Automatic Identification of Footwear Class Characteristics

Miranda Tilton and Dr. Susan VanderPlas, Iowa State University

csafe Center for Statistics and Applications in Forensic Evidence

NIST National Institute of Standards and Technology U.S. Department of Commerce

Objective

Identify class characteristics of footwear outsole images automatically.

Background

In shoe print analysis, it is often useful to determine the frequency of a given shoe print (or features of the print) in a local population. Machine learning tools, such as neural networks, are an inexpensive and efficient way to automatically identify these features, which may inform about the prevalence of such features.

Class Characteristics

Footwear class characteristics, including size, orientation, and position of geometric elements, are capable of distinguishing most shoes collected in samples from the general population [1], and can be used to speed up database searches for candidate shoe models [2].

Table 1: Geometric features. Categories modified from [3]

Unknown shoes and knockoffs can be classified using geometric features. Spatial relationships between features can be used to add complexity to the feature set.

Convolutional Neural Nets

- A convolutional neural network (CNN) is a tool for deep learning that is well-suited to image analysis.
- Inspired by the brain, CNNs learn global patterns using a hierarchy of local feature detection and pooling
- VGG16 is a CNN [4] pre-trained on ~1.3 million images spanning 1,000 categories from ImageNet [5].

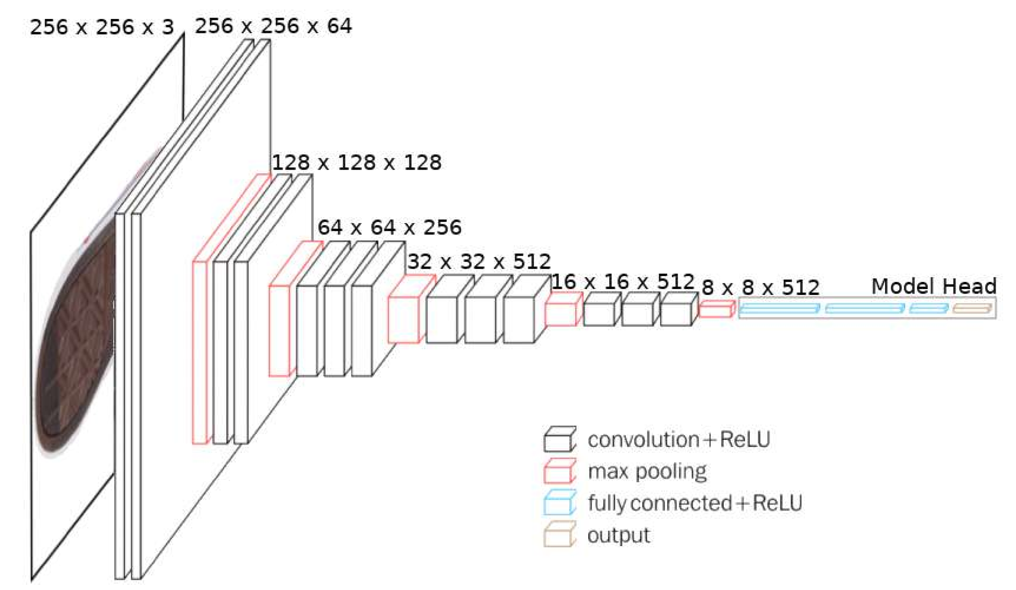
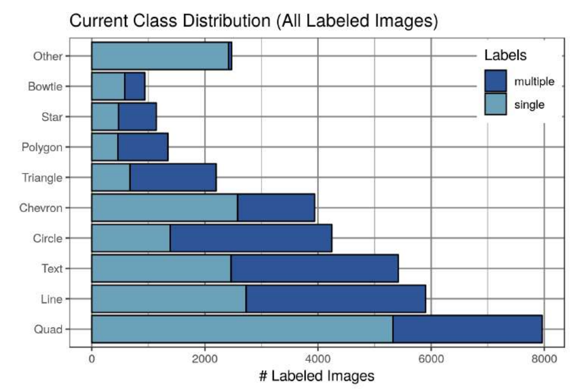


Fig. 1: Architecture of VGG16

Data



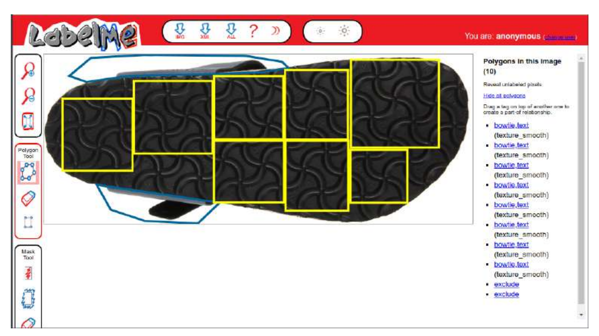


Fig. 2: Label Frequency **Fig. 3:** LabelMe Tool [6]

- 26k multi-label images from 4,400 shoes
- Validation set: 20%, used during training
- Test set: 20%, used after training
- Training set: 60%, with images augmented once (crop, zoom, skew, rotate, and color)

Fig. 4: Examples of original (left) and augmented (right) images.

