

AI 技術を活用した検査工程の省力化・効率化（第8報）

—欠陥画像分類におけるデータ拡張の効果—
渡辺博己*、生駒晃大*、松原早苗*

A study on artificial intelligence for labor savings or efficiency improvements of inspection process (VIII)

- Effects of data augmentation in defect image classification using deep learning -

WATANABE Hiroki*, IKOMA Akihiro* and MATSUBARA Sanae*

深層学習を利用した画像分類では、大量の学習用データを必要とするが、データ収集が困難な場合がある。本研究では、画像データに幾何変形を施す手法と、敵対的生成ネットワークを用いて新たな画像データを生成する手法により、少量の学習用データを拡張する方法を検討した。実験では、欠陥画像データセットを用いてデータ拡張を施した欠陥画像分類モデルを生成し、欠陥種別等に対する分類性能を検証した。その結果、データ拡張手法を用いて学習した欠陥画像分類モデルの有効性が確認できた。

1. はじめに

近年、深層学習により生成されたモデルの推論精度が高いことから、深層学習モデルが社会生活の様々な場面で使用されるようになりつつある。深層学習モデルが高い推論精度を持つ理由の一つとして、推論する上で非常に重要な特徴量を抽出する仕組みをモデルが獲得していることが挙げられる。しかし、このようなモデルを生成するためには、大量のデータが必要であることが知られており、データ収集が推論精度の高い深層学習モデル生成の課題となっている。

この課題を解決する手段として、前年度報告⁴⁾した転移学習等による手法以外に、元の学習データに幾何変形を施してデータ量を増加させるデータ拡張と呼ばれる手法が有効であると言われている。また、最近では、敵対的生成ネットワーク⁵⁾ (Generative Adversarial Network、以下、GAN) を利用するデータ拡張手法も研究されており、少量のデータでも深層学習モデルの汎化性能を向上させる取り組みが数多く行われている。

本研究では、15種の欠陥画像と4種の正常画像から構成される欠陥画像データセット（以下、データセット。図1）を用いて生成する欠陥画像分類モデル（以下、分類モデル）について、幾何変形、及びGANによるデータ拡張手法を検討した。実験では、幾何変形、GAN、及び幾何変形とGANの組み合わせによりデータ拡張を施した分類モデルを生成し、画像分類性能を検証した。本稿では、これらの内容について報告する。

2. データ拡張

2.1 幾何変形によるデータ拡張

データ量を増加させるデータ拡張手法の一つに、元画像に幾何変形を施す方法がある。幾何変形には、回転、平行移動、拡大・縮小、反転、剪断等の変形方法があるが、多くの場合、使用する画像に合わせて、適切に変形

