Manifold-based Test Generation for Image Classifiers

Taejoon Byun
taejoon@umn.edu
University of Minnesota
Minneapolis, MN

Sanjai Rayadurgam
rsanjai@umn.edu
University of Minnesota
Minneapolis, MN

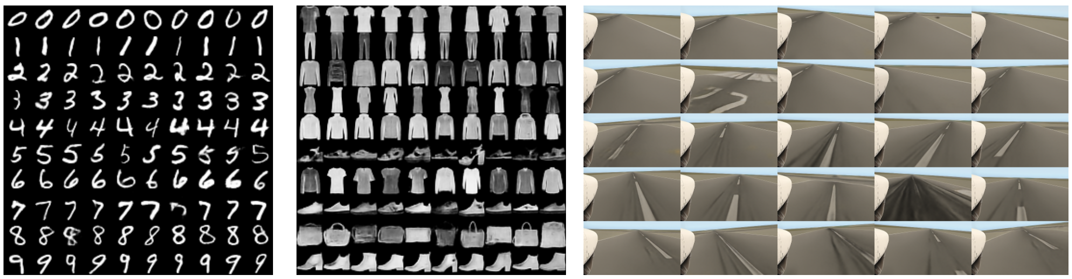

Figure 1: Images generated by the conditional Two-stage VAEs: sampled randomly from the second-stage manifold space. Images in each row are conditioned by the same class label (no cherry-picking).

ACM Reference Format:
Taejoon Byun and Sanjai Rayadurgam. 2020. Manifold-based Test Generation for Image Classifiers. In *IEEE/ACM 42nd International Conference on Software Engineering Workshops (ICSEW'20), May 23–29, 2020, Seoul, Republic of Korea.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3387940.3391460

1 SUMMARY

Neural network image classifiers are being adopted in safety-critical applications, and they must be tested thoroughly to inspire confidence. In doing so, two major challenges remain. First, the thoroughness of testing needs to be measurable by an adequacy criterion that shows a strong correlation to the semantic features of the images. Second, a large amount of diverse test cases needs to be prepared, either manually or automatically. The former can be aided by neural-net-specific coverage criteria such as surprise adequacy [3] or neuron coverage [4], but their correlation to semantic features had not been evaluated. The latter is attempted through metamorphic testing [5], but it is limited to domain-dependent metamorphic relations that requires explicit modeling.

This presentation discusses a framework which can address the two challenges together. Our approach is based on the premise that patterns in a large data space can be effectively captured in a smaller manifold space, from which similar yet novel test cases—both the input and the label—can be synthesized. This manifold space can also serve as a basis for judging the adequacy of a given test suite, since the manifold encodes all the necessary information for distinguishing among different data points. For modeling this manifold and creating a pair of encoder and a decoder that maps between manifold space and input space, we utilized a conditional variational autoencoder (VAE). The conditional VAE learns class-dependent manifold which enables class-conditioned test generation, solving the oracle problem by construction. For generating novel test cases, we applied search on the manifold to effectively find fault-revealing test cases. Experiments for test case generation show that this approach enables generation of thousands of realistic yet fault-revealing test cases efficiently even for well-trained models that achieve a high validation accuracy. Experiments for coverage measurement shows that manifold-based coverage exhibits higher correlation to semantic features—represented by class label—compared to neuron coverage or neuron boundary coverage. These results suggest that the concept of manifold-based testing is a promising direction for machine learning testing, and calls for a further investigation.

The original work is accepted to be presented in AI Test 2020 [2], and a part of the idea will also be presented in ICSE–NIER 2020 [1].

ICSEW'20, May 23–29, 2020, Seoul, Republic of Korea

ACM ISBN 978-1-4503-7963-2/20/05…$15.00
https://doi.org/10.1145/3387940.3391460

ACKNOWLEDGMENTS

This work was supported by AFRL and DARPA under contract FA8750-18-C-0099.

REFERENCES

[1] Taejoon Byun and Sanjai Rayadurgam. 2020. Manifold for Machine Learning Assurance. *arXiv preprint arXiv:2002.03147* (2020).
[2] Taejoon Byun, Abhishek Vijayakumar, Sanjai Rayadurgam, and Darren Cofer. 2020. Manifold-based Test Generation for Image Classifiers. *arXiv preprint arXiv:2002.06337* (2020).
[3] Jinhan Kim, Robert Feldt, and Shin Yoo. 2019. Guiding Deep Learning System Testing Using Surprise Adequacy. In *Proceedings of the 41st International Conference on Software Engineering* (Montreal, Quebec, Canada) *(ICSE '19).* IEEE Press, Piscataway, NJ, USA, 1039–1049.
[4] Kexin Pei, Yinzhi Cao, Junfeng Yang, and Suman Jana. 2017. DeepXplore: Automated Whitebox Testing of Deep Learning Systems. In *SOSP '17.* ACM Press, New York, New York, USA, 1–18. arXiv:1705.06640
[5] Yuchi Tian, Kexin Pei, Suman Jana, and Baishakhi Ray. 2018. DeepTest: Automated Testing of Deep-neural-network-driven Autonomous Cars *(ICSE '18).* ACM, New York, NY, USA, 303–314.