

VAE を用いた教師なし深層学習による 変状種別に依存しない河川護岸の異常検知

都築 幸乃^{1*}・吉田 龍人¹・大久保 順一¹・藤井 純一郎¹・山下 隆義²

¹正会員 八千代エンジニアリング株式会社 (〒111-8646 東京都台東区浅草橋 5-20-8CS タワー)

*E-mail: yk-tsuzuki@yachiyo-eng.co.jp (Corresponding Author)

²非会員 中部大学工学部情報工学科 (〒487-8501 愛知県春日井市松本町 1200)

教師あり学習に基づくインフラの変状検出は、教師データ収集の観点からモデルの構築が困難であり、未学習の変状が検出できないといった課題がある。一方で、異常データの収集が困難かつ検出対象となる異常に多様性がある画像認識タスクにおいて、正常データのみを用いた深層生成モデルによる異常検知手法の有用性が既往の研究で示されている。特に VAE を使った手法は均一な部品を対象とした異常検知に活用されている。本研究では、河川護岸における変状種別に依存しない異常検知手法を確立することを目的とし、VAE によるブロック単位での異常検知を検討する。実験を通じてデータセット作成に関する指針を明らかにするとともに、護岸に適した異常度算出法を示し、護岸ブロックを対象とする VAE 異常検知が多様な変状の検出に有効であることを示した。

Key Words: Deep Learning, Anomaly detection, Variational Auto-Encoder, Maintenance

1. はじめに

(1) 背景

近年、高度経済成長期に建設された多くの河川管理施設の経年劣化が進行し、維持管理が課題となっている。これに対し国土交通省では、平成 25 年に河川法の改正を行い維持管理の義務を明確化した¹⁾。また平成 27 年 3 月には堤防及び護岸、樋門・樋管について点検結果評価要領(案)を策定し、さらに平成 31 年 4 月に点検・評価要領²⁾として改定を行った。この点検・評価要領は、1 年に 1 回以上の頻度で実施される目視を主体とした点検と評価方法についてとりまとめている。しかし、国と地方がそれぞれ管理している現在の法河川の総延長は 144,005.4km であり³⁾、これをすべて目視点検することは現実的でない。また、人によって評価にばらつきがあることや、豊富な経験を有する技術者の数が減少していることが問題とされている。

道路橋やトンネルでは、平成 31 年に改定された点検要領⁴⁾により、近接目視と同等の情報が得られると判断された場合には代替手法の採用が可能とされている。また令和 4 年には、第 4 回デジタル臨時行政調査会において目視規制等を含むアナログ規制を定める法律や政省令を改正する方針が示された⁵⁾。このような背景から、

河川点検においても近接目視と同等の情報が得られる、または目視点検を省力化する新技術の開発・導入が促進されると考えられる。特に、近年著しく進展している機械学習の一種である深層学習を用いた維持管理手法が期待されている。

(2) 既往研究

都市部を流れる河道法面には、一般的にコンクリート護岸が張り巡らされている。コンクリート構造物の点検において、深層学習を用いた研究は盛んに行われている。まず分類モデルによって画像から変状を検出する手法がある。例えば、全ら⁶⁾は Random Forest と CNN を組み合わせたコンクリート表面のひび割れ検出手法を開発した。横山ら⁷⁾は CNN を用いてコンクリート構造物の表面写真からひび割れやエフロレッセンス、チョーク文字を自動検出する検出器を開発した。山根ら⁸⁾は Mask R-CNN を用いた Semantic Segmentation によりコンクリート表面のひび割れをピクセル単位で検出する手法を提案した。また、物体検出モデルによって変状とその位置を推定する手法もある。例えば、野村ら⁹⁾は YOLOv2 を用いてコンクリート表面のひび割れ箇所検出システムを開発した。青島ら¹⁰⁾は Faster R-CNN と YOLOv3 を用いてひび割れや漏水・遊離石灰、剥離・鉄筋露出といったコンクリート変

状の種類と位置を自動検出し、2手法の精度比較を行った。このように、コンクリート構造物の画像から変状を自動検出する手法は様々提案されているが、これらは全て入力データとそれに対する正解を組み合わせた教師データを必要とする教師あり学習に基づいている。

教師あり学習による変状検出手法には、現場への導入にあたって2つの懸念点が挙げられる。まず1つ目は、高精度なモデルを作成するためには大量の教師データを必要とする点である。青島ら¹²⁾の研究でも教師データの充実を今後の課題として挙げており、変状のある画像を収集することは容易でない。また、データ量に伴ってそれに正解を付与するアノテーション作業も増えるためコストがかかる。GAN¹³⁾を用いて教師データを生成する手法¹⁴⁾やAttention機構を用いて教師データ作成のコストを削減する手法¹⁵⁾も研究されているが、生成した教師データの多様性の欠如やノイズ増加などそれぞれ課題を抱えている。2つ目の懸念点は、学習した種類の変状しか検出できない点である。河川においても斎藤ら¹⁶⁾がSemantic Segmentationを用いて河川護岸のひび割れを検出する手法を提案しており、これは点検・評価要領で定められた管理指標の一つであるひび割れを検出することはできるが、ひび割れ以外の変状検出には対応していない。点検・評価要領に示される護岸の点検項目とその状況例を整理した表-1に示すように、実際の点検では欠損や陥没など様々な変状を記録する必要があるため、特定の変状を精度良く検出することとは別に、網羅的にあらゆる変状を検出することが望ましい。教師あり学習においても、教師データのクラス数を増やすことで検出対象の種類を拡張することは可能であるが、頻出しな変状は教師データ数が少ないため検出が難しく、また未知の変状が現れたとき、それを検出できない可能性がある。

一方、十分な教師データを用意できない場合、正常データのみを用いたGANやVAE¹⁷⁾等の教師なし学習による異常検知手法が提案されている。この手法は正常データのみを学習させるため、収集困難な異常データが不要である。さらに、正常データの分布に対する外れ値を異常と定義することで、あらゆる変状や未知の変状の検出が可能である。異常検知は医療での画像診断や工業製品の不良品検知など幅広い分野で使用されており¹⁸⁾⁻²⁰⁾、土木分野でも道路の監視^{21),22)}や道路橋の変状検出²³⁾における手法が研究されている。しかし、河川護岸への適用事例は存在しない。

