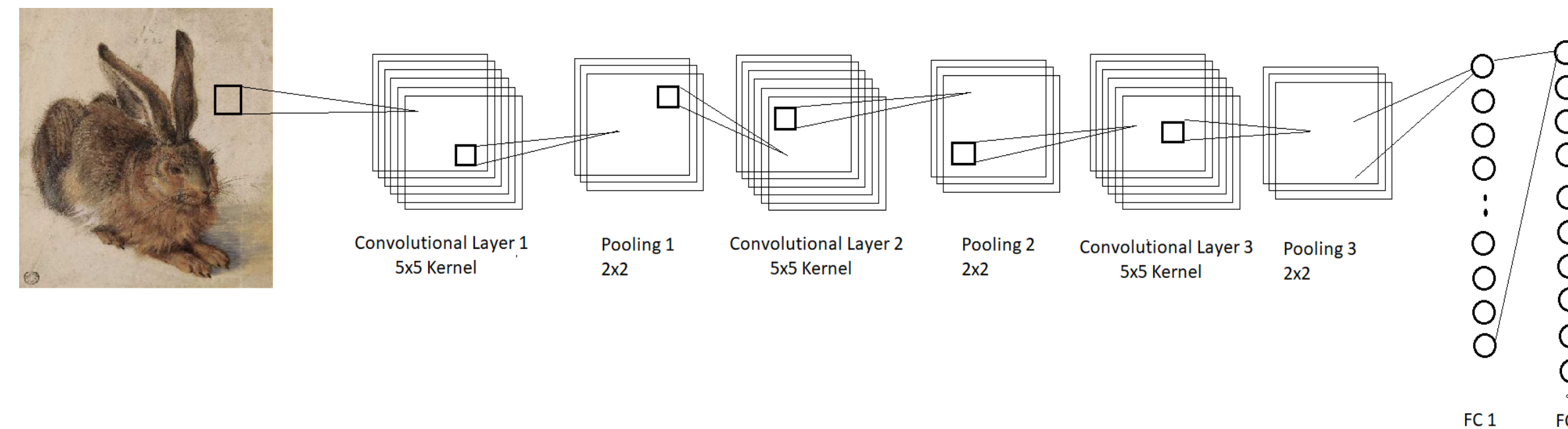
References

1. D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in IJCAI. AAAI Press, 2011, pp. 1237–1242.
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks." in NIPS, vol. 1, no. 2, 2012, p. 4
3. Suchecki, M., & Trzcinski, T. (2017). Understanding Aesthetics in Photography using Deep Convolutional Neural Networks. Retrieved from <http://arxiv.org/abs/1707.08985>
4. You, Q., Pang, R., Cao, L., & Luo, J. (2016). Image Based Appraisal of Real Estate Properties, 1–8. <https://doi.org/10.1109/TMM.2017.2710804>

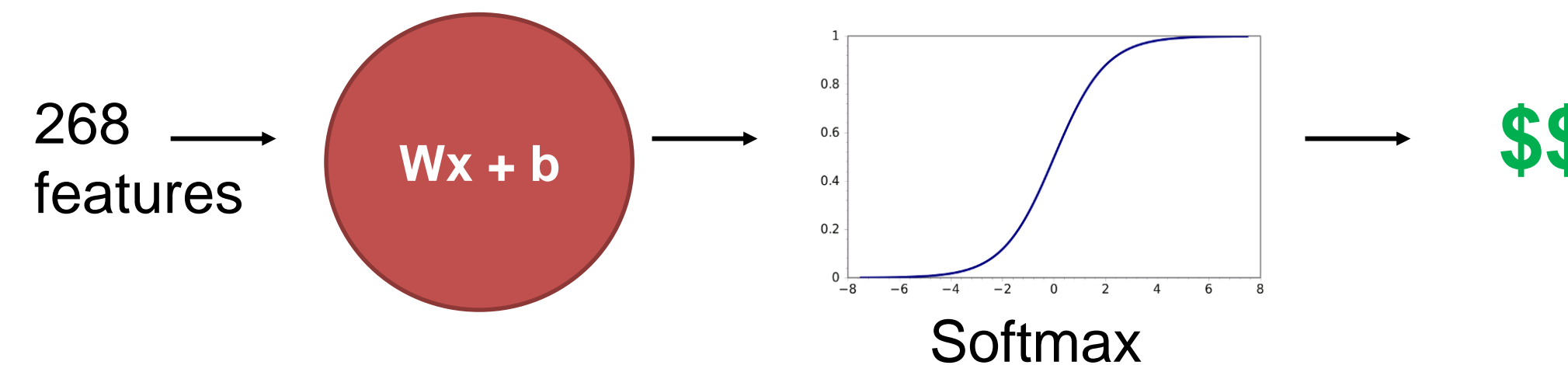
Models

CNN

- 3 convolutional/pooling layers, 2 fully connected layers
- Slight variations on the above (bin sizes, +/- 1 layer)
- AlexNet

