

Stable Diffusion を用いた画像拡張による欠陥検出モデルの性能向上

永田研究室 F121056 松富尊之

1. 目的

近年、製品の検査工程においては深層学習の技術画像認識に特化させた畳み込みニューラルネットワーク (CNN) を応用した不良品検査システムの普及が進んでいる。しかし、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収集が容易ではなく、結果的に、CNN を用いた欠陥検出システムを実装しようとした場合、汎化性の高いシステムの構築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製品、さらに従来の左右反転や上下反転などの画像拡張法を適用できない製品を対象とし、Stable Diffusion の Image-to-Image 機能を応用して欠陥が含まれた画像を拡張し、それらを用いて訓練させることで欠陥検出モデルの性能向上を試みたので報告する。

2. 研究内容

本研究では、まず、錠剤の欠陥検出に CNN と FCDD を適用し、性能を評価した。まず、オリジナル画像（良品 218 枚、不良品 90 枚）をもとに最大エポック数 200、ミニバッチサイズ 32、学習率 0.0001 にて両モデルを訓練後に、テスト用データ（良品 57 枚、不良品 70 枚）の分類実験を行った。次に、欠陥検出率が悪かった特徴を持つ画像を Stable Diffusion の Image-to-Image 機能の batch オプションを用いて拡張した。このオプションは、拡張したい画像の入っているフォルダを選択するとフォルダに入っているすべての画像を設定したパラメータで拡張することができる。さらに、Inpaint batch mask directory に変化を加えたくない場所を白以外の色でマスクした画像を配置することで、より目標に近い類似画像を生成することができる。設定パラメータは Sampling を 50、Sampling method を DPM++SDE、Photo size を 800×800、Denoising strength を 0.2 に設定し、画像拡張を実行することで訓練用の不良品画像を 113 枚増やした。その後、拡張後の訓練用データ（良品 218 枚、不良品 203 枚）をもとに最大エポック数 200、ミニバッチサイズ 32、学習率 0.0001 にて両モデルを訓練後に、さきほどと同じテスト用データ（良品 57 枚、不良品 70 枚）の分類実験を行い、性能を評価した。

3. 結果

オリジナル画像で CNN を訓練した場合、表 1 のように不良品画像 65 枚が誤分類され、分類精度は 48.41% であった。FCDD の場合は、良品 2 枚、不良品 10 枚が誤分類され、分類精度は 90.47% であった。color, scratch, crack の欠陥検出率が悪かったため、この 3 つの特徴を持つ画像を Stable Diffusion を用いて拡張した。図 1 はオリジナル画像と Stable Diffusion で生成した画像の差分画像の例を示す。オリジナル画像と同じ画素値のピクセルは白で、異なる画素値のピクセルは黒で示すように差分図を出力しているが、欠陥 crack の特徴を持ちオリジナル画像と似ているものの、画素値レベルでは異なる領域が良好に生成されていることがわかる。拡張後のデータセットで CNN を訓練した場合、表 2 のように良品 13 枚、不良品 13 枚が誤分類され、分類精度は 79.36% であった。拡張前と比較すると分類精度は約 31% 向上した。FCDD の場合、良品 1 枚、不良品 1 枚が誤分類され、分類精度は 98.41% であった。拡張前と比較すると分類精度は約 8% 向上した。以上の結果から、不良品の発生頻度の低い製品、さらに従来画像拡張法を適用できない製品に対して、Stable Diffusion の Image-to-Image 機能を応用して欠陥の特徴が含まれた画像を拡張し、それらを用いて訓練させることで欠陥検出モデルの性能を向上できることが確認された。



Fig. 1 Example of subtracted image.

オリジナル画像と同じ画素値のピクセルは白で、異なる画素値のピクセルは黒で示すように差分図を出力しているが、欠陥 crack の特徴を持ちオリジナル画像と似ているものの、画素値レベルでは異なる領域が良好に生成されていることがわかる。拡張後のデータセットで CNN を訓練した場合、表 2 のように良品 13 枚、不良品 13 枚が誤分類され、分類精度は 79.36% であった。拡張前と比較すると分類精度は約 31% 向上した。FCDD の場合、良品 1 枚、不良品 1 枚が誤分類され、分類精度は 98.41% であった。拡張前と比較すると分類精度は約 8% 向上した。以上の結果から、不良品の発生頻度の低い製品、さらに従来画像拡張法を適用できない製品に対して、Stable Diffusion の Image-to-Image 機能を応用して欠陥の特徴が含まれた画像を拡張し、それらを用いて訓練させることで欠陥検出モデルの性能を向上できることが確認された。

Table 1 Before image augmentation.

CNN			FCDD		
予測 正解	OK	NG	予測 正解	OK	NG
OK	56	0	OK	54	2
NG	65	5	NG	10	60

Table 2 After image augmentation.

CNN			FCDD		
予測 正解	OK	NG	予測 正解	OK	NG
OK	43	13	OK	55	1
NG	13	57	NG	1	69