

# 外観検査 AI モデルのための合成データ生成技術

## Synthetic Data Generation Techniques for Visual Inspection AI Models

大 島 誉 寿 技術開発本部技術基盤センター先進生産プロセスグループ 主査  
水 山 佳 乃 技術開発本部技術基盤センター先進生産プロセスグループ  
北 村 俊 也 技術開発本部技術基盤センター先進生産プロセスグループ

製造業の外観検査自動化において欠陥パターンを網羅した多様な不良品データの収集が困難であったが、近年は人工的に不良品画像を生成するという合成データ (Synthetic data) 技術によってデータ数を確保する方法が注目されている。そこで、金属部品を対象に画像加工や光学シミュレーションなど複数の手法によって不良品を模擬した合成データを生成した。また、合成データのみを訓練に使用した検査 AI モデルを構築することで、学習用データとしての有用性を確認した。

In the automation of visual inspection in the manufacturing industry, the difficulty of collecting defective product data has been a significant challenge. Recently, the use of synthetic data, which involves the artificial generation of images of defective products, has gained attention as a method to ensure a sufficient volume of data. In this study, synthetic data simulating defective metal components was generated using multiple approaches, including image processing and optical simulation. Furthermore, the usefulness of synthetic data as training data was verified by constructing an AI inspection model trained exclusively on synthetic data.

## 1. 緒 言

工業製品の外観検査の自動化は製造業共通の課題として実現が期待されている。解決手段としてカメラと画像認識 AI モデルを用いた開発が一般的なアプローチとなりつつあるが、AI モデルの学習用・検証用データとして用いる不良品画像の収集には、不良品発生率の低さや、発生パターンを網羅しきれないという困難がある。良品画像のみで検査モデルを構築する方法はあるものの、過検出率が高い傾向にあるため、運用時に検査員によるダブルチェックが頻発し、AI 導入によるコストメリットに見合わない場合も多い。

解決策の一つとして、実際の不良品撮影画像ではなく、コンピュータ上の計算によって生成された人工の画像データを AI の学習・検証に使用する方法が近年注目されている。人工的に生成された画像データは合成データ (Synthetic data) と呼称され、自動車業界では自動運転アルゴリズムの学習・検証用データとして道路シーンを模擬した合成データがすでに活用されている。製造業でも不良品画像を模擬した合成データの活用による外観検査自動化の実現が期待されており、生成アルゴリズムを提供するサービスも幾つか登場してきている。

国際規格においても合成データについての言及がある。例えば非破壊検査の D-RT (Digital-Radiographic Testing) 画像を対象とした半自動欠陥検出モデル構築のガイドラインである ASTM E3327<sup>(1)</sup> の文中では、実際の製造工程であらゆる欠陥パターンを網羅するデータ取得が現実的には困難なことから、実際の取得データに加え、合成データの使用が推奨されている。

不良品画像の合成データは適用する検査対象および検査方法 (撮影方法) によって作成コストが大きく左右される。大別すると現状三つの生成手法が考えられる。すなわち「画像加工」、「光学シミュレーション (3D-CG)」、「生成 AI」を使用した手法である。本稿では IHI で検討してきた具体的な実施例を交えながらおのおのの生成手法を紹介し、学習用データとしての合成データの評価結果を示す。

## 2. 画像加工による生成手法

最も低コストで不良品の合成データを生成する方法は、良品画像の一部箇所を疑似不良箇所として編集・加工することである。疑似不良箇所は実際の不良品画像の一部を切り貼りするか、画像編集ソフトウェアやプログラムを使用して生成することが多い。ただしこの生成手法は撮影対象

や条件が単純なケースに適用に限られる。具体的な例として壁面や配管内をカメラが走査しながら撮影した検査画像や、超音波、赤外線、X線といった非破壊検査手法で得た検査画像が挙げられる。これらは二次元平面的で奥行きのない画像情報を扱うため、画像加工手法のみで低コストながら十分な品質の不良品合成データを得ることができる。