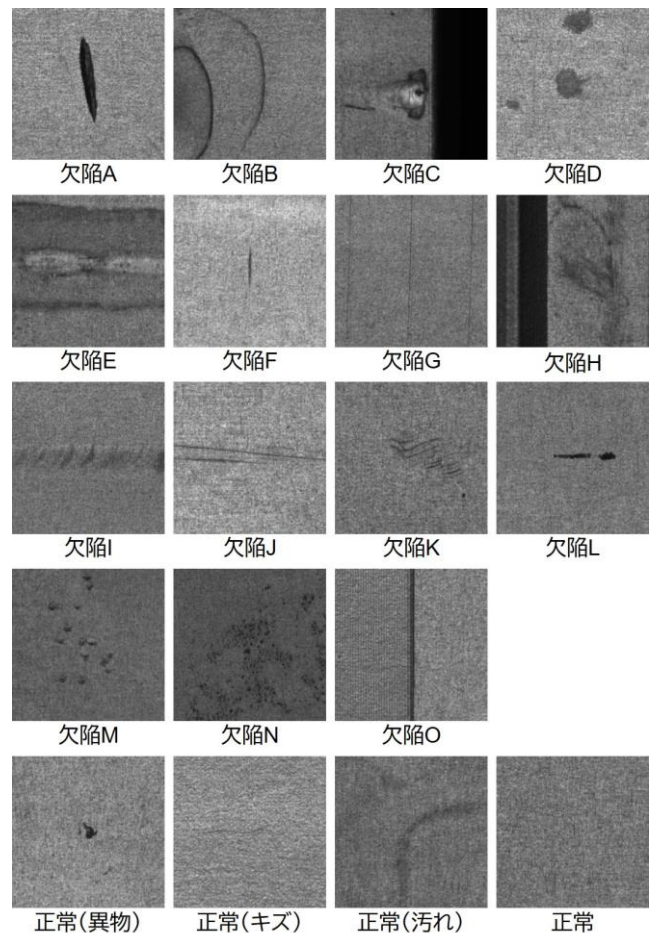


図1 欠陥画像データセットの画像例

方法を選定する必要がある。

本研究で取り扱うデータセットは、工業部品の製造過程で付加された傷等を撮影した画像から構成されており、傷の方向や大きさ・深さ等は、欠陥の要素として重要な情報を持つと考えられる。そのため、回転、拡大・縮小等の変形による情報損失は回避すべきである。また、画像は、4つの照明条件下で撮影されており、左方向や右方向から点灯した画像が含まれているため、傷等

* 情報技術部

の陰影に関する情報損失も回避すべきである。そのため、本研究で施した幾何変形は、平行移動のみとした。

2.2 GANによるデータ拡張

GANは、画像を生成するGeneratorと、入力画像か生成画像かを識別するDiscriminatorから構成される。

Discriminatorは、正しい入力画像を識別できるよう学習し、Generatorは、Discriminatorに正しい画像であるかのように見せかける画像を生成できるよう学習する。データ拡張にGANを用いる場合は、GANで機械的に画像を生成し、データ量を増加させる。

GANには、用途により様々な種類があるが、前々年年度の報告³⁾では、DCGANを用いて欠陥画像を生成した。しかし、DCGANでは、ランダムなノイズデータから目的とする画像を生成することから、十分な量の学習データが必要であることが分かった。

そこで、本研究では、Image-to-Image Translationを目的としたGANであるDCLGAN²⁾を利用して、正常画像から欠陥画像、欠陥画像から正常画像に変換するモデル(以下、変換モデル)を学習することにより、目的とする画像を生成することにした。DCLGANは、ドメイン(データの値の範囲や制約)間で双方向の変換を行う2つのGeneratorと、各ドメインについて正しい入力画像か否かを判定する2つのDiscriminatorから構成され、対照学習にDual Learningを組み合わせたGANである。なお、対照学習は、ラベル付けを行うことなくデータ同士を比較する仕組みを用いて、膨大なデータでもそのまま学習できる自己教師あり学習の一つである。また、Dual Learningは、あるデータを異なるドメインに変換し、変換したデータから元のデータを復元できるかに基づいて、双方向の変換を学習する手法である。

3. DCLGANによる画像生成実験

本研究で生成する分類モデルは、マルチタスク学習に対応したモデル⁵⁾(以下、従来モデル)であるため、全画像データについて、各タスクに応じた正解ラベルが必要である。タスクには、主タスクとして、19クラスの欠陥種別に分類するタスクA以外に、補タスクとして、欠陥種別を正常、欠陥の2クラスに大別するタスクB、部品部位である胴周部と端部・テーパ部との2クラスに分類するタスクC、4クラスの照明条件に分類するタスクDがある。そのため、DCLGANで変換モデルを学習する際には、タスクA、C、Dの正解ラベルの取得が容易になるよう、欠陥種別、部位別、照明条件別に学習データを分割し、個々に変換モデルを学習した。

学習に使用したデータ数を表1に示す。なお、正常(汚れ)クラスについては十分なデータ数があるため、変換モデルの学習の対象外とした。各変換モデルの学習にあたっては、入力データとして正常クラスと正常以外のクラスの画像を欠陥種別、部位別、及び照明条件別に与えた。また、各変換モデルによる画像生成時には、正