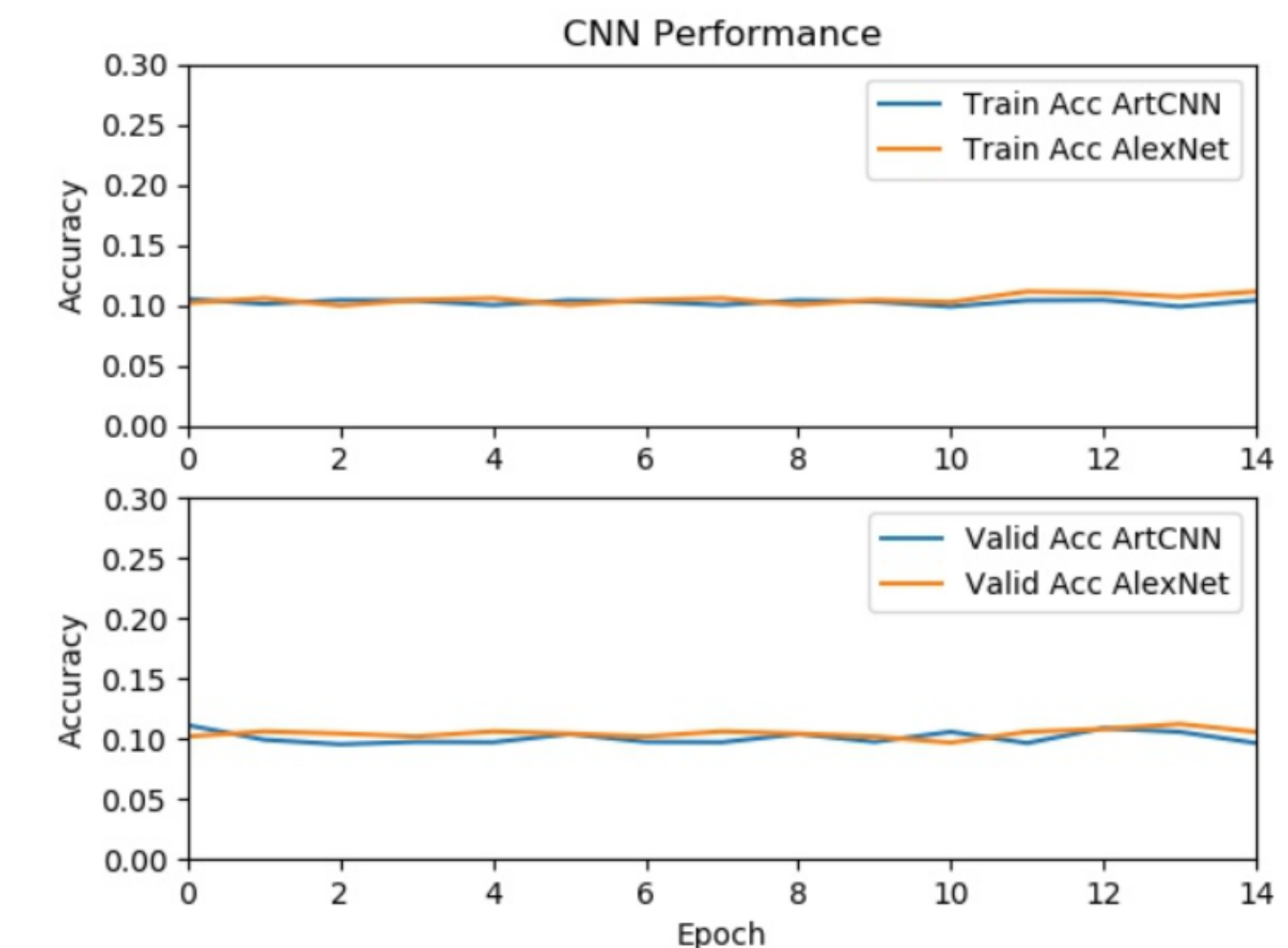


Logistic Regression



Results

Our trained CNN models consistently underperformed on our validation and test sets. The predictions were essentially random, with the accuracy directly influenced by the size and number of our output bins.



Model	Train Size	Test Size	Train Acc	Test Acc
Art CNN	90,000	10,000	0.106	0.101
AlexNet	90,000	10,000	0.103	0.110
Log Regr	600	100	0.413	0.275

Discussion

- CNNs performed abysmally despite direct use of visual features of artwork.
- Context of artwork, including artist name and art medium, was suspected to be a large factor in sell price, and this is supported by the higher accuracy achieved by logistic regression
- Extremely low accuracies by the CNN may suggest an inherent subjectivity to artwork that cannot be objectively captured by algorithms that only rely on visual features

Future

- Future avenues of approach include using the same models to classify subsets of the full dataset (e.g. classifying all work by one artist or all pieces rendered in the same medium).
- In regards to the logistic regression approach, ablative analysis of the metadata features used would indicate the most relevant features for sell price classification.