Training

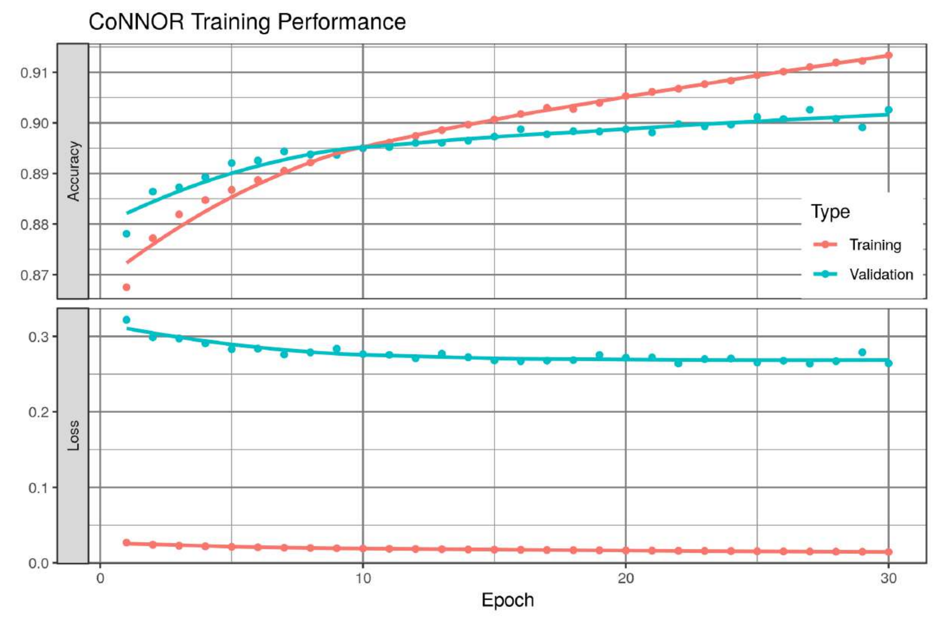


Fig. 5: Training and Validation accuracy and loss for each epoch of the fitting process.

Prediction Accuracy

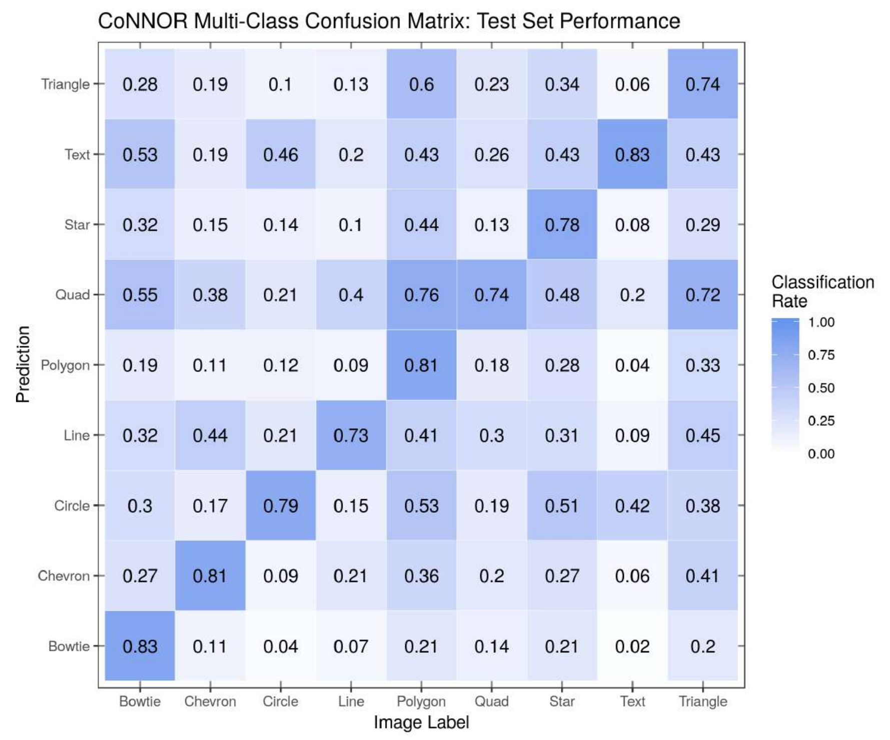


Fig. 6: Percent of test images identified as containing a class with probability $\geq$ equal error rate, after accounting for multiple labels