(3) 本研究の目的と検討内容

本研究は、河川護岸の点検において特定の変状種別に依存しない異常検知技術を確認することを目的とした基礎研究である。

異常検知に用いられる教師なし深層学習モデルのうち、

表-1 点検・評価要領における護岸の点検項目

変状種別	状況例
護岸の破損	<ul style="list-style-type: none"> ・目地の開き, クラック ・欠損 ・空洞化 ・湧き水 ・陥没, 沈下 ・段差 ・樹木の侵入
はらみ出し	<ul style="list-style-type: none"> ・はらみ出しによる目地の開き, クラック ・はらみ出しによる段差 ・はらみ出しによる護岸の破損
基礎部の洗掘	<ul style="list-style-type: none"> ・根固工の沈下 ・基礎工の浮き上がり
端部の浸食	<ul style="list-style-type: none"> ・端部付近の洗掘 ・天端保護工の流出



図-1 護岸画像例

VAE (Variational Auto-Encoder, 変分自己符号化器) は工業製品などの均一な部品に対する異常検知手法として有効性が示されている¹⁸⁾。図-1に示すように、河川護岸もほぼ均一な部品(ブロック)で構成されていることから、一つのブロックを異常検知の計測単位として用いることが可能であると考えた。そこで本研究では、VAEによるブロック単位での河川護岸の異常検知を試みる。また正常/異常を判断する指標として、護岸の異常検知に適した異常スコアについて検討する。さらに本手法の現場への導入を想定し、実河川におけるブロックの異常検知を行う。既往研究⁷⁾⁻¹²⁾では、変状を精度良く検出することは取り込まれているが、領域全体に対する変状の割合やそれによる修繕優先度を示した事例は少ない。本研究では異常ブロック率=異常ブロック数/全ブロック数を求め、任意区間(例えば画像1枚)毎の健全性評価を行う仕組みを提案する。

2. 教師なし深層学習による異常検知手法

教師なし学習による異常検知では一般に生成モデルが用いられる。生成モデルとは、対象データの分布 $P(x)$ にできるだけ近い確率分布 $Q(x)$ を学習し、 $Q(x)$ に従ってデータをサンプリング $x' \sim Q(x)$ するモデルである。この x' の生成は再構成と呼ばれる。生成モデルが正常データのみを学習した場合、モデルは正常箇所の特徴を再構

成することは可能だが、異常箇所の特徴を再構成することはできない。この特性を利用し、一般に、入力と再構成の差分が大きいのを異常と判断することができる。

深層学習を用いた生成モデルを深層生成モデルと呼ぶ。画像を対象とした深層生成モデルによる異常検知では、GAN や VAE などがベースモデルとして広く使用されている。GAN は高品質、高解像度な画像を生成できるが、対象データの分布の多様性を捉えられないこと (mode collapse) や、学習が安定しないといった課題がある²⁴⁾。本研究では、学習が安定しており比較的取り組みやすい VAE を用いた異常検知を適用する。VAE には様々な派生モデルが提案されているが、本研究では予備実験 (付録 1 参照) において直観的に再構成精度が高いと判断したモデルの一つである DIP VAE²⁵⁾を採用する。VAE による再構成画像はぼやけてしまうことが問題とされているが²⁶⁾、ここでは AE-Grad²⁷⁾と呼ばれる手法を用いて一般的な再構成画像よりもぼやけのない鮮明な画像の生成を行う。

本章ではまず VAE とその派生モデルである β -VAE²⁸⁾、さらに β -VAE よりも再構成精度が高いモデルとして提案されている DIP VAE について述べた後、AE-Grad による画像生成手法について述べる。

(1) VAE

VAE は、入力データ x を低次元の潜在変数 z に圧縮する符号化器 (Encoder) と潜在変数 z から入力データ x を再構成する復号化器 (Decoder) で構成される深層生成モデルである。潜在変数とは、入力データの特徴や要約を表し、直接観測することができない変数である。VAE では、この潜在変数が正規分布からサンプリングされると仮定する。

生成モデルが定義する生成分布の事後分布 $P_\theta(z|x)$ は解析的に解くことは困難であるため、VAE では $P_\theta(z|x)$ に近似する確率分布 $Q_\phi(z|x)$ を推定する。2 つの確率分布の類似性は KL ダイバージェンスで求められ、KL ダイバージェンスの最小化問題を解くことで近似分布 $Q_\phi(z|x)$ が得られる。このとき、入力データ分布の対数周辺尤度と変分下限 (ELBO) の差は KL ダイバージェンスに等しく、対数周辺尤度 = 変分下限 (最大) のとき KL ダイバージェンスは最小となる。つまり、ELBO を最大化したとき KL ダイバージェンスは最小化され、 $Q_\phi(z|x)$ は $P_\theta(z|x)$ に最も近似した分布となる。以上より、VAE の目的関数は次式で表される。

$$\begin{aligned} & \operatorname{argmin}_{\theta, \phi} \mathbb{E}_x [KL(Q_\phi(z|x) \| P_\theta(z|x))] \\ & = \operatorname{argmax}_{\theta, \phi} \mathbb{E}_x \left[\mathbb{E}_{z \sim Q_\phi(z|x)} [\log P_\theta(x|z)] \right. \\ & \quad \left. - KL[Q_\phi(z|x) \| P(z)] \right] \end{aligned} \quad (1)$$

a) β -VAE

β -VAE は、式(2)に示すように、調整可能なハイパーパラメータ $\beta (> 1)$ によって目的関数における正則化項の寄与を大きくした VAE の派生モデルである。

$$\begin{aligned} & \operatorname{argmin}_{\theta, \phi} \mathbb{E}_x [KL(Q_\phi(z|x) \| P_\theta(z|x))] \\ & = \operatorname{argmax}_{\theta, \phi} \mathbb{E}_x \left[\mathbb{E}_{z \sim Q_\phi(z|x)} [\log P_\theta(x|z)] \right. \\ & \quad \left. - \beta KL[Q_\phi(z|x) \| P(z)] \right] \end{aligned} \quad (2)$$

