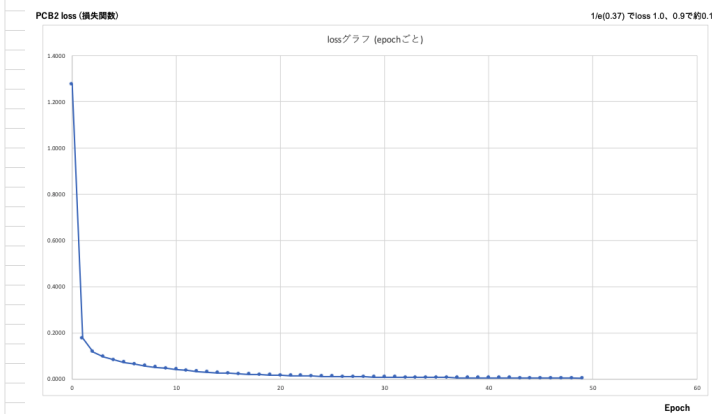


評価報告書内容(PCB2)

課題: 以下公開データセットを使って、不良箇所をピクセル単位で判定したい
 データセット: VisA: Visual Anomaly dataset 参照: <https://github.com/amazon-science/spot-diff>
 サブセット: PCB2: プリント基盤(Printed Circuit Boards) #2
 不良品クラス: 4クラス (1: Bent曲がり, 2: Melt溶解, 3: Scratch引っかき, 4: Missing欠損)
 解析タスク: セマンテック セグメンテーション

学習データ	総数	うち学習用	うち検証用	推論テスト	
良品数	1,001	182	84	0	*単位: 画像枚数
不良品数	100	84	16	16	良品、不良品のデータ不均衡を考慮し、バランシングのためランダムサンプリングを実施
合計	1,101	266	100	16	*ランダムサンプリング実施データ
画像サイズ	1404 x 1070	2.17 : 1 の割合 5.25 : 1 の割合			*不良品数が少ないため検証用と同一データを使用
ラベリング	グレースケール画像 (クラスIDでそのまま表現 0が不良品でない背景部分)				*アノテーションという呼称が一般的
入力画像サイズ	512 x 512 (RGB)。スライド384 (1/4である128pixel分を再利用するため)				
前処理	中心から1344 x 1024にcrop。(データセットの特性上、画像周辺の予測は不要であり、オリジナル画像サイズから64で割り切れない分をcrop)				
パラメータ調整	loss関数、Upsampling方式、転移学習有無				
最終Epoch数	50 (1 Epoch 約15分)				
予測モデル評価	*訓練された学習モデルの事				
使用データ	計100 (うち不良品16)				
採用モデルEpoch数	50 10Epochごとの評価比較				
lossグラフ	*学習の進み具合を知る簡易的な予測モデル性能の評価指標 過学習しているかどうかについてはこの数字では判断できない				



2クラス評価 (良品/不良品)

		予測			
		Defect (Positive)	BG (Negative)		
GT	Defect	58342	39902	*単位は画素数	
	BG	28732	157159424	FN: 見逃し数	*混合行列 (Confusion Matrix)
				FP: お手付き数	予測モデルの性能評価指標として一般的に用いられる。 縦軸視点が予測の内訳。横軸が教師データ内訳となる。 *GT: 教師データを指す(Ground Truth)

*1画像から6枚分の512x512データ作成。計600画像分のpixel数内訳となる。

IoU	0.4595	TP	TP+FP+FN
		58342	126976

5クラス評価(不良品クラス毎)

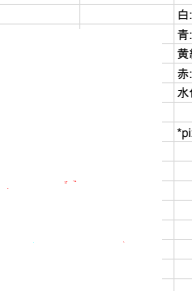
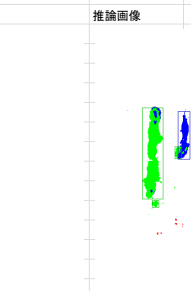
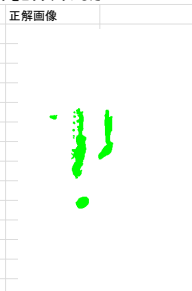
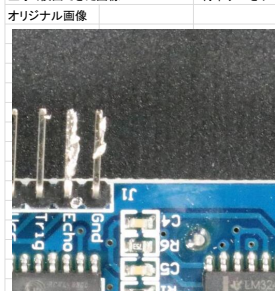
		予測				
		Bent	Melt	Scratch	Missing	BG (Negative)
GT	Bent	8309	2729	0	0	8968
	Melt	951	18598	0	1076	20618
	Scratch	0	0	2958	0	2589
	Missing	0	0	149	23572	7727
	BG	1566	7554	5921	13691	157159424
IoU	0.4208	TP	TP+FP+FN			
		53437	126976			

