

# AnoGAN を用いた IC 外観画像の不良品検知

中岡 黎<sup>†1</sup> 木本 智幸<sup>†1</sup>

**概要:** 工業製品などの異常データが少ない場合、ディープラーニングを用いて良品と不良品を分類する識別モデルを構築することは困難な場合がある。そこで、正常データのみを使い不良品検知を行うことが求められる。本論文では、IC 画像からの異常検知問題においてディープラーニングモデルである AnoGAN を適用し不良品検知を行いその結果を示す。

**キーワード:** 深層学習, GAN, AnoGAN, 画像認識, 不良品判定

## Detection of defective products of IC external appearance image using AnoGAN

REI NAKAOKA<sup>†1</sup> TOMOYUKI KIMOTO<sup>†1</sup>

**Abstract:** When there are few abnormal data such as industrial products, it may be difficult to construct an identification model that classifies good and bad products using deep learning. Therefore, it is required to detect defective products using only normal data. In this paper, we apply an AnoGAN which is a deep learning model in the anomaly detection problem from the IC image and show the result of defective product detection.

**Keywords:** DeepLearning, GAN, AnoGAN, Image Recognition, Defective item determination

### 1. はじめに

IC の後工程では樹脂で IC 全体を包むモールド処理が行われる。モールドは樹脂を金型に流し込むことで形成するが、IC 全体に行き渡らないこともあり、コンピュータによる画像判断で不良品の検出がなされている。しかし、検出精度が 100%ではなく、最終的には人の目に頼っている。先行研究では、ディープラーニングの識別モデル CNN を用いて画像認識を行っていた事例があるが、良品と不良品の両方の画像が大量に必要であるためモデルを十分に学習できず、不良品の検出が困難な場合があった。そこで本研究ではディープラーニングの生成モデルである GAN を用いて良品のみの画像で不良品の検出を行うことを目指す。

### 2. GAN (Generative Adversarial Network)

GAN は観測データ  $\mathbf{x} \sim p_{data}(\mathbf{x})$  から、 $p_{data}(\mathbf{x}) = p_g(\mathbf{x})$  となるような確率分布  $p_g(\mathbf{x})$  を持つ疑似画像を得ることを目的としており、生成モデルである Generator と識別モデルである Discriminator を交互に学習することでこれを実現している [1]。Generator から出力されるデータの確率分布  $p_g(\mathbf{x})$  はランダムノイズ  $\mathbf{z}$  を入力とし関数  $G(\mathbf{z}; \theta_g)$  で疑似画像を生成する。Discriminator は入力が本物であるかを識別し、その確率  $D(\mathbf{x}; \theta_d)$  を返す。ここで、 $\theta_g$  と  $\theta_d$  はそれぞれの結合荷重などのパラメータを示している。Generator は

Discriminator を騙すために  $D(G(\mathbf{z}; \theta_g); \theta_d)$  が大きな値を取るように訓練する。逆に、Discriminator は観測データ  $\mathbf{x}$  に対して大きな値を取り、Generator が生成したデータに対して小さな値を取るように訓練する。これらの関係は以下の価値関数の minmax 問題で表現することができる。

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{data}} [\log(D(\mathbf{x}; \theta_d))] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z}; \theta_g)))]$$

### 3. AnoGAN (Anomaly Detection GAN)

AnoGAN では訓練データに正常データのみを使用する [2]。Generator は正常データを生成するように学習されるため、異常データを発生する元は存在しても潜在空間中で圧倒的少数であることは妥当な仮定である。つまり、正常データならば、それに対応する潜在空間  $\mathbf{Z}$  の元が存在するという過程の下で異常度 (Anomaly score) を定義する。しかし、GAN では図 1 のように観測空間から潜在空間への逆写像  $G^{-1}(\mathbf{x})$  は学習できない。そこで次の 2 つの損失 (Residual Loss と Discrimination Loss) を導入し、その損失が小さくなるように勾配法によって画像  $\mathbf{x}$  に対応する潜在変数  $\mathbf{z}$  を探索する。