パラメータ β によって潜在変数 z の確率分布 $Q_\phi(z|x)$ を標準正規分布 $\mathcal{N}(0, I)$ に近づける効果が促進されるため、各 z は 0 に漸近する。しかしすべての z が 0 になると入力データに関わらず同じものが再構成されるため、必要最低限の z を残した確率分布 $Q_\phi(z|x) = \mathcal{N}(0, I)$ となる。共分散行列が単位行列ということは、お互いの次元の共分散が 0、つまり潜在変数 z の各次元がお互いに独立であることを表す。これにより、 β -VAE は入力データに内在する複数の特徴因子を独立した因子として取り出した disentangle な潜在表現を学習可能とされている。

b) DIP VAE

DIP VAE は、 β -VAE と同様に disentangle な潜在表現を学習するモデルとして提案されている。一般に、真の事後分布 $P(z)$ と生成モデルが定義する事後分布 $P_\theta(z)$ は近い場合もあるが、全く同じではない。従って ELBO を最大化すると $KL[Q_\phi(z) \| P(z)]$ の値も小さくなる可能性はあるが、 $KL[Q_\phi(z) \| P(z)]$ と $\mathbb{E}_x [KL(Q_\phi(z|x) \| P_\theta(z|x))]$ の差が大きくなり得る。そこで DIP VAE では、式(3)に示すように、正則化項として最小化すべき適切な距離 $D(Q_\phi(z) \| P(z))$ を VAE の目的関数に追加している。

$$\begin{aligned} & \operatorname{argmin}_{\theta, \phi} \mathbb{E}_x [KL(Q_\phi(z|x) \| P_\theta(z|x))] \\ & = \operatorname{argmax}_{\theta, \phi} \mathbb{E}_x \left[\mathbb{E}_{z \sim Q_\phi(z|x)} [\log P_\theta(x|z)] \right. \\ & \quad \left. - KL[Q_\phi(z|x) \| P(z)] - \lambda D(Q_\phi(z) \| P(z)) \right] \end{aligned} \quad (3)$$

ここでは D の最小化の代替として、 $P(z)$ と $Q_\phi(z)$ の 2 つの分布のモーメントのマッチングを行う。2 つの分布の共分散を合わせるということは、 $P(z)$ が $\mathcal{N}(0, I)$ の場合、 $z \sim Q_\phi(z)$ の次元を無相関化することに等しい。 $z \sim Q_\phi(z)$ の共分散は式(4)で表される。このとき、式(3)は式(5)に変換できる。

$$\begin{aligned} Cov_{Q_\phi(z)}[z] & = \mathbb{E}_{P(x)} Cov_{Q_\phi(z|x)}[z] \\ & \quad + Cov_{P(x)} \left(\mathbb{E}_{Q_\phi(z|x)}[z] \right) \end{aligned} \quad (4)$$

$$\begin{aligned} & \operatorname{argmax}_{\theta, \phi} ELBO(\theta, \phi) - \lambda_{od} \sum_{i \neq j} \left[Cov_{Q_\phi(z)}[z] \right]_{ij}^2 \\ & \quad - \lambda_d \sum_i \left(\left[Cov_{Q_\phi(z)}[z] \right]_{ii} - 1 \right)^2 \end{aligned} \quad (5)$$

ここで λ は目的関数における正則化項の寄与を制御するパラメータであり、 $\lambda_{od} = 5, \lambda_d = 10$ とした。論文²⁵⁾で

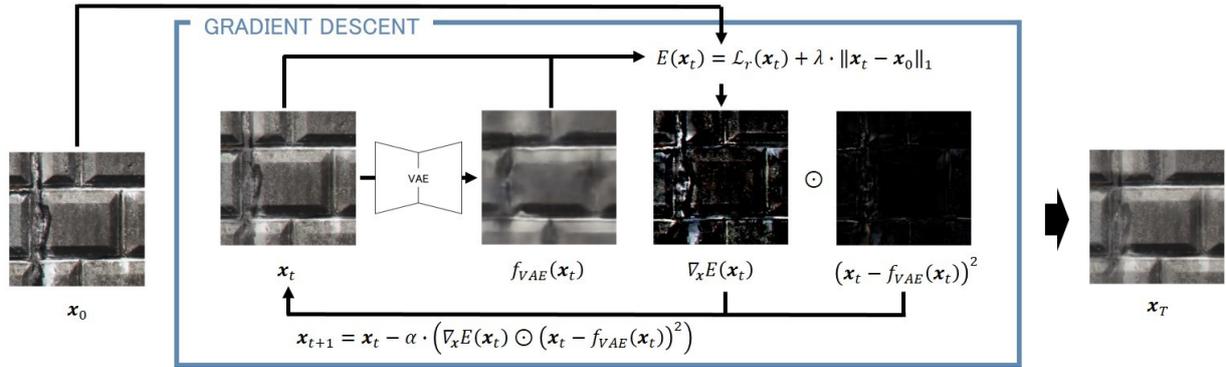


図-2 AE-Grad 概要

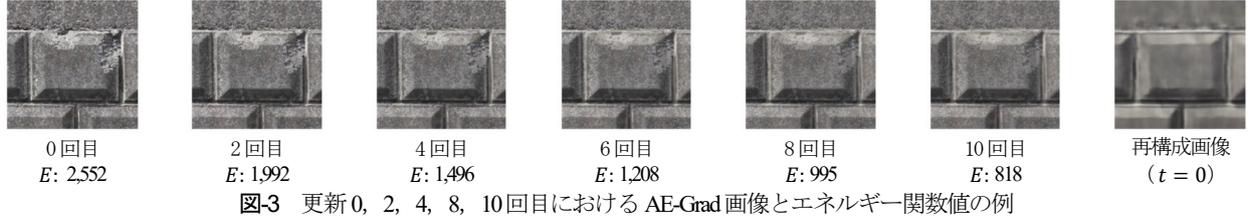


図-3 更新 0, 2, 4, 8, 10 回目における AE-Grad 画像とエネルギー関数値の例

は、VAE や β -VAE と比較して DIP VAE が disentanglement と再構成品質の両方において優位であることが示されている。

(2) AE-Grad

VAE では、確率分布 $P(x|z)$ を明示的にモデル化して最尤学習を行っているため、再構成画像がぼやけるとされている。AE-Grad は、式(6)に示す再構成誤差と正則化項で定義されたエネルギー関数 E を勾配降下させることにより、入力画像のうち正常箇所のパixelは入力画像と同じ値に保ちつつ、異常箇所のパixelを正常な状態 (VAE の再構成結果) に近づける手法である。