表1 DCLGANによる変換モデルの学習データ数

項目	胴周部				端部・テーパ部				計
	照明1	照明2	照明3	照明4	照明1	照明2	照明3	照明4	
欠陥A	299	52	236	257	112	11	67	45	1,079
欠陥B	199	44	172	194	75	27	67	48	826
欠陥C	0	0	0	0	348	118	228	190	884
欠陥D	244	120	194	203	55	23	28	35	902
欠陥E	247	124	177	156	68	21	57	50	900
欠陥F	289	0	256	250	58	0	45	45	943
欠陥G	64	0	225	209	13	0	49	44	604
欠陥H	0	0	0	0	493	0	232	178	903
欠陥I	339	14	163	98	143	30	69	53	909
欠陥J	427	200	128	32	137	62	54	22	1,062
欠陥K	218	135	23	30	115	64	12	14	611
欠陥L	472	369	22	19	28	13	0	0	923
欠陥M	306	20	216	189	31	9	19	19	809
欠陥N	394	25	210	180	86	12	19	39	965
欠陥O	81	0	42	18	27	0	238	194	600
正常(異物)	500	0	300	300	325	0	300	300	2,025
正常(キズ)	500	70	300	300	500	13	300	261	2,244
正常 学習用	500	65	300	300	500	10	300	300	2,275
正常 生成用	8,132	65	1,325	1,707	790	10	304	620	12,953

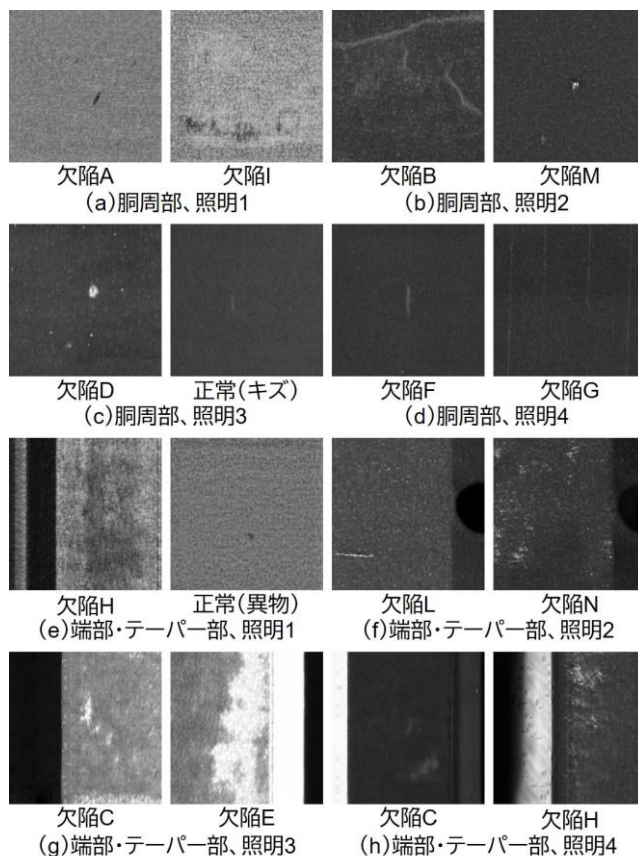


図2 DCLGANによる画像生成例

常クラスの画像については、同クラスの生成用データを、正常以外のクラスの画像については、各クラスの学習用データを入力し、正常以外のクラスの画像と正常クラスの画像にそれぞれ変換した。なお、正常クラスのデータ数についても十分であるため、本研究では、変換後の正常クラスの画像は利用しないことにした。

生成した画像例を図2に示す。生成画像を目視で確認したところ、欠陥Oクラスについては、欠陥の特徴表現が不十分であったが、欠陥O以外のクラスについては、全ての部位別、照明条件別変換モデルによる画像生

