EPJ Web of Conferences **287**, 13015 (2023) https://doi.org/10.1051/epjconf/202328713015
EOSAM 2023

Feature and texture distillation via neural network training

Altai Perry, *Xiaojing Weng*, *Ji Feng,* and *Luat T. Vuong*

University of California at Riverside, Riverside, California, USA

Abstract. Encoded-diffraction hybrid systems—optical encoding and simple electronic decoding—offers feature distillation via model training. Additionally, the most faithfully reconstructed images are not the ones that are best classified. We parametrize our results with singular value decomposition (SVD) entropy, a proxy for image complexity.

1 Introduction

There is recently significant interest in hybrid optical-electronic systems that leverage optical encoding and simple back-end electronic decoding of data. Optics provides high-speed, massively parallel, critical low-power data processing while electronics provides the variable sensor sampling, nonlinear logistic and thresholding needed for decision-making and data classification. While hybrid image processing with simple neural networks have achieved significant milestones [1], there is currently data that suggests it is necessary to have complex computing architectures in order to analyse more complex images [2]. Here, we explore the issues of image complexity, fractal texture analysis, and the potential for dimension reduction via neural network model training.

Image complexity often refers to the degree by which an image can be compressed. The upper bound for compression is the quantity famously dubbed Shannon entropy and is rooted in the statistics of information, defined as $H_{Shan} = -\mathbf{\Sigma}_i$ pi log p_i where p_i is the probability of a positive bit. This measure characterizes time-domain signal data and is also used to characterize image or multidimensional data. In fact, Shannon entropy reveals little about an image since a shuffle of the image pixels yields the same classical value for Shannon entropy. While other entropy measures have been proposed as a proxy for image complexity [3], creating a generalizable image complexity measure remains a difficult problem [4]. Recently, with optical neural networks and Fourier-domain inputs, we have shown that the choice of generalized training images leads to learned preferences that distil different salient features of the reconstructed images based on spectral correlations [5]. In this work, we have also shown that certain images are more easily learned than others, and faster learning times correspond to lower image complexity.

However, we have also shown that complicated neural networks may be avoided, even when images are complex if we are only concerned with image classification. That is, the most faithfully reconstructed images are not always best classified. Our results indicate that simpler optical neural network classifiers are available to analyse complex datasets via dimensionality reduction. Our analysis focuses on the application of SVD-entropy as a proxy for image complexity

$$H_{SVD} = \mathbf{\Sigma}_i\, \boldsymbol{\sigma}_i \log \boldsymbol{\sigma}_i$$

where $\boldsymbol{\sigma}_i$ are the singular values. When normalized by the size of the matrix or range of matrix data, H_{SVD} characterizes quantity of eigenimages needed to reconstruct an image in a manner that is consistent with definitions of image complexity and also invariant with fractal dimension [6].

We train image datasets with different fractal and random training datasets including speckle patterns. We show that the resulting model distils and achieves image segmentation not only for objects (which we have achieved previously [7]) but also for textures. Finally, when we reclassify these images, we achieve order-of-magnitude improvements in the accuracies, many of which rival decade-old computer vision state-of-the-art.

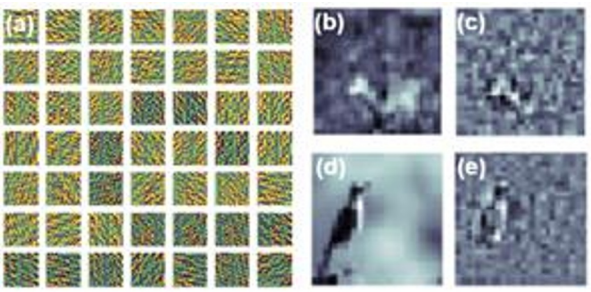

Figure 1: Object segmentation with CIFAR-10 natural-scenes. a) Textured, fractal training dataset and corresponding b-d)

* Corresponding author: author@e-mail.org

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

ground-truth images c-e) object segmentation and texture illumination.

Results

Figure 1 shows results training models for image reconstruction with synthetic fractal datasets. The resulting CIFAR-10 image data background textures are highlighted. Figure 2 shows training and subsequent classification with image classification using H_{SVD} and a spatial frequency density, Ω_k. While the classification of raw images achieves only a few percent accuracy, the reconstructed, filtered images achieve 30-58% accuracy. The parameterization was introduced in [5].

Conclusion

We achieve feature distillation and segmentation with texture using learned models that are trained with synthetic fractal and random data. We show that simple hybrid optical neural networks may analyse complex images via reduced dimensions from training.

Funding

FACE grant TJF20-056 and DARPA YFA D19AP00036.