予測テスト

平均推論速度 (推論後の画像解析(領域化等)の時間は含めない)
 CPU: 0.48秒/画像 (約2FPS)
 GPU: 0.0422秒/画像 (約24FPS) ...GeForce RTX 2080 Ti ...GPU1枚でも4枚でも差はなかった --> 作り方の問題か?
 上手く不良品を検出できたもの、できなかった評価データからピックアップして予測(推論)を実施し、その結果を紹介する。

上手く検出できた画像1

*特筆すべきデータをピックアップした



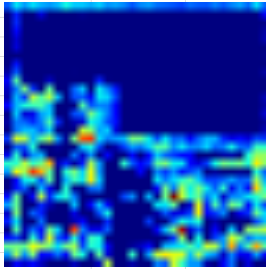
白: 背景
 青: Bent
 黄緑: Melt
 赤: Scratch
 水色: Missing

*pixel領域100以上に対して矩形表示

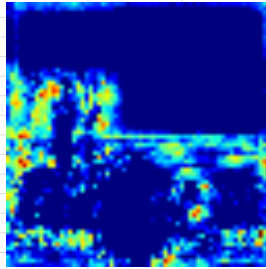
ヒートマップ

* Grad-CAMを使用。予測モデルが注目している領域ほどより赤(暖色)で反応

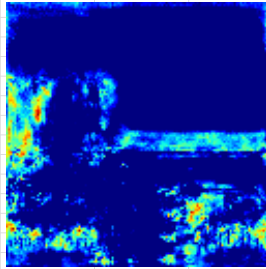
層1 (解像度32 x 32)



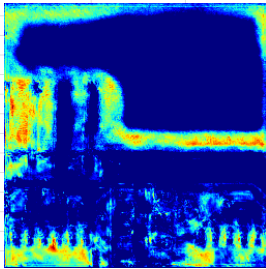
層2 (解像度64 x 64)



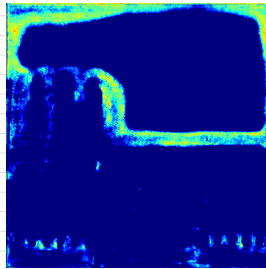
層3 (解像度128 x 128)



層4 (解像度256 x 256)

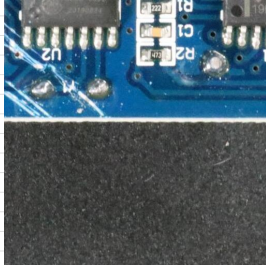


層5 (解像度512 x 512)



上手く検出できた画像2

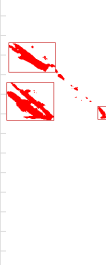
オリジナル画像



正解画像



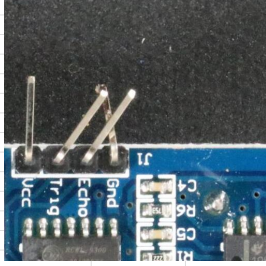
推論画像



ヒートマップは割愛

上手く検出できた画像3

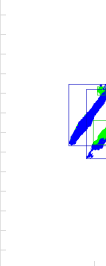
オリジナル画像



正解画像

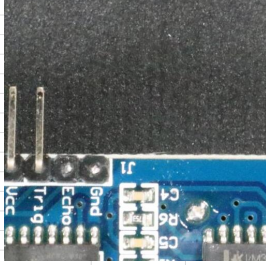


推論画像

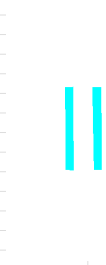


上手く検出できなかった画像

オリジナル画像



正解画像



推論画像



所見

局所的な認識は優勢だが、大局的な認識は比較的精度が出なかったのかなと思われる。
 ただ、経験的にデータ不良不均衡を抑えつつ、学習データ数が増えれば、全体的な精度はもっと上がると思われるため、このぐらいのデータ数においては、
 順当な学習結果だったと考えられる。最新予測モデルを採用し更なる精度向上を目指すという選択肢もあるが、最新のものほど複雑でありモデル構築コストも無視できない量となるため、
 それよりも学習に寄与できそうなデータを収集したり、外観検査システムとしての予測(推論)後の画像解析処理の方を工夫する方が、現実的な解決方法となろう。
 従って、この内容を以って今回のPoC検証作業を完了とする。