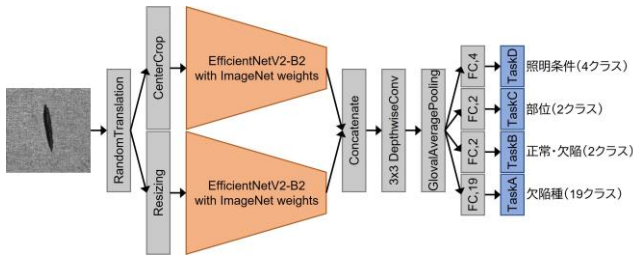


図3 拡張モデル A の構成

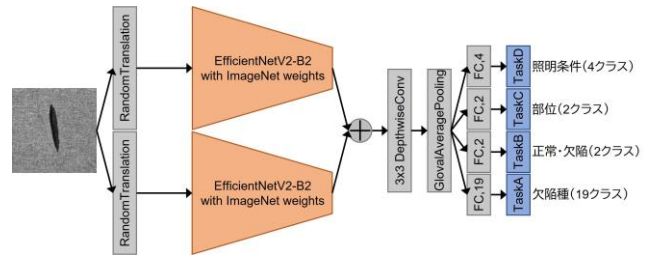


図4 拡張モデル B の構成

表2 従来モデルと拡張モデル A の平均正解率

	学習データ数	検証データ数	従来モデル	拡張モデルA
欠陥A	979	100	0.8740	0.8790
欠陥B	726	100	0.9460	0.9700
欠陥C	784	100	0.9890	0.9930
欠陥D	802	100	0.9870	0.9810
欠陥E	800	100	0.9680	0.9630
欠陥F	843	100	0.9420	0.9450
欠陥G	504	100	0.9920	0.9920
欠陥H	803	100	0.9970	0.9990
欠陥I	809	100	0.9720	0.9720
欠陥J	962	100	0.8750	0.8950
欠陥K	511	100	0.8330	0.8990
欠陥L	823	100	0.9770	0.9800
欠陥M	709	100	0.8650	0.8590
欠陥N	865	100	0.9670	0.9610
欠陥O	500	100	1.0000	1.0000
正常(異物)	5,688	100	0.8530	0.8590
正常(キズ)	6,451	100	0.7940	0.8210
正常(汚れ)	11,836	100	0.9800	0.9810
正常	12,853	100	0.9790	0.9910
欠陥種別計	48,248	1,900	0.9363	0.9442
欠陥(15種)	11,420	1,500	0.9767	0.9755
正常(4種)	36,828	400	0.9788	0.9840
正常・欠陥別計	48,248	1,900	0.9777	0.9797

表3 拡張モデル B の平均正解率

	拡張モデルBx0	拡張モデルBx1	拡張モデルBx2
欠陥A	0.8800	0.8830	0.8910
欠陥B	0.9460	0.9630	0.9740
欠陥C	0.9950	0.9990	1.0000
欠陥D	0.9700	0.9810	0.9830
欠陥E	0.9740	0.9790	0.9780
欠陥F	0.9490	0.9380	0.9380
欠陥G	0.9870	0.9910	0.9890
欠陥H	0.9950	0.9990	0.9990
欠陥I	0.9660	0.9740	0.9780
欠陥J	0.9030	0.9040	0.9210
欠陥K	0.8760	0.9000	0.8970
欠陥L	0.9770	0.9760	0.9810
欠陥M	0.8820	0.9020	0.8770
欠陥N	0.9590	0.9600	0.9590
欠陥O	1.0000	1.0000	1.0000
正常(異物)	0.8500	0.8470	0.8690
正常(キズ)	0.8090	0.8270	0.8390
正常(汚れ)	0.9800	0.9870	0.9850
正常	0.9710	0.9790	0.9830
欠陥種別計	0.9405	0.9468	0.9495
欠陥(15種)	0.9773	0.9805	0.9782
正常(4種)	0.9823	0.9850	0.9858
正常・欠陥別計	0.9798	0.9827	0.9820

成は困難であったものの、複数のモデルで欠陥の特徴表現が十分な画像が生成された。なお、画像生成時には、特徴表現のバラ付きを考慮し、各変換モデルについて、異なるエポックで生成された2つのモデルを使用した。

4. 分類性能の検証実験

4.1 幾何変形による分類モデル

幾何変形によるデータ拡張を施した分類モデルを検証するために、従来モデルに RandomTranslation 層を追加した分類モデル（以下、拡張モデル A。図3）を生成する実験を10回試行した。また、分類性能の検証にあたっては、生成した各拡張モデルについて、検証データに対する欠陥種別、正常・欠陥別の正解率を求めた。表2に、従来モデルと拡張モデル A の平均正解率を示す。

また、拡張モデル A では、 256×256 (px) の入力データの解像度を Resizing 層、または CenterCrop 層で 160×160 (px) に縮小してから EfficientNet により特徴量を抽出するが、入力データの解像度のまま特徴量を抽出するモデル（以下、拡張モデル B。図4）についても、モデル生成実験を10回試行した。なお、拡張モデル B については、RandomTranslation 層が両方の EfficientNet の前にある構成を x2、片方の場合を x1、どちらにもない場合を x0 としてモデルを生成し、検証データに対する欠陥種別、正常・欠陥別の正解率を求めた。表3に、拡張モデル B の平均正解率を示す。

