Pedro Blöss Braga pedro.bloess@fau.de
M.Sc. Computational and Applied Mathematics candidate

Friedrich-Alexander-Universität Erlangen-Nürnberg

Deep Anomaly Detection Using Geometric Transformations

Based on the method presented in "Deep Anomaly Detection using Geometric Transformations" (Golan and El-Yaniv 2018) [1], extensions to the image transformation set and normality score are inspected. A spatial feature importance analysis is provided analyzing the activation maps and gradients. Small-scale simulations demonstrate ROCAUC development.

I. Experiments

A. Additional Geometric Transformations

Apart from the 3 sets of transformations in [1], which are horizontal flip, rotation, and translation, new image transformations were tested: quantile histogram equalization, cropping, color jitter, and zoom.

B. Spatial localization of features

Extracting the activation map of relevant layers (e.g. convolutional and activation) of the trained network enables a feature localization method. A mask is created filtering high activations by a threshold.

An intuitive representation of general feature important is suggested, by averaging activation maps of the last convolutional layers. Results show high activation values in image borders, corners, and other regions.

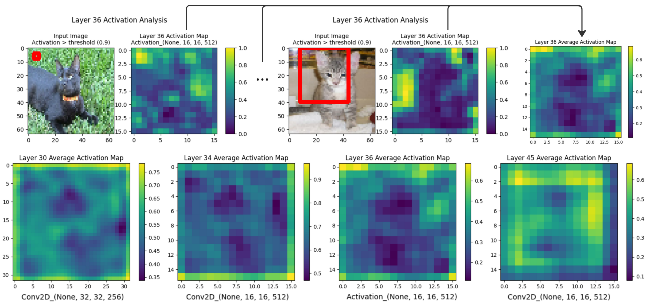

FIG. 1: Scheme for the average activation map, and examples of final layers.

Additionally, applying Grad-CAM [2], visual model explainability is enabled, by weighting 2D-activations with the average gradient.

FIG. 2: Original image (left); Grad-CAM on flip (middle) and rotate (right).

C. Normality score

As an alternative to the Dirichlet normality score, an entropy score [3] [4] is introduced via the mapping I.1, given probabilities p_i.

$$H(p) = -\sum_{i=1}^{N} p_i \log(p_i) \tag{I.1}$$

removing the necessity of a maximum likelihood estimation of parameters $\tilde{\alpha}_i$ via the fixed point iteration method, and producing higher ROC AUC values.

D. Uncertainty Estimation

Incorporating uncertainty into predictions helps to quantify model confidence and reliability. Monte Carlo Dropout [5] enables estimation of the posterior predictive distribution as $\hat{p}(y|x) \approx \frac{1}{N}\sum_{i=1}^{N} \hat{y}_i$, where $\hat{y}_i = f_D(x; \tilde{\theta}_i)$ represent predictions of a stochastic model f_D with parameters $\tilde{\theta}_i \sim \text{Dropout}(\theta)$, given input x. Subsequently, the uncertainty metric $\text{Var}[\hat{y}] = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - \mathbb{E}(\hat{y}))^2$ is introduced, to measure the variability in predictions.

II. Methodology

The experiments were mostly executed with samples of the "cats-vs-dogs" dataset, given the small number of classes, and high benchmark performance in [1]. Training was focused on class 0, with inference in both classes, and the implementation can be found at https://github.com/pedroblossbraga/DeepAnomDetecGeomTransf.

III. Results

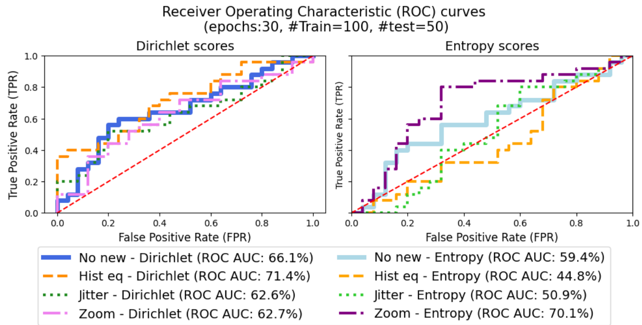

FIG. 3: ROC curve of different experiments, comparing different transformation combinations and normality scores. "No new" refers to the original set in [1], and the remaining refers to the original set adding 1 new transformation.

Overall, within Dirichlet scores, Quantile Histogram Equalization ($Q = 0.7$) demonstrated superior ROCAUC, whereas within Entropy scores, Zoom surpassed others with a wider margin. Comparing scores, in a small-scale experiment with only 10 Epochs entropy scores were preferred. However, by increasing to 30 epochs 3, there is only improvement with the zoom transformation. Furthermore, entropy computation is faster, by a ratio of 4.8 on a test with the original transformation set.

IV. Conclusion

Additional transformations, especially Zoom and Histogram Equalization, displays potential performance improvements. Moreover, the Entropy score I.1 shows higher ROCAUC, when Zoom is added, and it is computationally cheaper.

The uncertainty estimation leverages an additional layer of discrimination for normality. However, tests with much larger data sets and epochs should be more conclusive.

[1] I. Golan and R. El-Yaniv, in *Advances in Neural Information Processing Systems*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., 2018), vol. 31.
[2] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, International journal of computer vision **128**, 336 (2020).
[3] A. Taha and A. S. Hadi, ACM Computing Surveys (CSUR) **52**, 1 (2019).
[4] P. Bereziński, B. Jasiul, and M. Szpyrka, Entropy **17**, 2367 (2015).
[5] Y. Gal and Z. Ghahramani, in *international conference on machine learning* (PMLR, 2016), pp. 1050–1059.