論文の要旨

題目：**Feature extraction via autoencoder-based network for anomaly detection**

（異常検出のためのオートエンコーダベースのネットワークを介した特徴抽出）

氏名 Qien Yu

Optimizing the extraction of feature sets for anomaly detection tasks (AD) is still a fundamental and challenging problem in the field of deep leaning. Anomaly detection is an identification of instances, events or observations, which do not conform to an expected pattern or other instances in dataset.

To perform AD tasks, we extract the discriminative features and efficient coding in the latent space. Information from detection structure is encoded using convolutional-based schemes and later is mapped in the low-dimensional space. We propose several architectures to extract features of normal data for AD that encode the information according to multiple learning strategy. Our models are based on autoencoder, which can learn disrciminative information in encoder-decoder pipeline and a tight boundaries is set for normal data.

For the one-dimensional data, we proposed vector-based convolutional autoencoder (V-CAE) for one dimensional anomaly detection. Given the good performanceof convolution network on matrix data, it is promising to transform one-dimensional data into a matrix form and learn important relationship features by convolutional network. The core of this model is a linear autoencoder, which is used to construct a low-dimensional manifold of feature vectors for normal data. At the same time, we used vector-based convolutional neural network (V-CNN) to extract the features from vector data before and after the linear autoencoder (fully connected autoencoder with F-norm reconstruction error and linear activation function) that makes the model learn deep features for efficient anomaly detection. The V-CNN is used to extract non-linear feature vector from the input vector by 2-D convolutional neural network. In this study, we used input data as a vector form and only the features extracted from the normal input data are used to train our proposed model.

However, the reconstruction error and abnormal score introduced in the aforementioned studies used to tune the threshold only for latent sampled variables, and as a result, such methods reported poor reconstruction performance in the abnormal data. Unlike in the previous studies on applying CVAE to anomaly detection in which the intention of variational autoencoder (VAE) deviated from learning an acceptable pattern for anomalous data identification. As for AD classification through the normal data, examining the latent representation is promising and effective. Thus, we introduce maximization of mutual information (MMI) regularization that help in low-dimensional representation of learned features to emerge. The proposed convulutional variational autoencoder (CVAE) is optimized by combining the representations learned across the three different objectives targeted at MMI on both local and global variables with the original training objective function of Kullback-Leibler divergence distributions. This feature leads to the losses of information about the input data distribution and mapping to the prior probability distribution, thereby generating the output with the high false positives. Therefore, the application of VAEs to anomaly detection tasks needs to be facilitated by adding suitable regularization techniques. In the present study, we investigate the possibility to address this issue by regularizing multiple discriminator spaces aiming to estimate how precisely the output matches its input, rather than relying only on the encoder latent space.

Though, the above latent-based methods for AD achieved better results for vector datasets, it is not ideal for matrix datasets. Furthermore, those methods considered detection using only one latent space and did not consider the possibility of a mixture of low-dimensional nonlinear manifolds of multiple latent spaces. Linearly combining different manifolds in latent spaces can generate best latent representation. However, most of the existing AD methods solely based on the reconstruction errors or latent representation using a single low-dimensional manifold are often not ideal for the image objects with complex background. In this study to realize the promise of multi-manifold latent information for AD, we propose a mixture of experts ensemble with two convolutional variational autoencoders (CVAEs) and convolution network (MEx-CVAEC) which explicitly learns manifold relationships of data that make use of multiple encoded detections. In addition, in order to enhance the model detection performance, we re-encode the output of the CVAE by generating a new data manifold for AD. Thereby each expert is developed to comprise an encoder-decoder-

encoder pipeline (EDE) based on CVAE. Additionally, we use a tower structure in the mixture-of-expert model to assign a latent score to each latent representation.

Inspired by multi-space detection in autoencoder, orthogonal projection is introduced to capture the null subspace that consists of noisy information for AD, which is explicitly ignored in the existing approaches. The exploration of double subspaces, called normal space (NS) and abnormal space (AS) can improve the discriminative manifold information. All these insights have a direct application to the low-dimensonal representation of latent space in autoencoder-based methods for the field of anomaly detection. The range subspace and null subspace are two subspaces of the original space decomposed by their direct sum. To comprehensively exploit the manifolds in two subspaces for robust AD, in this study, we propose an autoencoder framework based on an orthogonal projection constraints (OPC) learning method. The primary objective involves the calculation of projected norms in the range and null subspace. By constraining the projection operator to approximate the orthogonal projections, the model can be trained in an end-to-end manner via BP. In the proposed autoencoder framework model, the features are firstly extracted from the raw input and projected into the subspaces by projection operator.

Compared with the state-of-the-art methods, the proposed methods achieve the best performances, which demonstrates the effectiveness and robustness of anomaly detection using the autoencoder-based method. In the future, we will try to redesign the discrimination objective of the generator to further enhance the generator'sability to recognize anomalies.