

生成AIを用いた異常画像生成技術

Data Augmentation Method Using Generative AI

画像AI学習のための画像オーグメンテーション手法

様々な領域でディープラーニングを用いた機械学習やAIの実用化が進んでいる。AIモデルを学習するためには一定量のデータが必要となるが、学習データの希少性が原因で十分なAIモデルの学習ができない事例は多い。本稿では、画像生成AIと幾何学模様のマスク画像を組み合わせ、少ない画像からAI学習用のデータを拡張する方法について紹介する。



執筆者
先端技術応用研究所
情報技術グループ
追良瀬 利也

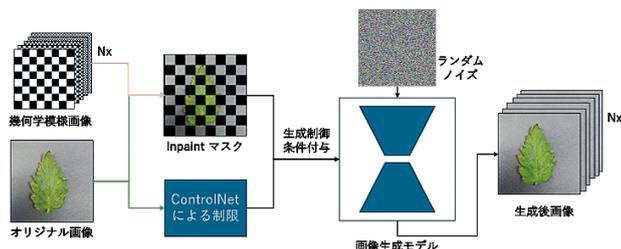
1 はじめに

ディープラーニングを用いたAI技術の進化に伴い、さまざまな産業でAIの活用が進んでいる。特にインフラや製造業では、カメラ画像を入力としてAIに画像内の異常を認識させる試みが盛んである。しかし、AIモデルを学習させるためには一定量のデータが必要であり、長期間の劣化による異常を画像で判別する場合などでは、十分な学習データを入手することが困難である。このようにデータが十分に得られない場合、少ないデータで学習を行うためにデータ拡張と呼ばれる、データに加工を加えて水増しする方法がよく用いられる。

データ拡張方法には、画像の回転や反転、明度の変更などがあるが、異常検出のようなタスクでは、現実世界の多様なパターンを網羅することが求められるため、これらの方法だけでは十分な判別精度を達成することが難しい。そこで近年発達した著しい画像生成AIを用いて、AIの学習のための画像データを生成する手法を研究している。本稿では、その概要について紹介していく。

2 提案手法概要

手法の概要を図1に示す。当該手法は、ある画像をもとに新たな画像を生成する、img2imgと呼ばれる画像生成手法と、幾何学模様によるマスクを核として構成されている。幾何学模様によりオリジナル画像をマスクし、マスクされていない部分を考慮しながら、マスクされた部分を中心に画像を再構成（生成）することにより新たな画像を作成する。生成までの流れを以下に示す。



第1図 手法の概要図

(1) 幾何学模様マスク画像の準備

オリジナル画像をマスクするためのマスク画像を用意する。図1では市松模様を例に示している。

(2) オリジナル画像の選択

拡張対象とするオリジナル画像、マスク画像を1枚選択する。

(3) 出力制限と画像生成

(1) (2) のマスク画像とオリジナル画像の組合せで、Inpaintと呼ばれる手法によりオリジナル画像にマスクを行う。この情報を画像生成モデルに入力し画像生成を行う。この際、生成画像が崩壊しないよう、ControlNet[1]と呼ばれる手法を用いて出力の制限を行う。

この手法のポイントは、同じオリジナル画像に対しても、幾何学模様のマスクを変更することで、隠される部分・隠されない部分が変化し、画像生成時の条件が容易に変更できることにある。これにより、1枚のオリジナル画像から、その特徴を持たせたまま「別の異常画像」をいくつも生成することが可能となる。

3 評価実験

(1) 実験概要

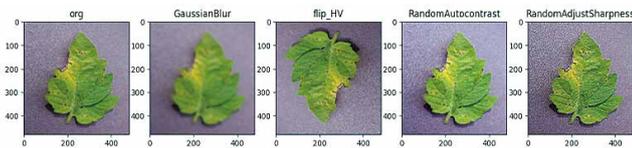
提案手法を用いて画像を生成し、その結果が画像判別モデルにどのような影響を及ぼすかを確認するため、実験を行った。この実験では、Kaggleのトマトの葉の病変分類データセット [2] から、Septoria leaf spot (白星病) およびEarly blight (輪紋病) クラスのデータを使用した。