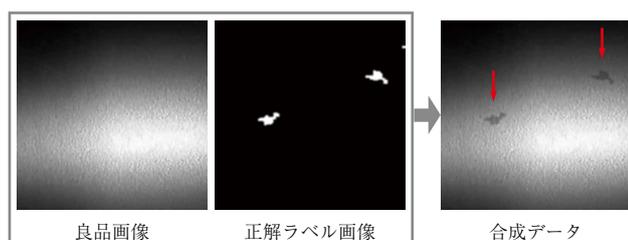
配管を模擬した円筒金属内面の良品部分を撮影した画像に対し、画像加工により腐食跡の局所減肉きずを付与した合成データを生成した例を第1図に示す。この例では腐食跡の減肉形状がセルオートマトンで近似できることを利用し、照明がつくる減肉のきずの影をセルオートマトンのプログラムによる画素値の編集で模擬した。

### 3. 光学シミュレーションによる生成手法

現実的には画像加工のみでもっともらしい不良品の合成データを生成することが困難なケースは多く存在する。具体的には、検査対象が複雑な三次元形状を有する、照明配置が複雑な撮影条件である、対象表面の光学特性に異方性がある、といった場合などが該当する。カメラへ入射する光線の物理的なシミュレーションを要するため、一般的に3D-CG (Computer Graphics) または物理ベースレンダリング<sup>(2)</sup> と呼ばれる手法で合成データを生成する。例えば、VQA (Visual Question Answering) と呼ばれる入力画像の説明文を出力するタスクを研究する分野ではCLEVR<sup>(3)</sup> というプロジェクトがデータセットを公開しているが、これは3D-CGによる合成データのみで構成されている。

#### 3.1 模擬翼へのきず導入の例

模擬翼を検査対象と想定し、3D-CADモデルを用いてその表面に局所的な凹みをきずとして導入し3D-CGによる撮影画像を合成データとして生成する手法と出力画像例を紹介する。想定する欠陥は50 $\mu$ m程度の深さを持ち、凹み形状の欠陥であるデント、細長いひっかき跡のような



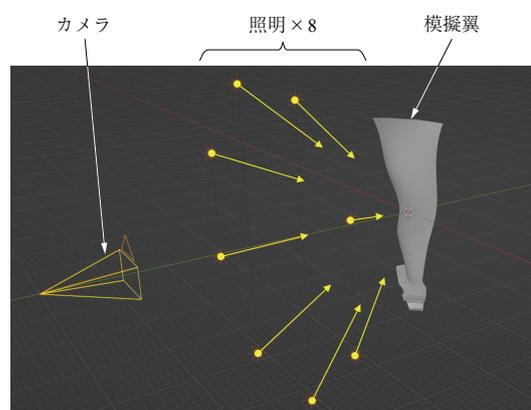
第1図 良品画像の加工によるきず指示付与の例  
Fig. 1 Example of adding defect annotations to images of non-defective products

欠陥であるスクラッチ、小さな欠けや切り欠き形状の欠陥であるニックの3種類である。Blender Ver.4.0を使用して、模擬翼3Dモデルを仮想空間上に作成し、カメラや照明を配置した撮影シーンを再現した様子を第2図に示す。今回の検出対象は凹みきずであるため、検査対象モデルのメッシュの各頂点をもとの位置から変位させることできずを疑似的に再現した。

