

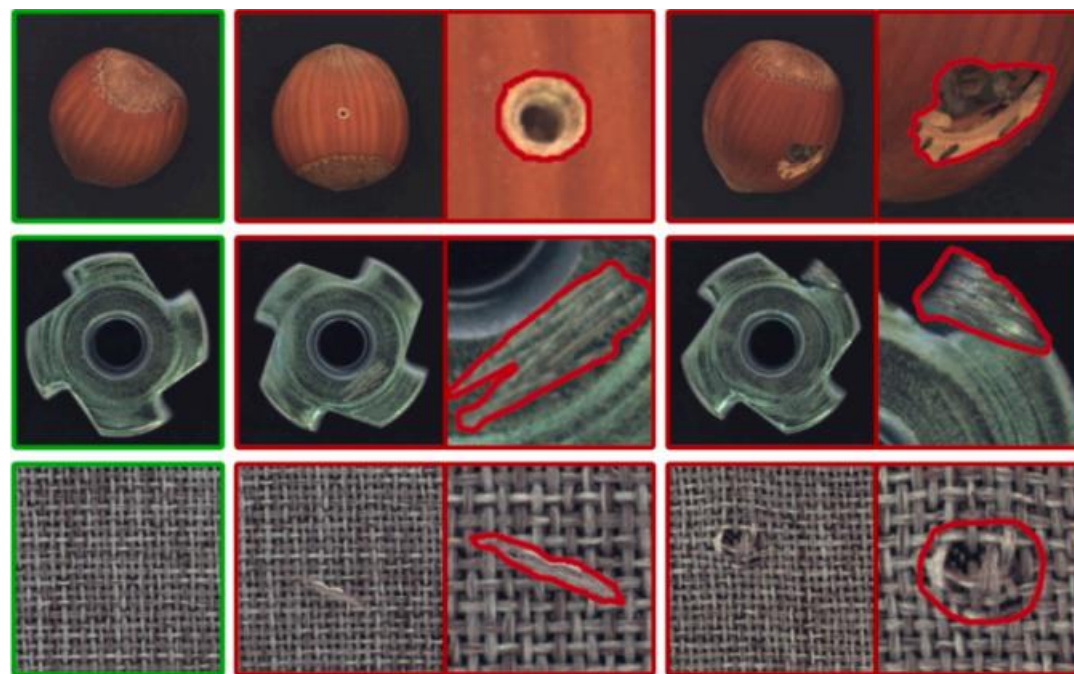
Unsupervised Anomaly Detection based Variational Autoencoder

Le Van Cong

異常検知とは

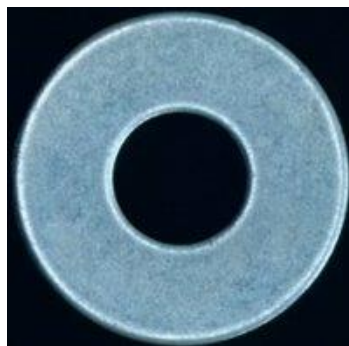
1. ある画像のよごれや傷などの異常を検知する (一般的な異常検知)

- 今回紹介する論文はこちらです.

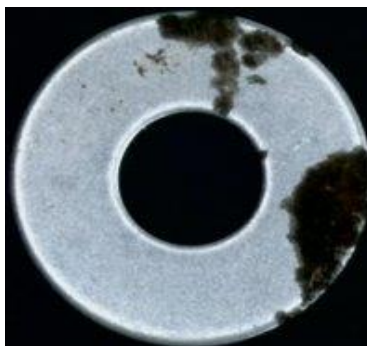


異常検知とは

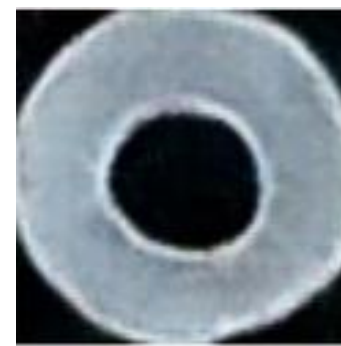
- 正常データのみを使ってAE(オートエンコーダ)ベースのモデルの学習を行ったあと、推論時に損失関数の勾配を利用することで、異常データがAEによって得られた正常データの多様体の最も近いところにマッピングされるように再構成し、その差をとることで異常箇所の特定を行う