References

1. G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, and D. Psaltis, "Inference in artificial intelligence with deep optics and photonics," Nature 588, 39–47 (2020).
2. T. Wang, M. M. Sohoni, L. G. Wright, M. M. Stein, S.-Y. Ma, T. Onodera, M. G. Anderson, and P. L. McMahon, "Image sensing with multilayer nonlinear optical neural networks," Nat. Photon. 17, 408–415 (2023).
3. Q. R. Razlighi and N. Kehtarnavaz, "A comparison study of image spatial entropy," in Visual Communications and Image Processing 2009, M. Rabbani and R. L. Stevenson, eds. (SPIE, 2009).
4. L. Madrid-Herrera, M. I. Chacon-Murguia, D. A. Posada-Urrutia, and J. A. Ramirez-Quintana, "Human Image Complexity Analysis Using a Fuzzy Inference System," in 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE, 2019).
5. A. Perry, X. Weng, E. Nozari, and L. Vuong, "Parameterized Learning and Distillation with Vortex-encoded Spectral Correlations," (2023).
6. X. Weng, A. Perry, M. Maroun, and L. T. Vuong, "Singular Value Decomposition and Entropy Dimension of Fractals," 2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML) (2022).
7. B. Muminov, A. Perry, R. Hyder, M. S. Asif, and L. T. Vuong, "Toward simple, generalizable neural networks with universal training for low-SWaP hybrid vision," Photon. Res. 9, B253 (2021).

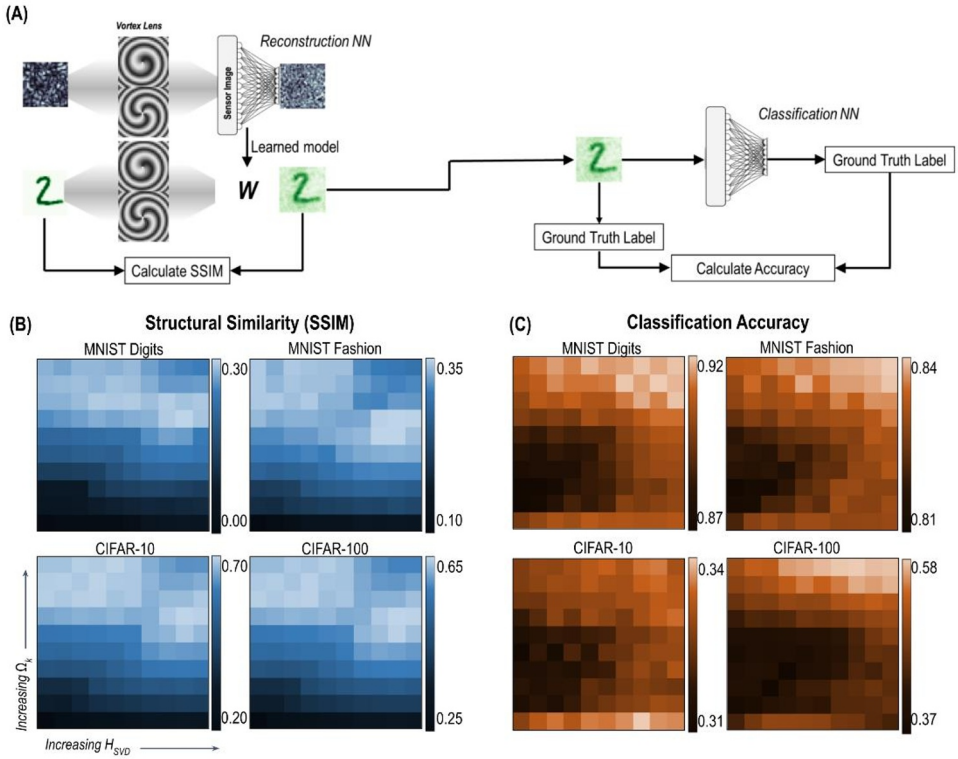

Figure 2: (A) A learned model is used to reconstruct test images (MNIST Digits, MNIST Fashion, CIFAR-10, and CIFAR-100). These reconstructed images are then used to train 2-layer classification NNs. Note that the (B) Average Structural Similarity Index (SSIM) of reconstructed test images. Each square corresponds to the SSIM of the particular dataset reconstructed using a model trained with a different dataset. Representative images of the datasets used are shown in a grid. (C) Classification accuracies of a 2-layer NN trained on images reconstructed and represented in (B). Models that were trained on the ground truth images themselves rather than the reconstructed images achieved 94\% - MNIST Digits, 86\% - MNIST - Fashion, 26\% - CIFAR-10, 8.2\% - CIFAR-100.