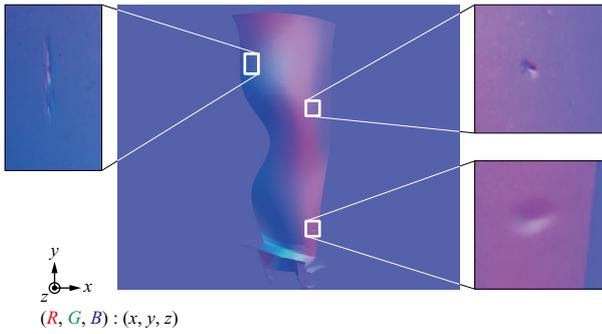
また撮影方法としてフォトメトリックステレオ法(照度差ステレオ法)<sup>(4)</sup>を採用したため照射方向の異なる複数の照明を配置した。フォトメトリックステレオ法は複数の照明の発光パターンごとに撮影した画像群の解析から撮影対象表面の法線を推定し法線マップとして出力することができる。法線マップから高さマップや曲率マップを計算することもできるため、検査対象の局所的な凹凸きずを強調する画像化手法としてよく用いられる。

3D-CGで実際の撮影画像に近い結果を得ようとした場合、重要となるのは現実の撮影対象表面の光学特性(反射、散乱、透過、吸収などの物理特性)の再現度の高さである。3.2節にて紹介するが、一般的にこの表面光学特性を測定するには大きなコストを要する。しかしながらフォトメトリックステレオ法における法線マップは対象の形状に注目した画像化方法のため、表面光学特性を忠実に再現できなかったとしても実際の撮影画像に近い結果を得られるメリットがある。

3D-CGで生成した合成データ(法線マップ)と凹みきず箇所を第3図に示す。法線マップのため画像のRGB値が被写体表面法線のxyz成分に対応する。外観検査における凹凸きずは表面的な広がり程度の面積だけでなく、表面からの最大深さ・高さで合否判定するケースが多い。3D-CGではこれらの数値をハイパーパラメータ



第2図 仮想空間に配置した撮影シーンの構成  
Fig. 2 Configuration of the shooting scene arranged in a virtual space



第3図 3D-CG で生成した合成データ（法線マップ）  
 Fig. 3 Synthetic data (normal map) generated by 3D-CG

として扱うことで、きず特徴を制御した生成が可能である。きずの形状、面積、最大深さをランダムに変化させることでさまざまなきずを撮影した合成データセットを構築した。生成した合成データを  $200 \times 200$  pixel に切り出したパッチ画像の例を第4図に示す。参考として、工程で実際に発生する欠陥形状を模擬して、検査対象良品に模擬きずを付与した供試体の撮影画像も第4図右側に示している。

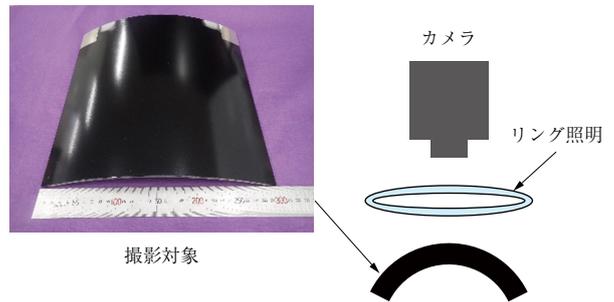
### 3.2 表面光学特性の取得・設定方法

3.1 節にてフォトメトリック撮影手法を採用することにより、表面光学特性の忠実な再現を省略しつつ実画像と近い結果が得られた例を示した。しかし、実際の撮影手法にはパターン投影法やマルチスペクトルカメラを使用した撮影手法などさまざまであり、検査対象表面の光学特性を無視できる手法は少ない。フォトメトリックステレオ法以外の撮影手法で得られる画像に対しても忠実に再現し、合成データを活用するためには、検査対象の表面光学特性の設定が重要である。これからその取得・設定方法について幾つかの例を紹介する。

#### 3.2.1 シェーダーのパラメータ調節による設定

3D-CG ソフトウェアにはオブジェクト表面の光のふるまいに関する数学的な手続きがまとめられたシェーダーが幾つか搭載されており、反射率や粗さ具合といったパラメータを調節することで光学特性を設定可能である。本稿で使用した Blender の場合は Principled BSDF<sup>(5)</sup> が該当する。一般的には設計者がソフトウェア上で現実を再現するように手でパラメータを調整するが経験と時間を要する。そこで、複数の撮影条件で実際に撮影した画像を目標に、3D-CG にて同じ撮影条件で生成（レンダリング）した画像を任意の評価値を用いて誤差を少なくするようにシェーダーのパラメータを最適化することも考えられる。