正常



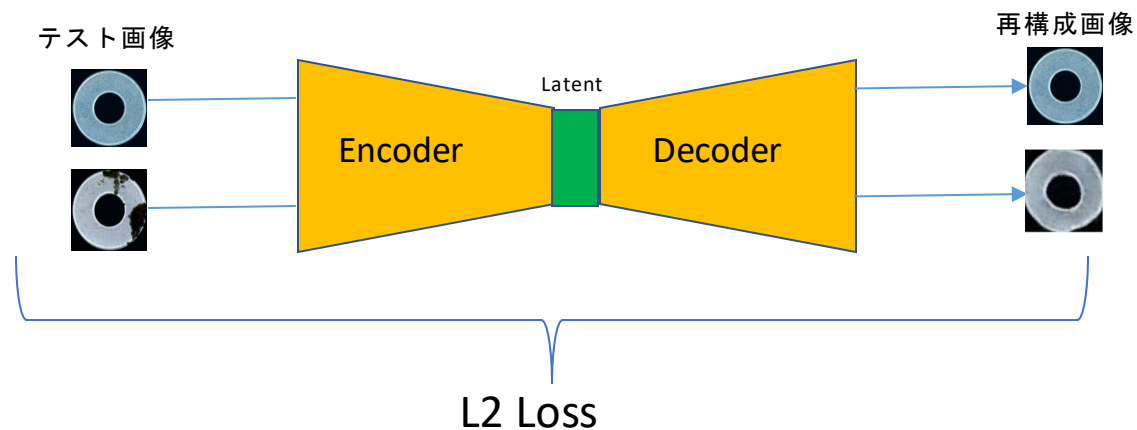
異常



再構成画像

背景知識: 異常部位特定の一般的な手法

- オートエンコーダベースの生成モデルは正常データの多様体を学習すると仮定
- テスト時に異常データが入力されると、異常データは訓練時がないので、元画像と最も近くなる正常データを出力する
- 元画像と再構成画像の差分を利用して異常部位を特定する



関連手法: VAEの下限を使った異常検知

- ある閾値Tを決めて、VAEの下限を上回れば異常とする方法

$$\begin{aligned}\log p(\mathbf{x}) &= \log \mathbb{E}_{\mathbf{z} \sim q(\mathbf{z}|\mathbf{x})} \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q(\mathbf{z}|\mathbf{x})} \\ &\geq \mathbb{E}_{\mathbf{z} \sim q(\mathbf{z}|\mathbf{x})} \log p(\mathbf{x}|\mathbf{z}) - D_{\text{KL}}(q(\mathbf{z}|\mathbf{x}) \| p(\mathbf{z})) = -\mathcal{L}(\mathbf{x})\end{aligned}$$

- しかし, [Matsubara+ 2018]によれば, KL項は良くない影響を及ぼすため, 再構成項だけを利用することを考えている

$$\mathcal{L}_r(\mathbf{x}) = -\mathbb{E}_{\mathbf{z} \sim q(\mathbf{z}|\mathbf{x})} \log p(\mathbf{x}|\mathbf{z})$$

- まあ結局, VAEの再構成もぼやけるため, 単純にVAEを使うのも微妙
 - もちろん再構成がぼやけないUnetのようなオートエンコーダを使っても, 多様体学習しないからダメ(傷も復元されてしまう)

結論

- リアルな正常画像を生成できている
- 画像に大きな異常を検出することができます
- 画像内の大きな異常と小さな異常の両方を検出できるように、さらにモデルを調査する必要があります