

Occlusion Sensitivity による CNN の判断根拠可視化時の Mask Size と Stride の効果

永田研究室 F120049 松浦 一史

1. 目的

様々な工業製品の検査工程においては目視検査に依存している部分が多くあり、検査員の作業負荷や製品の検出誤差など生産性の低下を招くという問題がある。このような問題を解決するため、深層学習の技術の中でも画像認識に特化した畳み込みニューラルネットワーク(CNN)を製品の欠陥検出に応用しようとする試みがなされている。また、CNNの判定根拠の可視化方法としてはGrad-CAMが提案されており、欠陥部分を可視化できているものの、明確に捉えることができていないという課題がある。本研究では、CNNによる画像分類において画像のどの部分が重要であるかを理解するためのもう一つの手法であるOcclusion Sensitivityを利用して判断根拠の可視化性能を評価する。

2. 研究内容

Occlusion Sensitivityを最適に利用するためには、画像内の一部を隠すマスクサイズMask Sizeとそのラスタースキャンに伴う移動量であるStrideを適切な値に調整する必要がある。本研究では、メーカーから提供された2590×1942ピクセルの良品234枚、不良品268枚の工業製品画像の中から、トレーニング用データとして良品187枚、不良品214枚、テスト用データとして良品47枚、不良品54枚を用いる。このとき、解像度をオリジナルの2590×1942の中心から1942×1942のサイズでトリミング後にNearest-neighbor法、Bilinear法、Bicubic法の補間方法を適用し224×224にダウンサイジングしたものをそれぞれ、データセットA1, B1, C1とする。これらのデータセットを用いて訓練したCNNモデルをそれぞれCNN_A1, CNN_B1, CNN_C1とする。同様にして、解像度を300×300, 512×512にダウンサイジングし、それぞれをデータセットA2, B2, C2及びA3, B3, C3とし、それらで訓練したCNNモデルをCNN_A2, CNN_B2, CNN_C2, 及びCNN_A3, CNN_B3, CNN_C3とする。なお、これらのCNNモデルは研究室オリジナルのsssNetをもとに入力層の解像度を変えて設計している。良品、不良品の2クラスの分類実験終了後、不良品と判定された製品画像の欠陥部分の根拠を可視化するためにGrad-CAMとOcclusion Sensitivityを適用し、それぞれの可視化性能を比較し、評価する。

3. 結果

9つのCNNモデルのテスト画像に対する汎化性能を比較したところ、CNN_C3が最も良かったため、このモデルを用いた時の可視化結果により考察した。図1、図2及び図3は、CNN_C3を用いて不良品のテスト画像を分類した際に不良品と判定され、その根拠となった領域をGrad-CAMとOcclusion Sensitivityで可視化したものである。このとき、CNN_C3のテスト画像に対する汎化性能は、良品100.0%、不良品83.3%の精度であった。Grad-CAMは欠陥部分以外にも注視領域として捉えているが、Occlusion Sensitivityは、Mask SizeとStrideを調整することで欠陥部分を明確に注視できている。Mask SizeとStrideの値をさまざまな値に変更し試行錯誤的に調整することで、最も欠陥部分を注視できている関係値を決定したところ、Mask Size 45, Stride 21となった。Mask SizeとStrideの値は試行錯誤により得られた結果であったため、今後は入力画像の解像度や画像内の欠陥部分の大きさなどからシステマティックに算出できるような関係を見つけていきたい。また、不良品と判断された画像の中で糸片のような目立たない欠陥に対しても可視化できるような方法について検討していきたい。

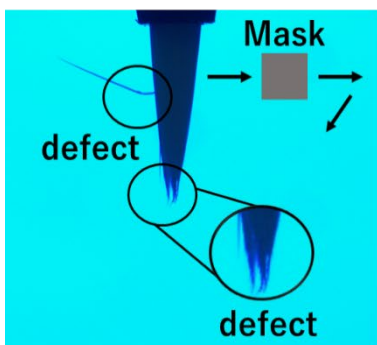


Fig. 1 Input Image (512×512)

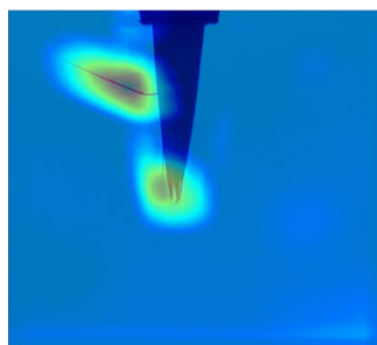


Fig. 2 Grad-CAM

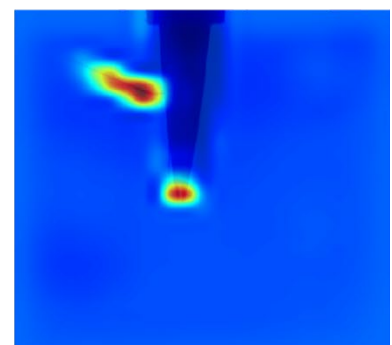


Fig. 3 Occlusion Sensitivity