AE-Grad の概要を図-2 に示す。まず画像 x_t を VAE に入力すると再構成画像 $f_{vae}(x_t)$ が出力される。その後エネルギー関数 $E(x_t)$ を求め、入力 x_t から勾配ベクトル $\nabla_x E(x_t)$ を引き、画像 x_{t+1} を生成する。このとき式(8)に示すように、特に誤差が大きいパixelに重みを付けることで収束を早めている。ここで \odot はアダマール積であり、画像を行列と考えた場合の要素・成分ごとの積を表す。さらに、生成画像 x_{t+1} を再び VAE に入力して同様の操作を行うと x_{t+1} から x_{t+2} に画像が更新される。このループを任意の回数 (T 回) 繰り返して生成された画像 x_T が AE-Grad による最終的な出力画像となる。以上のように、VAE への入力を反復的に更新することにより、VAE のボトルネックによる高周波情報の損失 (ぼやけ) を回避することができ、従来の再構成画像よりも鮮明な画像を生成できる。

$$E(x_t) = L_r(x_t) + \lambda \cdot \|x_t - x_0\|_1 \quad (6)$$

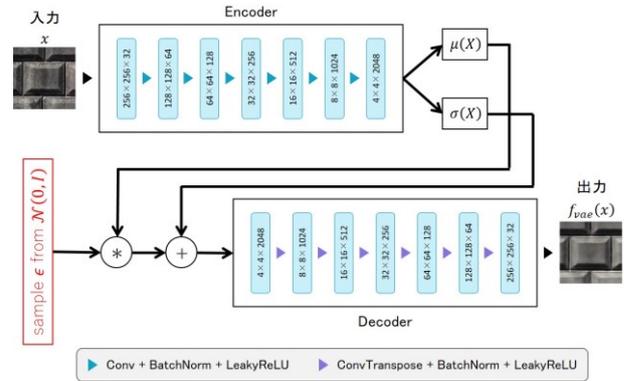


図-4 モデル概要

表-2 計算条件一覧

計算条件	条件内容
入力サイズ	(256, 256, 3)
エポック数	300
バッチサイズ	144
学習率	0.0005
潜在変数次元数	128

$$L_r(x_t) = \frac{1}{n} \sum_{i=1}^n (f_{VAE}(x_t)_i - x_{t_i})^2 \quad (7)$$

$$x_{t+1} = x_t - \alpha \cdot \nabla_x E(x_t) = x_t - \alpha \cdot (\nabla_x E(x_t) \odot (x_t - f_{VAE}(x_t))^2) \quad (8)$$

ここで λ は正則化重み、 α は学習パラメータである。本研究では、AE-Grad の適用後、ブロックの変状が十分に正常な状態に近づくことを確認した $\lambda = 0.5$ 、 $\alpha = 0.01$ に値を設定した。勾配ベクトル $\nabla_x E(x_t)$ は初期値を 0 とし、画像を生成する度に再構成誤差の差分を微分することで求められる。

AE-Grad を用いて生成した画像 (以降、AE-Grad 画像

被写体	データセット A	データセット B	画像例
ブロックのみ	950	950	
ブロックと 水抜口、梯子 (植生も含む)	-	431	
ブロックと 植生	-	1,848	
元画像合計枚数	950	3,229	
Data Augmentation 後 画像合計枚数	13,300	45,206	

図-5 学習データ内訳



図-6 Data Augmentation (これらに加えて左右反転も行っている)

	クラス	画像例				
テストデータA	正常 植生なし					
	異常 植生なし	表面剥離 	目地開き 	エフロレッセンス 	溶出物 	欠損
テストデータB	正常 植生あり					
	異常 植生あり	表面剥離 	目地開き 	エフロレッセンス 	溶出物 	

図-7 テストデータ画像例

と呼ぶ) 例を図-3に示す。更新回数の増加に伴ってエネルギー関数の値が減少し、異常箇所のピクセルが正常な状態に近づいていることが確認できる。たとえば更新回数 10 回目の AE-Grad 画像と再構成画像を比較すると、AE-Grad 画像の方がぼやけが少なく鮮明な画像であるといえる。

AE-Grad は元々変状箇所を正確にセグメント化することを目的として提案された手法であるが、副次的な効果として異常検知精度の向上につながる可能性があると考え、本研究ではこれについて検証を行った。

3. VAE による河川護岸の異常検知実験

本章では、正常データ(健全なブロック画像)を用いて VAE モデルの学習を行った後、テストデータに対して異常検知を実施しその結果を評価する。また、河川護岸に適した評価指標(異常スコア)の算出法を検討する。

(1) VAE モデルとデータセットの作成

a) ネットワーク設計

本研究では、深層生成モデルとして DIPVAE を採用す

正常/異常	被写体または変状種類	入力画像	モデル A		モデル B	
			再構成結果	AE-Grad 画像	再構成結果	AE-Grad 画像
正常	ブロックのみ					
	水抜口					
	植生					
異常	欠損					
	目地開き					
	表面剥離					
	溶出物					
	エフロ 小					
	エフロ 大					
	エフロ 植生					

図-8 モデル A・Bにおける再構成結果 (DIP VAE) と 10回更新を行った AE-Grad 画像例

る。モデルの概要を図-4に示す。Encoder と Decoder は共に 7層の畳み込み層で構成し、計算条件は表-2の通りである。

b) 学習データ

本研究で使用するブロック画像は、図-1に示すような護岸の全体画像から手作業で各ブロックの切り出しを行っている。本研究では、被写体がブロックのみの画像で構成されたデータセット A と、これに水抜口や梯子、植生を含んだ画像を加えたデータセット B の 2 種類の学習用データセットを作成した。内訳を図-5に示す。回転やシフト、左右反転の Data Augmentation を行い (図-6) 、