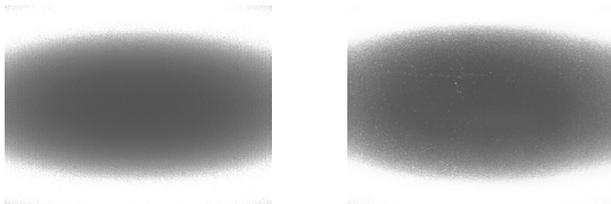
黒色塗装が施された曲げ金属板を対象とし、本設定手法を用いた例を紹介する。撮影状況の模式図を第5図に示す。同様の撮像シーンを 3D-CG 上に再現し、実際の撮影画像と 3D-CG のレンダリング画像の類似度を最大化するように、シェーダーのパラメータを最適化計算により設定した。最適化後のレンダリング画像と実際の撮影画像の比較例を第6図に示す。検査対象の表面性状や形状による制



第5図 黒色塗装金属板の撮影状況の模式図  
 Fig. 5 Schematic representation of the photography scene setup for a black-painted metal plate

きず形状	合成データ	実際の模擬きず画像
デント状		
スクラッチ状		
ニック状		

第4図 合成データと実際のきず画像の比較例  
 Fig. 4 Comparative example of synthetic data and real defect images



(a) 3D-CG作成の合成データ (b) 実際の撮影画像

第6図 黒色塗装金属板の撮影画像

Fig. 6 Captured images of a black-painted metal plate

約で光学特性の取得が困難な場合には、このような簡易なパラメータ調節方法で模擬することが可能である。

### 3.2.2 テクスチャ画像による設定

3.2.1項のシェーダーのパラメータに表面光学特性を反映したPBR (Physical Based Rendering) テクスチャと呼ばれる素材画像を指定することが可能である。テクスチャ画像を設定することで表面の模様や粗さレベルの微小な表面凹凸も再現できるため、より現実に忠実なレンダリングが可能となる。代表的なものとして色味を表すアルベドマップ、微小凹凸を表すノーマルマップやバンプマップ、粗さを表すラフネスマップなどがある。

PBR テクスチャを得るには、実際の検査対象の一部平面の素材に複数の照明方向から対象を撮影するため撮影スタジオを要するものの、世の中にはこれらの撮影を請け負うサービスも多く存在する。金属製粗さ試験片 ( $Ra = 1.6 \mu\text{m}$ ) を対象として生成したテクスチャ画像を第7図に示す。また、それらを用いて直方体表面性状を適用した場合の3D-CG上の外観を第8図に示す。表面上に線状の粗さが付与されていることが見て取れる。

### 3.2.3 BRDF 測定結果による設定

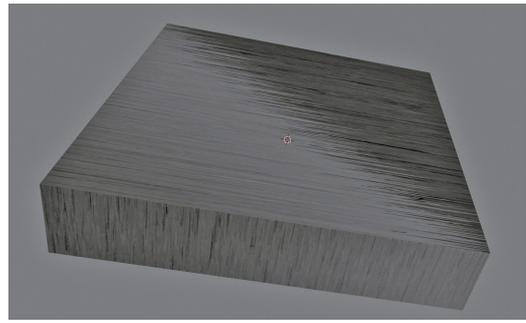
3D-CG上で光の反射特性をより忠実に再現する方法として、物体表面の入射光の方向と観測光の方向に対する表面の反射率を定量的に測定するBRDF (Bidirectional Reflectance Distribution Function, 二方向反射分布関数)<sup>(6)</sup>を利用することが考えられる。BRDF測定の模式図と入射角 $5^\circ$ 、 $25^\circ$ 、 $60^\circ$ 、 $70^\circ$ で固定したときの測定結