表4 DCLGAN でデータを拡張したデータセットにより学習した従来モデルの平均正解率

	データセットα		データセットβ	
	学習データ数	従来モデル	学習データ数	従来モデル
欠陥A	5,979	0.8670	10,979	0.8700
欠陥B	5,726	0.9350	6,920	0.9140
欠陥C	3,288	0.9890	3,288	0.9890
欠陥D	5,802	0.9790	10,802	0.9670
欠陥E	5,800	0.9640	10,800	0.9640
欠陥F	5,843	0.9290	10,843	0.9090
欠陥G	5,504	0.9850	10,504	0.9880
欠陥H	3,137	0.9950	3,137	0.9990
欠陥I	5,809	0.9610	10,809	0.9630
欠陥J	5,962	0.8660	10,962	0.8710
欠陥K	5,511	0.8450	10,511	0.8650
欠陥L	5,823	0.9740	10,823	0.9730
欠陥M	5,709	0.8720	6,903	0.8530
欠陥N	5,865	0.9460	10,865	0.9330
欠陥O	500	1.0000	500	1.0000
正常(異物)	10,688	0.8370	15,688	0.8470
正常(キズ)	11,451	0.7800	16,451	0.7910
正常(汚れ)	11,836	0.9740	11,836	0.9760
正常	12,853	0.9830	12,853	0.9820
欠陥種別計	123,086	0.9306	185,474	0.9292
欠陥(15種)	76,258	0.9720	128,646	0.9671
正常(4種)	46,828	0.9783	56,828	0.9798
正常・欠陥別計	123,086	0.9751	185,474	0.9734

4.2 DCLGAN による分類モデル

DCLGAN によるデータ拡張については、変換モデルにより生成した欠陥種別の画像をデータセットに追加し、従来モデルを使用して分類モデルを生成する実験を10回試行した。表4に、追加後の学習データ数と検証データに対する従来モデルの平均正解率を示す。なお、実験では、学習データ数の異なる2つのデータセット

表5 DCLGAN でデータを拡張したデータセットにより学習した拡張モデルの平均正解率

	データセット α				データセット β			
	拡張モデルA	拡張モデルBx0	拡張モデルBx1	拡張モデルBx2	拡張モデルA	拡張モデルBx0	拡張モデルBx1	拡張モデルBx2
欠陥A	0.8880	0.8470	0.8810	0.8870	0.8810	0.8380	0.8820	0.8780
欠陥B	0.9670	0.9550	0.9580	0.9630	0.9610	0.9470	0.9590	0.9650
欠陥C	0.9910	0.9970	0.9950	0.9940	0.9950	0.9980	0.9920	0.9960
欠陥D	0.9750	0.9730	0.9780	0.9760	0.9770	0.9670	0.9750	0.9730
欠陥E	0.9740	0.9720	0.9760	0.9690	0.9710	0.9730	0.9740	0.9780
欠陥F	0.9380	0.9300	0.9420	0.9420	0.9380	0.9350	0.9520	0.9490
欠陥G	0.9880	0.9850	0.9880	0.9880	0.9910	0.9840	0.9850	0.9870
欠陥H	0.9990	0.9980	0.9960	0.9940	1.0000	0.9970	0.9990	0.9980
欠陥I	0.9650	0.9690	0.9660	0.9690	0.9690	0.9740	0.9690	0.9700
欠陥J	0.8940	0.8920	0.9160	0.9290	0.8900	0.8990	0.9010	0.9180
欠陥K	0.8890	0.8860	0.8970	0.8890	0.8810	0.8770	0.9130	0.8890
欠陥L	0.9780	0.9780	0.9780	0.9770	0.9750	0.9760	0.9770	0.9760
欠陥M	0.8590	0.8570	0.8790	0.8680	0.8410	0.8440	0.8700	0.8780
欠陥N	0.9710	0.9530	0.9620	0.9550	0.9630	0.9520	0.9560	0.9580
欠陥O	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
正常(異物)	0.8550	0.8350	0.8700	0.8660	0.8650	0.8530	0.8540	0.8560
正常(キズ)	0.8220	0.7750	0.8160	0.8410	0.8250	0.7910	0.8180	0.8220
正常(汚れ)	0.9860	0.9710	0.9820	0.9870	0.9890	0.9710	0.9830	0.9890
正常	0.9850	0.9750	0.9780	0.9790	0.9830	0.9720	0.9720	0.9800
欠陥種別計	0.9434	0.9341	0.9452	0.9459	0.9418	0.9341	0.9437	0.9453
欠陥(15種)	0.9749	0.9739	0.9764	0.9743	0.9726	0.9721	0.9767	0.9764
正常(4種)	0.9853	0.9815	0.9865	0.9875	0.9870	0.9868	0.9873	0.9835
正常・欠陥別計	0.9801	0.9777	0.9815	0.9809	0.9798	0.9794	0.9820	0.9800

