Deep One-Class Classification: Supplementary Material\textsuperscript{1}

Lukas Ruff  Robert A. Vandermeulen  Nico Görnitz  Lucas Deecke
Shoaib A. Siddiqui  Alexander Binder  Emmanuel Müller
Marius Kloft

2018

\textsuperscript{1}Supplementary material of the paper Deep One-Class Classification published in Proceedings of the 35\textsuperscript{th} International Conference on Machine Learning, 2018. Copyright 2018 by the author(s).
1 MNIST

Figure 1: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples per class on MNIST according to Deep SVDD anomaly scores ordered by score from top left to bottom right.
Figure 2: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples per class on MNIST according to Deep SVDD anomaly scores ordered by score from top left to bottom right.
Figure 3: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples per class on CIFAR-10 according to Deep SVDD anomaly scores ordered by score from top left to bottom right.
Figure 4: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples per class on CIFAR-10 according to Deep SVDD anomaly scores ordered by score from top left to bottom right.
3 Adversarial Attacks on GTSRB stop signs

Figure 5: The 20 adversarial examples generated by using Boundary Attack.

Figure 6: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples on GTSRB stop signs training set according to Deep SVDD anomaly scores ordered by score from top left to bottom right.

Figure 7: Example from one experiment (seed) of the 32 most normal (left) and 32 most anomalous (right) examples on GTSRB stop signs test set according to Deep SVDD anomaly scores ordered by score from top left to bottom right. The example shows that Deep SVDD detects adversarial attacks.