(a) アルベドマップ (b) ノーマルマップ (c) ラフネスマップ

第7図 作成したPBRテクスチャ画像の例

Fig. 7 Example of created PBR texture images



第8図 粗さ試験片の表面性状を直方体へ適用した例

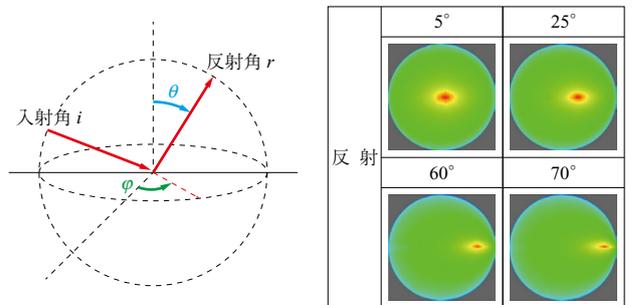
Fig. 8 Example of applying the surface texture of roughness test pieces to a cube

果例を第9図に示す。離散的なBRDF測定結果を連続関数化したモデルをシミュレーション上の検査対象表面へ適用する。実検査対象の光学特性の測定値を基に設定するため厳密性はあるが、専用の測定装置や光学解析シミュレーションソフトを要するためコストが大きい欠点がある。

### 3.2.4 表面光学特性の設定手法の比較評価

BRDF測定結果を適用可能な光学解析シミュレーションソフトであるSpeos (ANSYS社, 米国)を用いて、3.2.2項で紹介したテクスチャ画像による手法と3.2.3項で紹介したBRDF測定による手法の比較評価を実施した。Speos上での構成模式図を第10図に示し、平板試験体を対象として表面光学特性を各手法で設定したときのレンダリング画像を第11図に示す。第10図のようにシミュレーション上で、レンズ設計データやIES (Illuminating Engineering Society) データを用いて、実際の撮影環境を再現することも可能である。テクスチャ画像とBRDF測定結果を同時に設定することも可能であり、両者を設定することで、実物の同様の反射特性を保持する表面上に発生する微小な凹凸形状を再現可能であることを期待した。

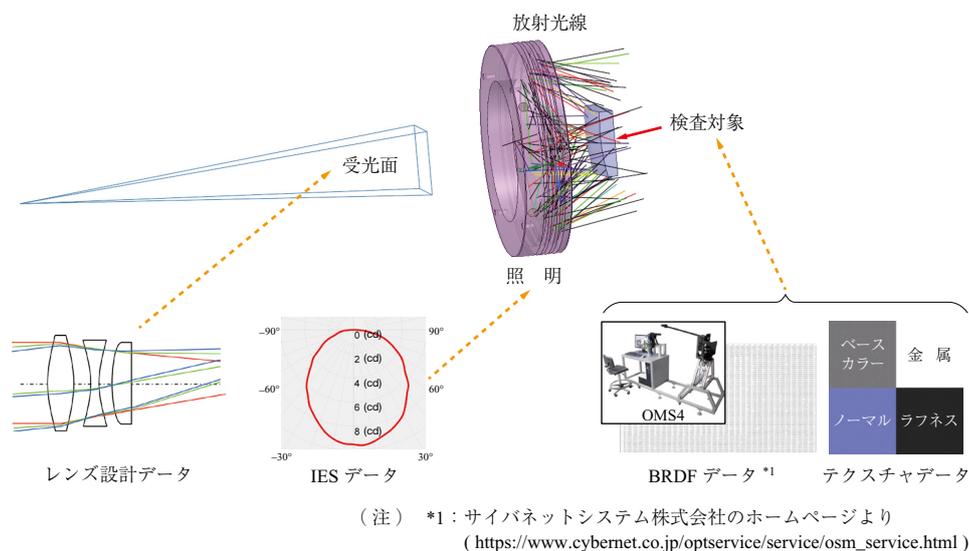
画像類似度を示す一般的な指標であるSSIM (Structural



(a) BRDF測定の模式図 (b) 測定結果 (conoscopic map) の例

第9図 BRDF計測の模式図と計測結果の例

Fig. 9 Schematic representation of BRDF measurement and example results



第 10 図 Speos 上での構成模式図  
Fig. 10 Schematic representation of configuration in Speos