<sup>†1</sup> 大分工業高等専門学校  
National Institute of Technology, Oita college

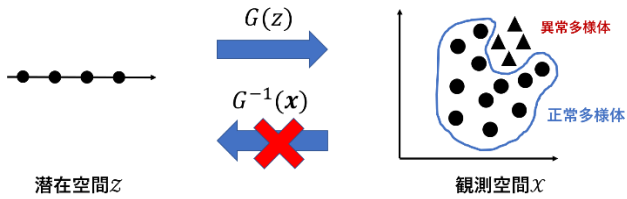


図 1 AnoGAN のイメージ

### (1) Residual Loss

Residual Loss は異常かを判別したい画像  $\mathbf{x}$  と Generator から生成された画像  $G(\mathbf{z})$  のピクセル差の合計値で、正常データであれば関数値は 0 に収束する。

$$\mathcal{L}_R(\mathbf{z}) = \Sigma |\mathbf{x} - G(\mathbf{z})|$$

### (2) Discrimination Loss

Discrimination Loss は異常かを判別したい画像  $\mathbf{x}$  と Generator から生成された画像  $G(\mathbf{z})$  の Discriminator の中間層の差の合計値である。ここで関数  $f$  は Discriminator の中間層の出力を表す。

$$\mathcal{L}_D(\mathbf{z}) = \Sigma |f(\mathbf{x}) - f(G(\mathbf{z}))|$$

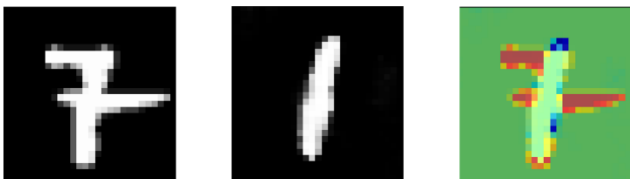
AnoGAN では上の 2 つの損失の加重和である以下の式を勾配法で最小化することにより、画像  $\mathbf{x}$  に近い画像を生成できる潜在変数  $\mathbf{z}_\gamma$  が求まる。ただし、 $\mathbf{z}_\gamma$  は勾配法で  $\gamma$  ステップ後の潜在変数を表し、最小化の対象は  $\mathbf{z}_\gamma$  であるため潜在変数を更新する際、Generator と Discriminator のパラメータは更新しない。

$$\mathcal{L}(\mathbf{z}_\gamma) = (1 - \lambda) \cdot \mathcal{L}_R(\mathbf{z}_\gamma) + \lambda \cdot \mathcal{L}_D(\mathbf{z}_\gamma)$$

そして、 $\Gamma$  回の更新後の値  $\mathbf{z}_\Gamma$  を用いて最終的な以下の異常度 (Anomaly score) を定義する。ここで  $R(\mathbf{x}) = \mathcal{L}_R(\mathbf{z}_\Gamma)$ 、 $D(\mathbf{x}) = \mathcal{L}_D(\mathbf{z}_\Gamma)$  である。

$$A(\mathbf{x}) = (1 - \lambda) \cdot R(\mathbf{x}) + \lambda \cdot D(\mathbf{x})$$

AnoGAN は元の画像と探索された潜在変数により再構成された画像の差分を取ることで、異常が発生した可能性が高い部位の確認が可能になる。図 2 は MNIST の 1 の画像を正常データとし、それ以外の画像を異常データとして異常検知を行なった例である。



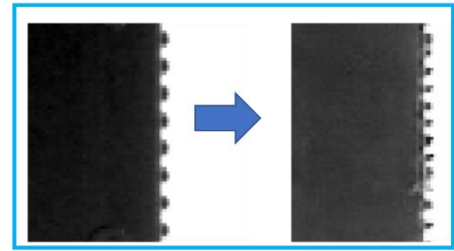
(a) 異常画像 (b) 再構成された画像 (c) (a) と (b) の差分

図 2 MNIST での異常検知

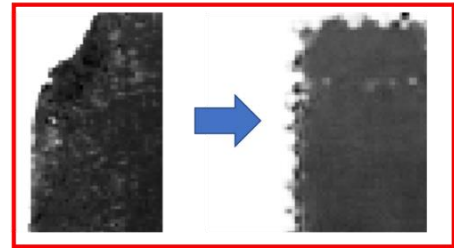
## 4. 実験

今回の実験では、IC の充填不良画像を AnoGAN により検出する。まず、64\*64 ピクセルのグレースケール画像である良品 3277 枚で DCGAN を学習させる[3]。その後、AnoGAN を用いて充填不良の 755 枚の画像の異常度を計算し、未学習の良品の画像 1100 枚の異常度との比較を行う。図 3 はそれぞれ良品と充填不良の元画像と Generator によって再構築された画像の比較である。ここで縦軸は頻度、