データセット A は 13,300 枚(元画像 950 枚)、データセット B は 45,206 枚(元画像 3,229 枚)の画像を用意した。本研究における Data Augmentation の有効性については付録 2 で言及する。

医療や工場の現場では、一般的に背景や光源などの撮影条件が常に同じ画像を用いた異常検知を行うことが可能である。しかし護岸画像では、コンクリートの質感や色合いが環境によって異なり、また植生や水抜口などの付属物が散在し、正常データにおける多様性が高い。このことから、生成モデルが正常データの分布を推定することが難しく、従って護岸への異常検知の適用は他分野

に比べて困難であると想定した。そのため本研究では、まず被写体をブロックのみに絞ることで正常データの特徴を単純にしたデータセットAによって護岸への異常検知適用の基礎検討を行う。その後データセットBにより、正常データの特徴が多く複雑な場合にも護岸への異常検知適用が可能であるかを段階的に検討する。以降、データセットAを学習したものをモデルA、データセットBを学習したものをモデルBとする。

c) テストデータ

テストデータも学習データと同様に2種類のデータセットA・Bを作成した。画像例を図-7に示す。ここでは被写体がブロックのみの正常/異常画像と、植生や水抜口、梯子を含む正常/異常画像を各100枚用意した。異常画像に含まれる変状は表面剥離、目地開き、エフロレッセンス、溶出物、欠損である。

(2) 再構成結果

各モデルにおけるテストデータの再構成結果と10回更新を行ったAE-Grad画像を図-8に示す。

正常画像では両モデルとも入力画像に非常に似た画像が再構成されており、モデルが健全なブロックの特徴を学習したことが確認できる。植生や水抜口を学習したモデルBでは、それらが再構成結果に反映されている。ただし、コンクリート部分に比べると再構成精度は劣り、特に植生は詳細な形状を再構成できていない。一方異常画像では、欠損等の異常箇所は再構成されていないことが確認できる。ブロック全体に変状が生じている場合、輪郭のはっきりしない不安定なブロック画像が再構成された。

正常画像では、AE-Gradを用いても入力画像からの変化が少なく、AE-Grad画像が再構成結果よりも鮮明であることが確認できる。またモデルBでは再構成結果に水抜口や植生が反映されているため、AE-Gradの適用後もそれらが残存している。異常画像では、欠損のような部分的な変状の場合、AE-Gradによって異常箇所が正常な状態に変化したことが明確に分かる。

(3) 異常スコアの検討

本研究における異常スコアは、一つのブロックに対して正常か異常かを判断する指標である。ここでは一般的な画像の類似度指標であるSSIM²⁹⁾やZNCC³⁰⁾、MSEだけでなく、AE-Gradにおけるエネルギー関数に基づく新しい異常スコアを提案し、合計6種類の異常スコアについて比較を行う。図-9に勾配降下に伴うエネルギー関数の変位を示す。オレンジ色が異常画像を入力した場合、青色が正常画像を入力した場合の一般的な挙動である。以下に各異常スコアについて述べる。

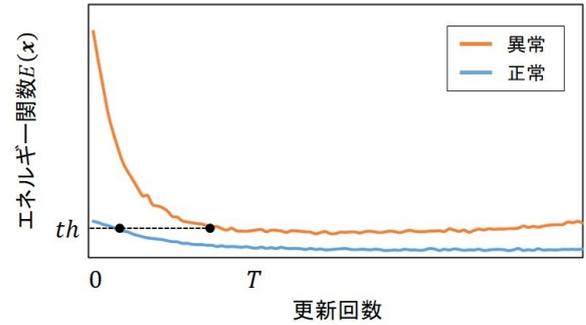


図-9 勾配降下に伴うエネルギー関数の変位

a) 画像類似度指標

画像間の類似性を表す評価尺度には様々なものが提案されている。代表的なものとして、SSIM (Structural SIMilarity) やZNCC (Zero-mean Normalized Cross Correlation, 零平均正規化相互相関)、MSE (Mean Squared Error, 平均二乗誤差) が挙げられる。符号化前と符号化後の画像をそれぞれ x 、 y とすると、それぞれ以下の式で表される。

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

$$ZNCC = \frac{(x - \bar{x}, y - \bar{y})}{\|x - \bar{x}\| \|y - \bar{y}\|} \quad (10)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2 \quad (11)$$

SSIMは、輝度とコントラスト、構造の3つの要素が符号化前後でどの程度変化したかによって画質を評価する指標である。式(9)における μ はピクセル平均値、 σ_x と σ_y は標準偏差、 σ_{xy} は共分散である。また、 c_1 と c_2 は分母の値が非常に小さくなった場合に評価値が不安定にならないための定数である。

ZNCCは、ベクトルのなす角のコサインを類似度とするNCC (Normalized Cross Correlation, 正規化相互相関) に平均値の減算(中心化)とノルムの除算(正規化)を組み合わせた指標である。ZNCCは輝度の変動に頑健な類似度算出手法とされており、テンプレートマッチングにも活用されている。式(10)における (x, y) は内積、 $\|x\|$ はノルム、 \bar{x} は平均値である。

MSEは、画像間で同じ位置同士のピクセル値の差分の2乗を計算し、それを平均したものである。これはVAEの損失関数としても用いられる。

b) 異常スコアb: 再構成誤差

入力画像と再構成画像のMSEを求める。これは一般的に再構成誤差と呼ばれ、異常検知における代表的な異常度の一つである。AE-Gradにおいて、更新回数0回目のエネルギー関数は正則化項が $\lambda \cdot \|x_0 - x_0\|_1 = 0$ のため再構成誤差に等しくなる($E(x_0) = \mathcal{L}_r(x_0)$)。図-9より、更新回数0回目の値を比較すると、正常画像より