画像種類	実画像	レンダリング画像		
		BRDF のみ	テクスチャのみ	BRDF + テクスチャ
光学特性の設定データ	-	BRDF のみ	テクスチャのみ	BRDF + テクスチャ
画像				
実画像との SSIM	-	79.61%	71.06%	78.76%

第 11 図 平板試験体による表面光学再現手法の検証結果  
Fig. 11 Verification results of surface optics reproduction method using flat test specimens

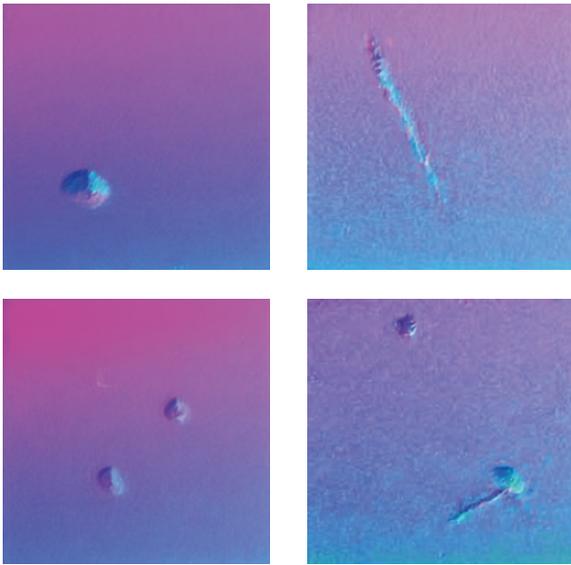
Similarity Index)<sup>(7)</sup> を評価指標として使用し、実画像と生成画像の SSIM を第 11 図の各画像の下に示している。BRDF 測定のみを適用したシミュレーション画像が、テクスチャ画像のみを適用した結果よりも実物との類似度が高いことから、BRDF 測定結果を適用することにより検査対象の表面が持つ反射特性をより正確に再現したことを確認した。BRDF 測定結果のみを適用した画像 (第 11 図の左から 2 列目) の方が BRDF 測定結果とテクスチャ画像を適用した画像 (第 11 図の左から 4 列目) より類似度が僅かに高くなっているが、これはテクスチャ画像をスキャンした表面と実画像の撮像範囲の位置合わせをしていないことが原因の可能性がある。評価指標である SSIM は輝度、コントラスト、構造の比較項からなり、局所領域ごとに評価値を計算し、それらを平均した値であるため、表面上の微細な凹凸の位置ずれにより局所領域ごとの評価値が低くなり、画像全体の評価値が実際の見た目よりも低くなる場合がある。生成した画像を適切に評価可能な評価

指標を確立することが課題であることが分かった。

#### 4. 生成 AI (拡散モデル) による生成手法

3D-CG でのモデリングが困難なほど複雑な不良品合成データを生成したい場合は、一般的に画像生成 AI と呼ばれる拡散モデル<sup>(8)</sup> のファインチューニングにより生成できる可能性がある。

第 12 図にニックやスクラッチといった局所的な凹みきずのフォトメトリックステレオ撮影による法線マップの合成データを拡散モデルにより生成した例を示す。ファインチューニングは Stable Diffusion v1.4<sup>(9)</sup> モデルをベースに 20 枚の実際のきず画像を学習素材として DreamBooth<sup>(10)</sup> と呼ばれる手法を用いて学習した。注意点として本方法では生成結果が必ずしもきずの合成データを示すわけではないため、各生成結果に対してそれがきずであるか判別が可能な専門家による目視の仕分け作業、つまりアノテーションコストが発生する。また生成手法の理論的な背景から、