横軸は異常度  $A(\mathbf{x})$  の値を表している。図 3 より良品は再構築でき、充填不良は再構築できていないことがわかる。また、図 4, 5 はそれぞれ潜在変数  $\mathbf{z}$  の探索を 500 ステップ、100000 ステップ行った時の異常度のヒストグラムである。



(a) 良品画像



(b) 充填不良

図 3 元画像(左)と再構築画像(右)の比較

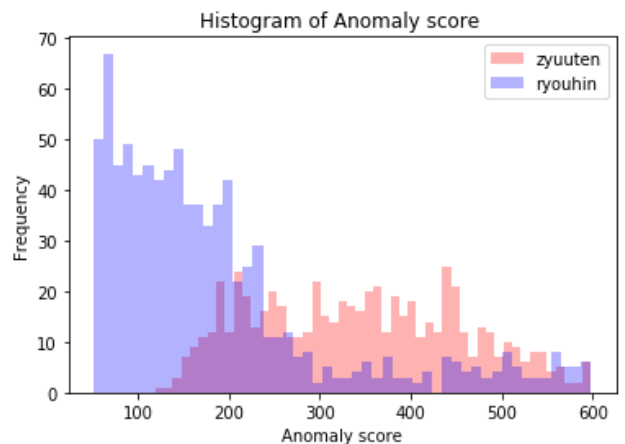


図 4 IC 画像の異常度の分布(500 ステップ後)

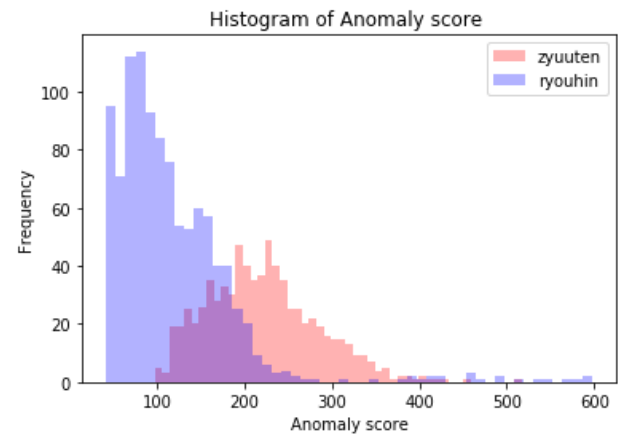


図 5 IC 画像の異常度の分布(100000 ステップ後)

図 4, 5 より充填不良の異常度の平均が良品の異常度の平均より大きいことがわかる。しかし、探索した潜在変数  $z$  により充填不良が再構築できてしまった場合や、良品を再構築できなかった場合などの原因によりそれぞれの異常度の分布に重なりが来ている。また、ステップ数を増やすことにより良品の再構築される枚数が多くなり、異常度の分布の重なりが多少ではあるが小さくなった。

## 5. まとめ

今回の実験では AnoGAN を IC 外観検査の不良品検知に用いた。実験結果より、充填不良の多くは潜在変数によって再構築できないことがわかり異常検知ができることが確認できた。しかし、充填不良が再構築されてしまった場合や良品を再構築できなかった場合もあった。そのため、良品の訓練画像枚数の増量、モデルのパラメーターチューニングなど行い検出精度を上げる必要がある。また、今回用いた AnoGAN は画像一枚あたりの異常度を計算するのに大変時間がかかる。今後は、計算時間の問題を回避した Efficient GAN を実装する必要がある[4]。

**謝辞** 本研究は、大分県 LSI クラスター研究開発事業費補助金の助成を受けて行われた。

## 参考文献

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems. (2014) 2672–2680
- [2] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In International Conference on Information Processing in Medical Imaging, pp. pp. 146157, 2017.
- [3] Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434 (2015)
- [4] Houssam Zenati, Chuan-Sheng Foo, Bruno Lecouat, Gaurav Manek, Vijay Ramaseshan Chandrasekhar.: EFFICIENT GAN-BASED ANOMALY DETECTION. arXiv:1802.06222v1 (2018)