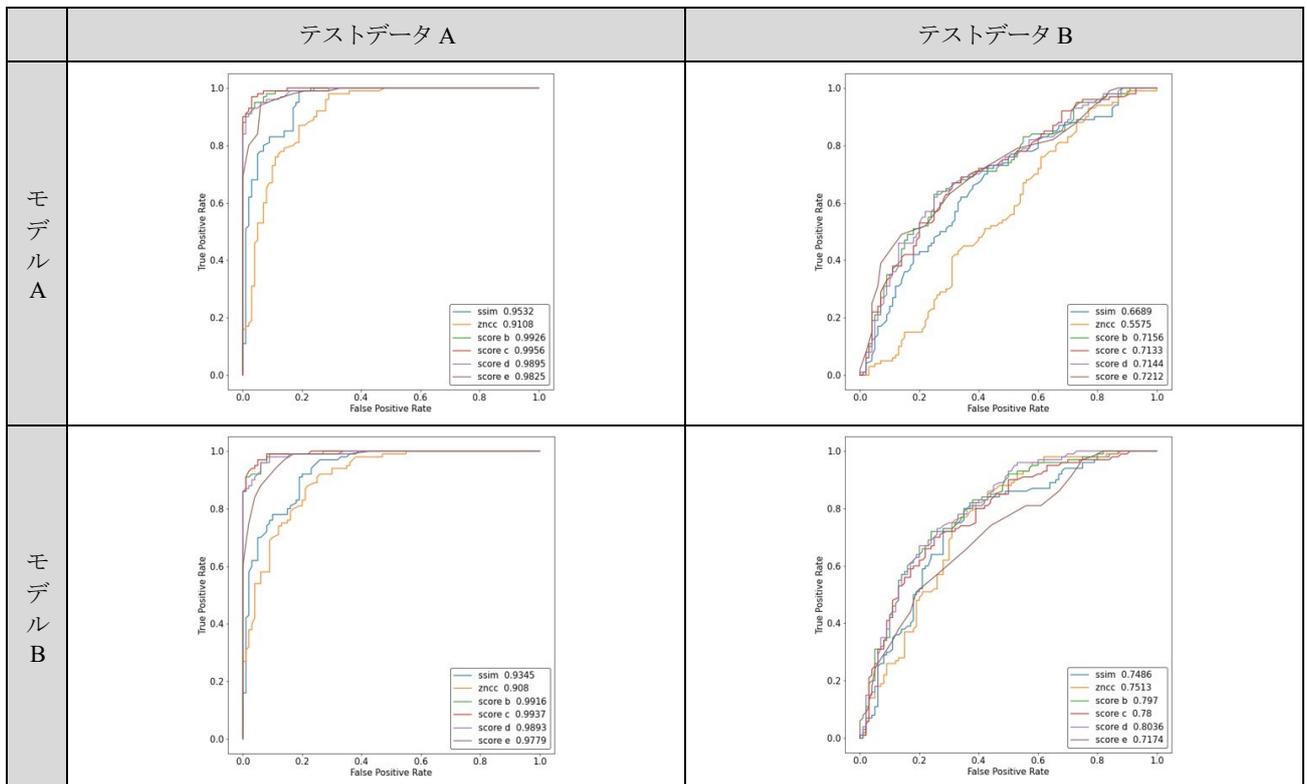


図-10 モデル A・Bにおける各テストデータに対する ROC 曲線

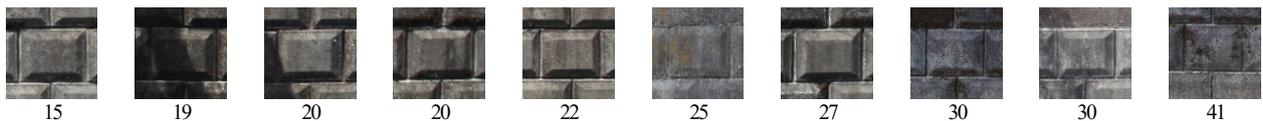


図-11 異常スコアが低い正常画像と異常スコア値 (異常スコア c)

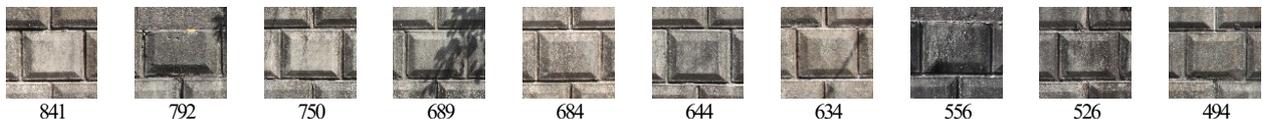


図-12 異常スコアが高い正常画像と異常スコア値 (異常スコア c)



図-13 異常スコアが低い異常画像と異常スコア値 (異常スコア c)



図-14 異常スコアが高い異常画像と異常スコア値 (異常スコア c)

も異常画像の方が値が大きく、異常スコア b は有効な指標であるといえる。

c) 異常スコア c: 入力画像と AE-Grad 画像の MSE

入力画像と AE-Grad 画像の MSE を求める。図-9 より、更新回数 0 回目から T 回目 (本研究では 10 回) までのエネルギー関数の変位は異常画像の方が大きく、AE-Grad 画像のピクセル値は入力画像のピクセル値から大きく変化しているはずである。従って、正常画像よりも異常画像の方が異常スコア c の値が大きくなると考えた。

d) 異常スコア d: 異常スコア b, c の差分

異常スコア b と異常スコア c の差分を求める。両スコアとも入力画像との MSE を求めていることから、異常スコア d は入力画像を基準とした再構成画像と AE-Grad 画像の類似度といえる。AE-Grad は再構成結果が正常であるという前提の下、入力画像を再構成画像に徐々に近づけていく手法であることから、エネルギー関数は最小化すべき AE-Grad 画像と再構成画像の距離を表していると考えた。図-9 より、収束点における異常画像のエネルギー関数の値が大きいことから、異常スコア d は異常スコア b と異常スコア c の差分を求めると考えた。