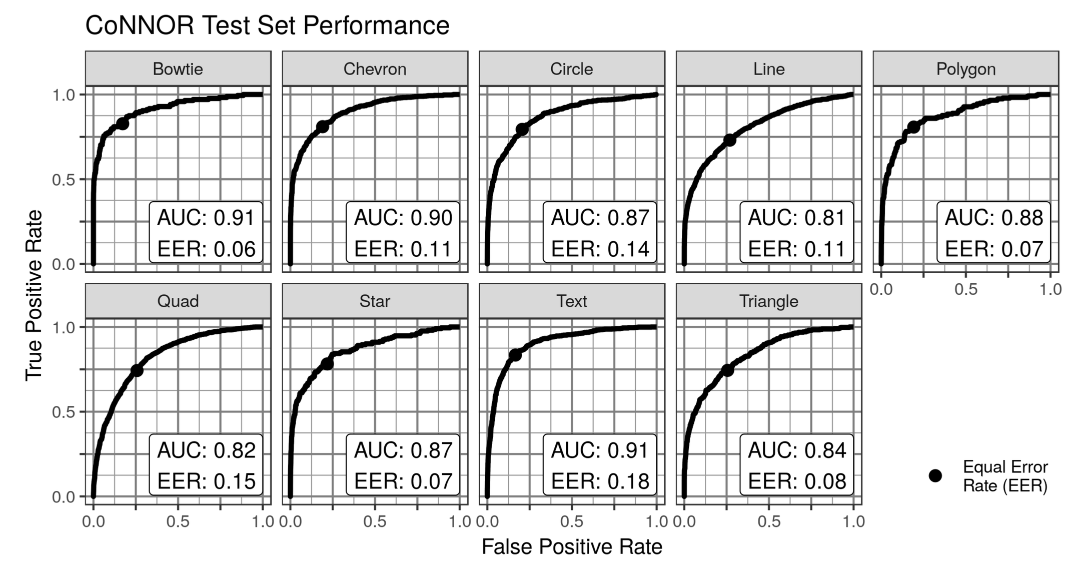


Fig. 7: True Positive Rate by False Positive Rate of classification for each class

Features as Distance

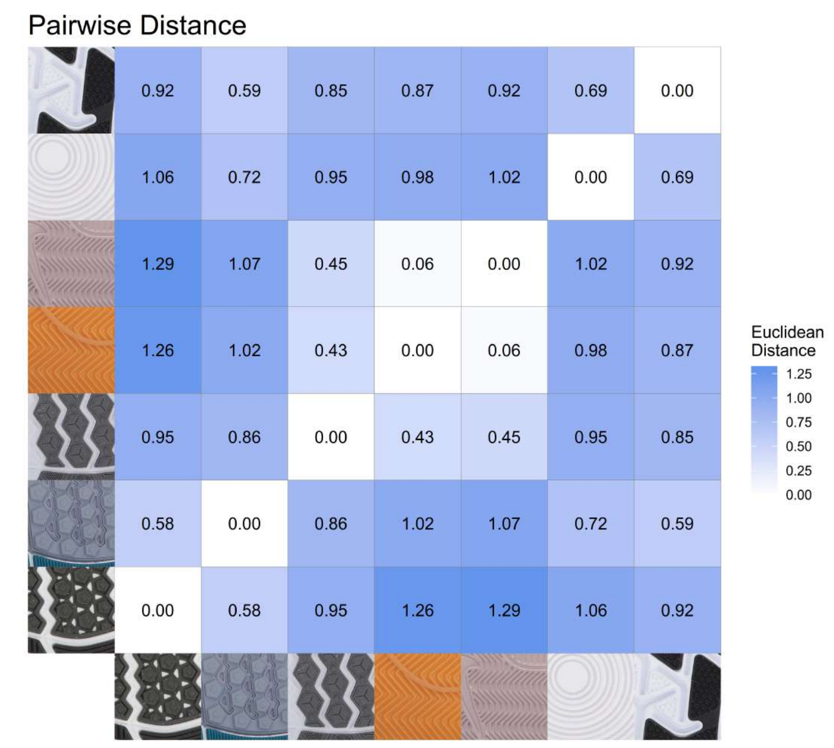


Fig. 8: Pairwise Euclidean distance between selected example images

Future Applications

- Input features for statistical models assessing match strength
- Speed up database searches: feature indexing
- Estimate frequency of similar shoes given information about local population

References

[1] Sheida Hancock, Rian Morgan-Smith, and John Buckleton. The interpretation of shoeprint comparison class correspondences. *Science and Justice*, 52(4):243–248, 2012.

[2] Maria Pavlou and Nigel M Allinson. Automatic extraction and classification of footwear patterns. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 721–728. Springer, 2006.

[3] Susan Gross, Dane Jeppesen, and Cedric Neumann. The variability and significance of class characteristics in footwear impressions. *Journal of Forensic Identification*, 63(3):332, 2013.

[4] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. September 2014.

[5] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.

[6] Bryan C. Russell, Antonio Torralba, Kevin P. Murphy, and William T. Freeman. Labelme: A database and web-based tool for image annotation. *Int. J. Comput. Vision*, 77(1-3):157–173, May 2008.

This work was partially funded by CSAFE through Cooperative Agreement # 70NANB15H176 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, University of California Irvine, and University of Virginia