第 12 図 拡散モデルで生成した合成データ（法線マップ）のきず  
 Fig. 12 Defects in synthetic data (normal map) generated by a diffusion model

確率的な生成となるため、3D-CG のようにきずの性状をコントロールすることはできない。このように使用上の注意点は幾つか存在するものの、技術の発展が著しい分野であるため合成データ生成の手段として検討すべき選択肢である。

### 5. 合成データの有効性評価

3D-CG により作成した合成データの有効性を検証するため、合成データのみで学習させた画像認識 AI（深層学習モデル）、実際のきず画像をテストデータとして推論させた場合の検出率の評価を行った。検証の簡略化のため、深層学習モデルにはシンプルな画像分類モデルである MobileNetV2<sup>(11)</sup> を使用した。第 4 図のように合成データを 200 × 200 pixel に切り出したパッチ画像を学習データとして使用し、パッチ画像を OK（良品）か NG（きずあり）へ分類する 2 クラス分類問題として学習した。

テストデータの推論結果と正解クラスとの対応を整理した表である混同行列形式にしたものを第 1 表に示す。TPR（検出率）は  $TP/(TP + FN)$  で計算されるため 89.6% であり、FPR（過検出率）は  $FP/(FP + TN)$  で計算される

第 1 表 合成データ学習モデルの推論結果の混同行列  
 Table 1 Confusion matrix of synthetic data-based model predictions

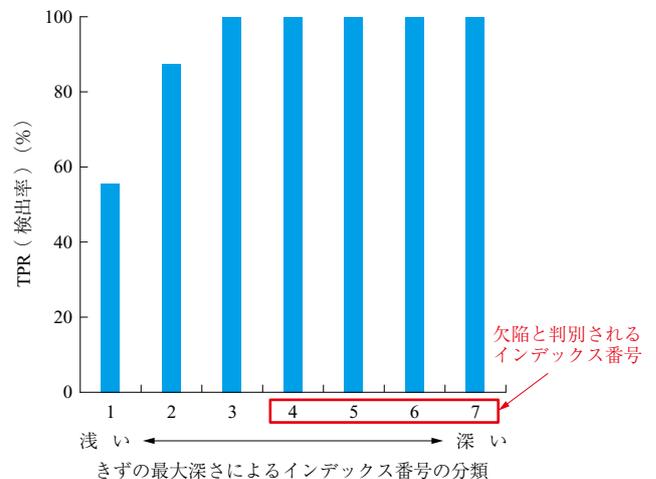
		予測値	
		NG（きずあり）	OK（良品）
正解値	NG（48 枚）	43 枚 (TP)	5 枚 (FN)
	OK（2,365 枚）	288 枚 (FP)	2,077 枚 (TN)

ため 12.2%であった。  $TP/(TP + FP)$  で計算される Precision（適合率）は 13.0%のため誤検知率は 87.0%と多い。ただし、実際の検査モデルを構築する際にはより複雑なネットワークモデルを採用し、実際の撮影データも組み合わせるためこれらの性能値には改善の余地がある。

また導入した凹みきずの最大深さによってきず程度を分類し、深さ順にインデックス番号を振り分けた。インデックス番号ごとの TPR（検出率）を第 13 図に示す。ここから見逃しは最大深さが浅いきずで発生していることが分かる。本検査対象はインデックス番号 4 以上の最大深さを持つきずを欠陥として判別しなければならない要求を設定されているため、用意したテストデータの欠陥検出率は 100%となり 3D-CG で生成した合成データは検査モデルの学習データとして有効であることが確認できた。