α 、 β を構成して従来モデルを学習し、4.1と同様に正解率を求めた。

4.3 幾何変形とDCLGANによる分類モデル

幾何変形とDCLGANの組み合わせによるデータ拡張を施した分類モデルを生成するために、4.2のデータセット α 、 β を4.1の拡張モデルA、Bで学習する実験を行った。実験では、各拡張モデルの学習を10回試行し、4.1と同様に正解率を求めた。表5に、各データセットにおける各拡張モデルの平均正解率を示す。

4.4 考察

表2より、拡張モデルAの欠陥種別計の平均正解率が、従来モデルより0.79%向上した。また、拡張モデルBについては、表3の欠陥種別計の平均正解率をx0とx1、x1とx2で比較すると、順に0.63%、0.27%向上し、RandomTranslation層が増えるほど平均正解率が高くなるという結果が得られた。これらのことから、幾何変形によるデータ拡張には、欠陥画像に対する分類性能を向上させる効果があることが確認できた。

また、表4のデータセット α 、 β を用いて学習した従来モデルについては、表2の従来モデルの欠陥種別計の平均正解率を下回った。この要因として考えられるのは、DCLGANによる変換画像の表現性である。変換画像の欠陥の数、位置、大きさ等については変化があり、複製された画像は見受けられないため、GANの学習時に発生するモード崩壊は起きていないと考えられるが、同一クラス内の欠陥の特徴表現の乏しさが原因であると考えられる。そのため、DCLGANによる変換画像の表現性の拡大について検討する必要がある。

さらに、表5についても表4の場合と同様に、表2、3で対応する各モデルと比較して、欠陥種別計の平均正解率が下回る結果となった。しかし、表2、3の結果と同様に、欠陥種別計の平均正解率は、従来モデルより拡張モデルAが高くなり、拡張モデルBではx0、x1、x2

の順に高くなった。そのため、GANによりデータ拡張を施した場合でも、幾何変形によるデータ拡張は、欠陥画像の分類性能を向上させるのに有効であることが確認できた。

5. まとめ

本研究では、データ拡張手法を用いた場合の画像分類性能について検討した。データ拡張手法としては、幾何変形として画像を平行移動する方法と、DCLGANを用いて正常画像を正常以外の画像に変換し、データ量を増加させる方法について検証した。その結果、DCLGANを用いたデータ拡張については、有効性を確認できなかったが、平行移動によるデータ拡張については、分類性能を向上させる効果があることが確認できた。

今後は、GANによる生成画像の表現性について検討し、データ拡張に関する欠陥画像分類性能の検証を継続する予定である。

【謝辞】

本研究を遂行するにあたり、欠陥画像データセットをご提供いただきました株式会社前田精工の皆様には深く感謝の意を表します。

【参考文献】

- 1) I. Goodfellow, et al., Proc. NIPS 2014, pp.2672-2680, 2014
- 2) J. Han, et al., arXiv: 2104.07689, 2021
- 3) 渡辺ら, 岐阜県産業技術総合センター研究報告 No.2, pp.93-94, 2021
- 4) 渡辺ら, 岐阜県産業技術総合センター研究報告 No.3, pp.99-100, 2022
- 5) 渡辺ら, 岐阜県産業技術総合センター研究報告 No.3, pp.101-102, 2022