ギ関数の値は異常画像の方が大きいため、AE-Grad 画像と再構成画像の距離は異常画像の方が大きく、類似度が低いと予測する。(類似度が低いほどスコアは大きい)

e) 異常スコア e : AE-Grad の更新回数

エネルギー関数がある閾値に減少するまでの AE-Grad の更新回数を求める。図-9 より、ある閾値 th (本研究では 800 に設定) に到達するまでの更新回数は、正常画像よりも異常画像の方が多いと考えた。

(4) 性能評価

各モデルにおいて、入力画像と再構成画像の ZNCC と SSIM, 異常スコア b, c, d, e の合計 6 種類の異常スコアによる性能評価を行った。ここでは、立花ら¹⁸⁾や青島ら²³⁾と同様に ROC 曲線 (Receiver Operating Characteristic Curve) の下部面積 (ROC-AUC) によって分析を行う。ROC 曲線は閾値を変化させたときの偽陽性率 (False Positive Rate: x 軸) と真陽性率 (True Positive Rate: y 軸) の変化をプロットしたものである。この ROC 曲線の下部面積を AUC (Area Under the Curve) と呼び、AUC は値が 1 に近いほど異常検知の性能が高いことを表す。モデル A とモデル B における各テストデータに対する ROC 曲線を図-10 に示す。まずテストデータ A では、両モデルとも SSIM と ZNCC による識別精度は低く、異常スコア b, c, d, e ではどれも AUC が 0.97 以上の高い異常検知性能を示した。中でも異常スコア c による評価がモデル A で 0.9956, モデル B で 0.9937 と最も AUC が高い結果となった。これは、VAE の再構成では全体的なぼやけによって正常箇所における入力との差分が大きくなる可能性があったが、AE-Grad 画像では正常箇所における入力との差分が抑えられ、異常箇所の差分が際立ったためであると考察する。一方テストデータ B では、モデル A で異常スコア e による評価が 0.7212, モデル B で異常スコア d による評価が 0.8036 と最も AUC が高い結果となった。

モデル A では植生や水抜口を含む画像は学習していないため、テストデータ A の結果よりも AUC の値が低いのは妥当である。一方モデル B は、モデル A より AUC の値は高いが、テストデータ A の結果に比べると識別精度が低い。このことから、生成モデルに定義させる正常データの多様性が高くなると、異常検知性能が低下することが示唆される。

テストデータ A に対し、モデル A を用いて異常スコア c による異常検知を行った結果のうち、正常/異常画像における異常スコア下位画像と上位画像各 10 枚をそれぞれ図-11 から図-14 に示す。正常画像では、比較的明るい色味のブロックに対して異常スコアが高い傾向にある。一方異常画像では、変状面積の小さいエフロレッセンスや欠損、目地開きは異常スコアが低く、変状面積の大きいエフロレッセンスや明るい色味のブロックに対する異

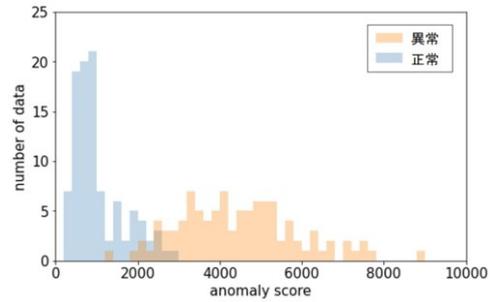


図-15 モデル A のテストデータ A に対する異常スコア b の度数分布

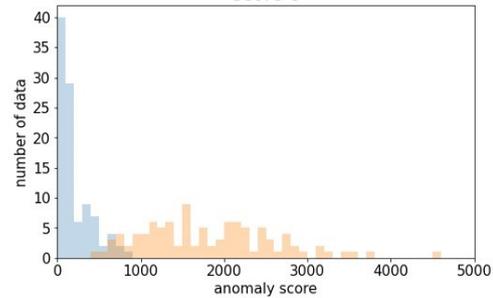


図-16 モデル A のテストデータ A に対する異常スコア c の度数分布

常スコアが高い傾向にある。

テストデータ A に対するモデル A による異常スコア b と異常スコア c の度数分布の結果を図-15 と図-16 に示す。これらを比較した結果、AE-Grad 画像は単純な再構成画像よりも入力画像に似ているためスコア c はスコア b に比べて MSE の値の大きさ (x 軸) は小さくなるが、全体的な分布傾向に大きな変化は見られなかった。このため、異常判別の精度に関してスコア b とスコア c に大きな差は見られなかったのだと考察できる。

4. 実河川への適用

これまでは個別のブロック画像の学習データ・テストデータを用いた実験結果を示した。本章では、本手法の点検現場への導入を想定し、実河川への適用を試みる。具体的には、河川護岸全体を撮影した画像に対し、前章で作成したモデル A を用いて異常スコア c による異常検知を行い、異常スコアが閾値未満のブロックを正常、閾値以上のブロックを異常と判断する。ここでは、テストデータ A に対するモデル A による異常スコア c の度数分布の結果 (図-16) から、正常画像がすべて正常と判断される異常スコア=900 に閾値を設定した。ただし、実河川画像では個別ブロックの正常/異常の網羅的な正解データは存在しないため、異常検知結果の目視による定性評価を行う。結果を図-17 に示す。左が元画像、右が異常ブロックの検出結果である。異常と判断されたブロックを四角の枠で囲い、枠の色は赤色に近いほど異常ス

