

Appendix for CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows

A. Relationship with the flow framework

The loss function for the reverse $D_{KL} [p_Z(z; \cdot) \| p_Z(z)]$ objective [1], where $p_Z(z; \cdot)$ is the model prediction and $p_Z(z)$ is a target density, is defined as

$$L(\cdot) = E_{p_Z(z; \cdot)} [\log p_Z(z; \cdot) - \log p_Z(z)]: \quad (5)$$

The first term in (5) can be written using (4) definition for a standard MVG prior ($u \sim N(0; I)$) as

$$\log p_Z(z; \cdot) = \log(2\pi)^{-D/2} E^2(u) = -2 + \log | \det J |; \quad (5.1)$$

where $E^2(u) = \|u\|_2^2$ is a squared Euclidean distance of

Similarly, the second term in (5) can be written for MVG density (2) using a square of Mahalanobis distance as

$$\log p_Z(z) = \log(2\pi)^{-D/2} + \log \det \Sigma^{-1/2} - \frac{1}{2} M^2(z); \quad (5.2)$$

By substituting (5.1-5.2) into (5), the constants $\log(2\pi)^{-D/2}$ are eliminated and the loss is

$$L(\cdot) = E_{p_Z(z; \cdot)} \left[\frac{M^2(z)}{2} - \log | \det J | \right]; \quad (6)$$

B. CFLOW decoders for likelihood estimation

We train CFLOW-AD using a maximum likelihood objective, which is equivalent to minimizing the forward D_{KL} objective [1] with the loss defined by

$$L(\cdot) = D_{KL} [p_Z(z) \| p_Z(z; c; \cdot)]; \quad (7)$$

where $p_Z(z; c; \cdot)$ is a conditional normalizing flow (CFLOW) model with a condition vector $c \in \mathbb{R}^C$.

The target density $p_Z(z)$ is usually replaced by a constant because the parameters do not depend on this density during gradient-based optimization. Then by analogy with unconditional flow (4), the loss (7) for $p_Z(z; c; \cdot)$ can be written as

$$L(\cdot) = E_{p_Z(z)} [\log p_U(u) + \log | \det J |] + \text{const} \quad (7.1)$$

In practice, the expectation operation in (7.1) is replaced by an empirical train dataset $\mathcal{D}_{\text{train}}$ of size N . Using the

definition of base distribution with $p_U(u)$, the final form of (7) can be expressed as

$$L(\cdot) = \frac{1}{N} \sum_{i=1}^N \left[\frac{\|u_i\|_2^2}{2} - \log | \det J_i | \right] + \text{const} \quad (7.2)$$

where the random variable $u_i = g^{-1}(z_i; c_i; \cdot)$ and the Jacobian $J_i = \partial_z g^{-1}(z_i; c_i; \cdot)$ depend both on input features z_i and conditional vector c_i for CFLOW model.

References

- [1] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, 2021.