

Stable Diffusion Web UI を用いた画像拡張による工業材料の 欠陥検出用 CNN モデルの性能改善

永田研究室 F120035 ZHENG ZHELIN

1. 目的

深層学習においては最適化アルゴリズムが不均衡な量のデータを処理する場合、多数派のカテゴリにほとんどの時間が費やされ、少数派のカテゴリの学習時間が少なくなってしまうという傾向がある。このため、この問題を放置しておくとう学習後のモデルは少数カテゴリを多数カテゴリとして誤分類してしまう可能性が高まってしまふ。工業製品によっては欠陥や不良品の発生頻度が非常に低いためにそれらの画像を収集することが容易ではなく、このような訓練データの不均衡といった状況が発生する。結果的に CNN や SVM 等の多次元データの分類モデルによる欠陥検出システムを構築しようとした場合、汎化性の高いシステムの実現が困難となる場合が少なくない。本研究では、不良品の発生頻度の低い工業材料を対象とし、Stable Diffusion のインペインティング機能を応用して欠陥が含まれた画像の高品位な拡張を行うことで欠陥検出用 CNN モデルの汎化性能の向上を試みたので報告する。

2. 研究内容

今回対象とした工業材料の画像サイズは 2590×1942 ピクセルであったが、Stable Diffusion では内部処理上の制約により、8 の倍数の画像サイズのみ生成が可能となっている。このため、対象となる材料部分(Target area)が含まれている 1288×1288 の領域(Cropping area)をオリジナル画像から切り取り、この対象領域での画像拡張を行った。まず、単一の RGB 値 (0, 242, 246) を持つピクセルで元のサイズまで背景部分を埋めることでオリジナル画像のサイズに復元した。実験では、Stable Diffusion で拡張した画像を CNN の訓練に用いることが汎化性能にどのような影響及ぼすのかを評価するために、二つのデータセットを用意した。データセット A はオリジナル画像の良品: 156, 不良品: 196 のみで構成し、データセット B はデータセット A に対して Stable Diffusion で拡張した画像を加えた良品: 361, 不良品: 475 で構成した。図 1 にはオリジナル画像と Stable Diffusion で再構築された画像の Target area の重要部分(先端部分)における差分画像を示す。オリジナル画像に近い画素値のところは白に近く、異なったところは黒に近くなるように差分図を出力させている。この結果から、Stable Diffusion を用いることで範囲指定されたエリアに対してのみ、本物の欠陥に非常によく似た特徴を持つ新たなパッチ画像を再構築できることが確認された。

3. 結果

データセット A とデータセット B で訓練した CNN モデルはそれぞれ、 CNN_A と CNN_B とした。 CNN_A と CNN_B をそれぞれ 5 回ずつ試行的に訓練し、分類結果のベスト精度と平均精度を比較したところ、 CNN_B の精度は CNN_A のそれよりそれぞれ、0.85% および 1.42% 高い結果となった。このように生成系 AI の Stable Diffusion を用いることで、画像内の指定領域に対して望ましい特徴を持つパッチ画像を組み込むことが可能になり、さらに、生成された画像を用いて画像拡張を行い、CNN モデルの訓練に用いることで分類精度の向上も確認できた。しかしながら、今回使用したデータセットの場合、表 1 と表 2 の混同行列から CNN_B の再現率は逆に下がっていることが確認されたため、今後はこの再現率の改善に取り組むたい。

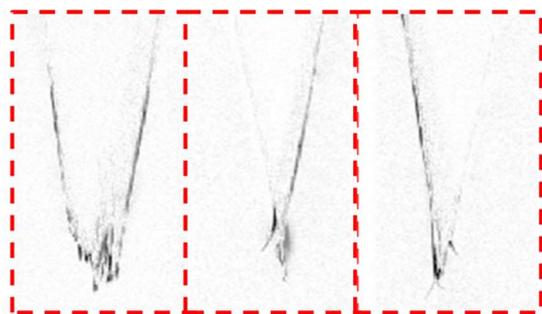


Fig. 1 Examples of subtraction between original images and augmented images.

Table 1 Classification result by CNN_A model.

	Predicted	Anomaly (NG)	Normal (OK)
True			
Anomaly (NG)		145	11
Normal (OK)		8	188

Table 2 Classification result by CNN_B model.

	Predicted	Anomaly (NG)	Normal (OK)
True			
Anomaly (NG)		150	6
Normal (OK)		10	186