

モデルの不確実性を考慮した外観検査の効率化

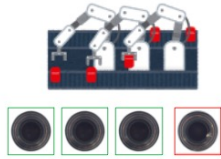
Improving Industrial Inspection Efficiency by Using Model Uncertainty

太田 真人 株式会社 電通国際情報サービス

要約

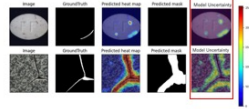
検査員は異常検知の結果をいつ信頼するのか？AI導入して本当に業務効率化になるのか？

検査対象の撮影



- 正常画像は200枚程度
- 異常画像は評価用に数枚

教師なし異常検知 予測の信頼性の定量化



- 異常検知手法にはPaDiM [Defard 21]を使用
- 不確実性の定量化手法に深層アンサンブルを使用

目視検査対象の選定

予測結果		自信あり		自信なし	
正常	異常	正常	異常	正常	異常
500枚	4枚	100枚	5枚	100枚	5枚

目視検査対象

- 不確実性を用いて人とAIの協調を目指す
- 不確実性を用いて偽陰性候補を検出する指標を提案

検査員による 目視検査の効率化



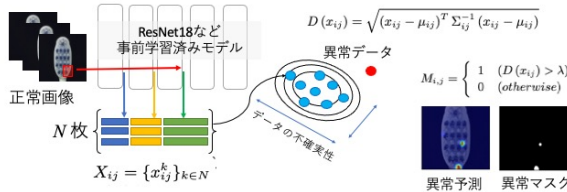
AIの正常は本当に正常か？
異常の見逃しは無くしたい
結局、目視検査数が多い

本手法適用後

目視検査後の最終精度の改善
見逃し候補を中心に検査
目視検査数が削減

関連研究

PaDiM：特徴量ベースな教師なし異常検知手法 [Defard 21]



1. バッチごとの正常データの特徴量から多変量ガウス分布を計算
2. バッチごとの異常データの特徴量とのマハラノビス距離を計算
3. 閾値以上の場合、バッチ部分を異常と定義

2種類の不確実性

データの不確実性：データノイズ・複雑さによるもの
各パッチの特徴量分布の共分散行列
撮影方法を変えると不確実性が減少

モデルの不確実性：モデルの表現力によるもの（予測の不一致）
予測に一貫性がないと不確実性が高くなる
深層アンサンブルによる予測分散

提案手法

- 人の目視検査対象画像を絞りつつ、偽陰性を減らす方法
- ✓ 人がモデルの不確実性の高い画像を検査
- ✓ 偽陰性画像候補を中心に選別する指標を提案

モデルの不確実性の定量化

- 深層アンサンブル [Lakshminarayanan 16]
- 学習済みモデルを利用するため、計算コストが低い

各パッチの予測分散

$$\sigma_{ij}^2 = \frac{1}{K} \sum_{k=1}^K (M_{ij}^k - \bar{M}_{ij})^2, (i, j) \in [1, W] \times [1, H]$$

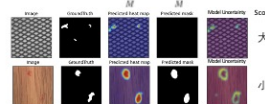
パッチ位置

\bar{M}_{ij} ：パッチ位置に対する平均予測マスク

目視検査対象の選別指標の提案

- 選別指標が閾値以上の画像を目視検査
- 指標の意味合い：画像に対する予測マスクの面積が小さく、予測分散の総和が大きいと検査対象となる

$$Score(n) = \frac{\sum_{(i,j)} \sigma_{ij}^2}{1 + \sum_{(i,j)} \bar{M}_{ij}}, n = 1, \dots, N$$



実験結果・考察

外観検査異常検知用 MVTec データセット[Bergmann 21]
モデル数 (K=5) ResNet-18, 34, 50 101, Wide-ResNet50-2を使用

- Q1：選別指標は偽陰性 (FN) を検出できるのか
- Q2：目視検査枚数を削減でき、最終検知性能は改善するのか

- A1：指標を用いて8/15クラスでFNを全て検出
3クラスのみ全てのFNを検出できず
- A2：既存手法に比べ、目視検査枚数を平均 32%削減
誤分類の検出率を平均 4%改善

Toothbrush クラスの目視検査前後の分類結果

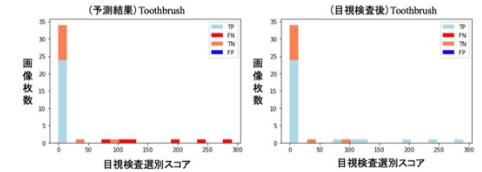


表 1: 実データによる目視検査効率化の評価

Type	Deep Ensembles (提案指標)		ResNet18 (A ± 0.1)	
	TPR/FPR	検出率 (FN, FP, N _o)	TPR/FPR	検出率 (FN, FP, N _o)
Carpet (117)	1.00/0.00	17.64% (6, 0, 37)	1.00/0.00	10.52% (6, 0, 57)
Grid (78)	0.98/0.00	31.94% (15, 0, 47)	0.96/0.00	16.00% (7, 1, 50)
Leather (124)	0.99/0.00	40.00% (10, 0, 25)	1.00/0.00	11.29% (7, 0, 62)
Tile (117)	1.00/0.00	27.50% (1, 10, 40)	1.00/0.06	48.97% (0, 24, 49)
Wood (79)	1.00/0.05	11.11% (2, 0, 18)	1.00/0.05	19.23% (0, 5, 26)
全テキスチャ	0.99/0.01	25.63% (6.8, 2.0, 38.4)	0.99/0.02	21.20% (4.0, 6.0, 48.8)
Bottle (83)	1.00/0.00	7.14% (0, 1, 14)	1.00/0.00	3.70% (0, 1, 27)
Cable (150)	1.00/0.29	44.89% (1, 21, 49)	0.98/0.38	66.66% (0, 32, 48)
Capsule (132)	1.00/0.08	58.82% (30, 0, 51)	1.00/0.08	14.66% (9, 2, 75)
Hazelnut (110)	1.00/0.22	6.66% (0, 2, 30)	1.00/0.22	19.14% (1, 8, 47)
Metal (115)	1.00/0.04	47.82% (0, 11, 23)	1.00/0.04	79.16% (0, 19, 24)
Fill (167)	1.00/0.53	37.50% (2, 7, 24)	1.00/0.34	36.36% (0, 16, 44)
Screw (160)	0.96/0.09	58.92% (64, 2, 112)	0.96/0.00	33.75% (42, 11, 157)
Toothbrush (42)	1.00/0.00	66.66% (6, 0, 9)	1.00/0.00	22.72% (5, 0, 22)
Transistor (100)	1.00/0.18	32.60% (0, 15, 46)	1.00/0.20	48.00% (0, 24, 50)
Zipper (151)	1.00/0.21	15.38% (1, 3, 26)	1.00/0.03	13.33% (2, 10, 90)
全オブジェクト	1.00/0.16	37.63% (10.4, 6.2, 38.4)	0.99/0.12	33.77% (5.9, 12.3, 58.4)

FN(FN)：偽陰性(偽陽性)画像の検査枚数 N_o：目視検査枚数

今後の展望

- 予測の不確実性による意思決定支援の検証
- 実際の目視検査における業務効率、負担、精度改善
- 不確実性を画像に表示することによる意思決定の変化調査