本研究で目的とするのはデータ数が限られた場合でのAI精度向上であることから、元の訓練データセット（各クラス1000枚）から、それぞれ5, 10, 15, 20, 25枚を無作為抽出した実験用データセットを作成した。この実験用データセットに対して、①そのままのデータでモデルを訓練した場合②既存のデータ拡張手法により水増しした上でモデルを訓練した場合③提案手法によりデータ

拡張した上でモデルを訓練した場合とで比較を行った。学習対象のモデルはEfficientNetV2[3]と呼ばれるモデルを用いており、画像生成時の主要な条件は表1に示す通りである。また、②の比較対象のデータ拡張方法は図2の4種である。

第1表 画像生成時の主要条件

項目	内容
マスク模様	市松模様
マスク画像数	12枚（正方形1辺の長さ5,10,15,20,25px・白黒反転）
ControlNet 出力制限方法	CannyEdge
出力枚数	4枚
生成AIモデル	StableDiffusion v1.5
生成時プロンプト	leaf
生成画像サイズ	512*512



第2図 比較対象のデータ拡張方法4種（左からオリジナル・ぼかし・上下反転・コントラスト変化・鮮鋭化）

(2) 実験結果

実験結果を表2に示す。表の各数字は各条件で学習させたモデルの検証データにおける正解率を示している。表から、無作為抽出された画像枚数が5, 10, 15毎の場合提案手法が最も精度が高いことがわかる。また20, 25枚においては上下左右反転が最も精度が高く、提案手法は2番目に良い精度を示している。いずれの場合も、提案手法は拡張を何も加えない場合に比べて画像判別精度が向上しており、特にクラスごとのデータ数が少ない場合において、他のデータ拡張手法に比べ良い結果を示している。また、生成されたサンプルを図3に示す。元の画像の特徴を保ったまま、オリジナル画像では特に異常のない箇所が病変のように画像が変化していることが確認できる。



第3図 生成画像サンプル（左：オリジナル、右：生成後）

第2表 実験結果

	拡張無し	ガウシアン フィルタ	上下左右反転
5枚	0.74	0.77	0.768
10枚	0.806	0.844	0.816
15枚	0.874	0.893	0.889
20枚	0.865	0.881	0.894
25枚	0.895	0.896	0.905
	コントラスト 変更	鮮鋭化	提案手法
5枚	0.743	0.713	0.816
10枚	0.823	0.828	0.848
15枚	0.873	0.867	0.892
20枚	0.872	0.878	0.887
25枚	0.891	0.897	0.902

4 おわりに

本稿では、限られたデータ環境下でのAI活用を促進するための、画像生成AIを用いた異常画像の生成技術について紹介した。実験結果から、データが限られている際のAIの学習過程において、当該手法がAIの精度向上に寄与することが確認出来た。この技術は、これまでデータの不足が原因でAIの活用が進まなかった領域に対して、AIの開発・テストのためのデータを提供し、AI技術の活用の幅を広げられる可能性がある。

今後は実際の電力設備の画像における実験や、画像判別以外の物体検出やセグメンテーションを行う他のAIモデルにも広く適用出来るよう、手法の改良や、更に効果的な画像を生成するための手法開発を予定している。こうした活動を通じ、社内のAI適用する業務を増やしていき、社内のDX化に貢献していきたい。

5 参考文献

- Zhang, L., Rao, A., & Agrawal, M. (2023). Adding conditional control to text-to-image diffusion models. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 3836-3847).
- Kaggle. Tomato leaf disease detection, <https://www.kaggle.com/datasets/kaustubhb999/tomatoleaf>, 2025-01-07 アクセス。
- Tan, M., & Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In International conference on machine learning (pp. 10096-10106). PMLR.
- Stability-AI. Stable Diffusion Public Release, <https://stability.ai/news/stable-diffusion-public-release>, 2025-01-07 アクセス。