今回は学習用データとしてどれほど有用性があるかを評価したが、合成データの評価指標は研究段階でありさまざまな指標が提案されている。例えば合成データと実画像の統計的類似性を評価する手法として、特徴量分布の違いを評価する FID (Frechet Inception Distance)<sup>(12)</sup> や、合成データの多様性を評価する IS (Inception Score)<sup>(13)</sup>、3.2.4 項で採用した構造的類似度を評価する SSIM などがある。また、データの識別性を評価する TSTR (Train on Synthetic, Test on Real) や TRTS (Train on Real, Test on Synthetic)<sup>(14)</sup> という手法もあり、今回は TSTR を採用した評価に相当する。

合成データの評価指標は確立されておらず、合成データの特徴や評価の目的に応じての使い分けが必要である。



第 13 図 きずの最大深さごとに算出した検出率

Fig. 13 Detection rate calculated for each maximum depth of defects

## 6. 結 言

外観検査自動化に向けた AI 検査モデルに供する合成データについて「画像加工」「光学シミュレーション (3D-CG)」「生成 AI」の三つの生成手法について具体例を紹介した。特に光学シミュレーションについては合成データのみを学習用データとして使用した検査分類モデルを構築し、未知の実画像を推論させることで合成データの評価を実施した。TPR89.6%, FPR12.2%の精度で良品か不良品の分類が可能であり、合成データが学習用データとして有効であることを確認できた。

他にも検証用データとしてや検査モデルの転移学習元データとしての使用用途も考えられる。今後の合成データの活用方法としては工程に導入する外観検査 AI モデルの性能検証や品質保証が考えられ、これに伴い検査画像の評価指標の検討を行う。また外観検査装置を新規導入する際の光学系設計のエンジニアリングツールなどの開発を実施する予定である。

## 参 考 文 献

- (1) ASTM E3327/E3327M-21 : Standard Guide for the Qualification and Control of the Assisted Defect Recognition of Digital Radiographic Test Data, (2021)
- (2) B. Burley : Physically Based Shading at Disney, ACM SIGGRAPH, (2012), pp. 1 - 7
- (3) J. Johnson, B. Hariharan, L. Maaten, F. Li, L. Zitnick and R. Girshick : CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2017)
- (4) 井本治孝, 町田賢一郎, 山浦佑介, 蓼沼 周, 前川 卓 : 反復照度差ステレオ法を用いた微細形状計測装置の開発, 日本機械学会論文集, Vol. 82, No. 835, 2016 年
- (5) B. Burley : Extending the Disney BRDF to a BSDF with Integrated Subsurface Scattering, ACM SIGGRAPH, (2015)
- (6) C. Schlick : An Inexpensive BRDF Model for Physically-based Rendering, Computer Graphics Forum, Vol. 13, No. 3, (1994)
- (7) Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli : Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, Vol. 13, Iss. 4, (2004), pp. 600 - 612
- (8) R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer : High-resolution Image Synthesis with Latent Diffusion Models, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), (2022), pp. 10,684 - 10,695
- (9) Stability AI : <https://huggingface.co/CompVis/stable-diffusion-v1-4>, (参照 2022. 8. 20)
- (10) N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein and K. Aberman : DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), (2023), pp. 22,500 - 22,510
- (11) M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen : Mobilenetv2: Inverted Residuals and Linear Bottlenecks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2018), pp. 4,510 - 4,520
- (12) M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler and S. Hochreiter : GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium, Advances in Neural Information Processing Systems 30 (NIPS 2017), (2017)
- (13) T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford and X. Chen : Improved Techniques for Training GANs, Advances in Neural Information Processing Systems 29 (NIPS 2016), (2016)
- (14) L. Xu, M. Skoularidou, A. Cuesta-Infante and K. Veeramachaneni : Modeling Tabular Data Using Conditional GAN, Advances in Neural Information Processing Systems 32 (NeurIPS 2019), (2019)