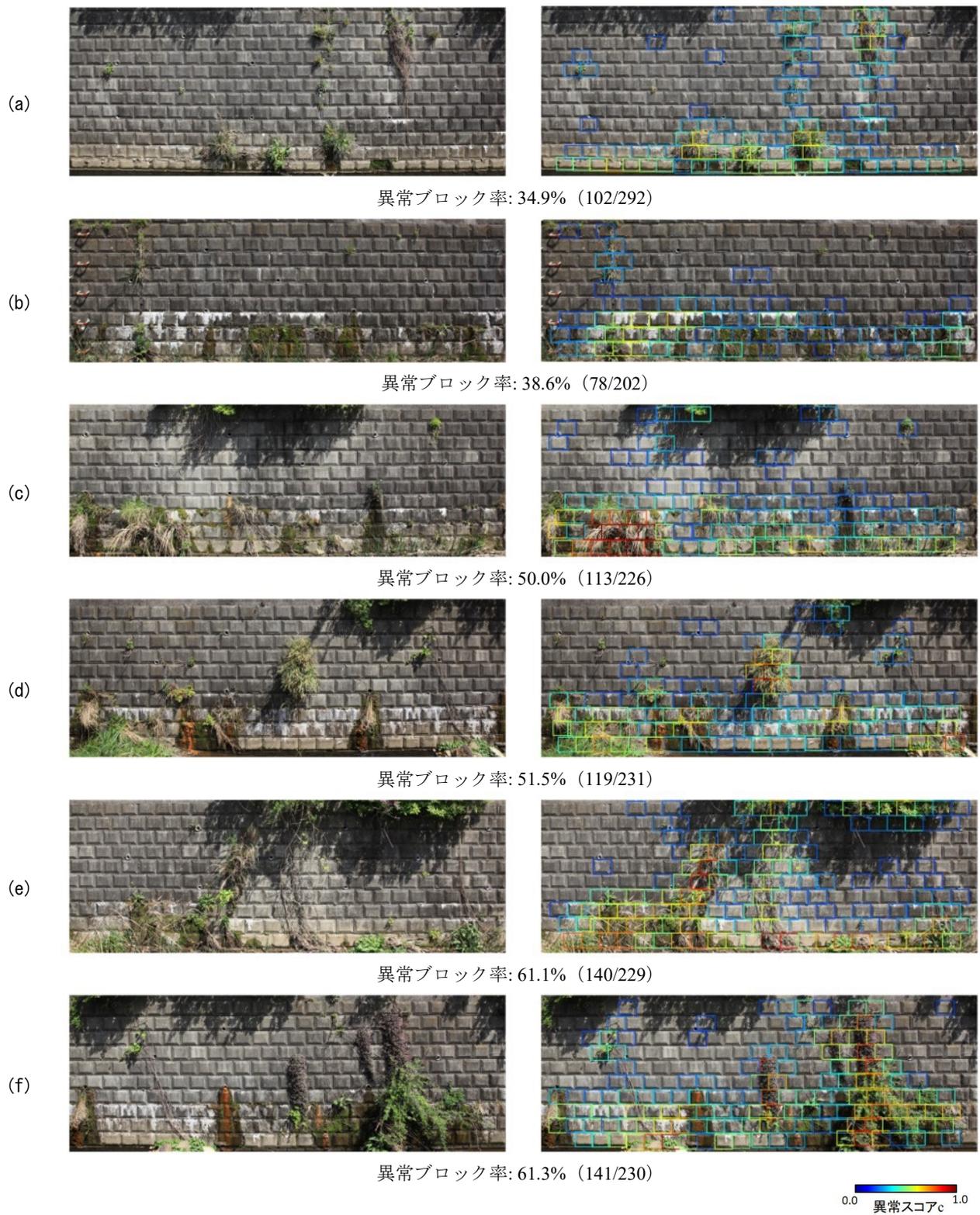


図-17 検出結果例（左：元画像，右：検出結果）と異常ブロック率（異常ブロック数/全ブロック数）



図-18 護岸ブロック外れ（左：元画像，右：検出結果）



図-19 各 VAE モデルの再構成結果

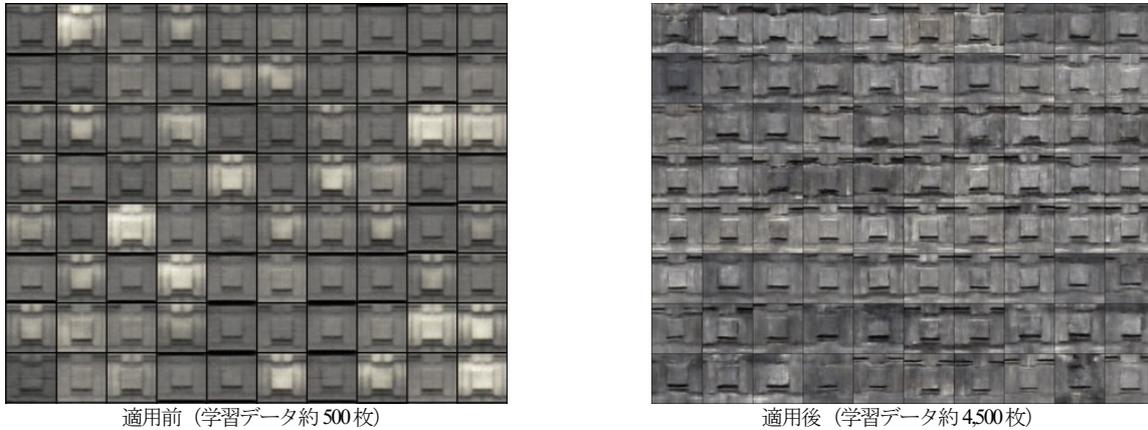


図-20 Data Augmentation 適用前後の再構成 (DIP VAE) の比較結果

コアの値が高く、青色に近いほど値が低いことを表す。ブロックの外れやエフロレッセンス、溶出物などの変状が検出されており、本手法が多様な変状を検出可能であることが確認できる。また図-18 に示すブロックの外れはデータ数の少ない変状の一つであり、こうした教師あり学習による検出が難しい変状も本手法では検出できていることが分かる。ただし、管理上異常として扱わない植生や水抜口、水面付近のブロック等も異常として検出されており、これらの対応は今後の課題である。

異常検知結果の活用例として、画像毎に異常ブロック率=異常ブロック数/全ブロック数を算出した。図-17 に示す各画像の異常ブロック率を比較すると、値が高いほど護岸全体に対する変状や植生の繁茂が多いことが確認できる。このことから、異常ブロック率をもとに任意区間（例えば画像 n 枚）での健全性を評価できる可能性が示唆される。

5. おわりに

本研究では、河川護岸の点検において特定の変状種別に依存しない異常検知技術を確立するために、VAE を用いた教師なし深層学習による異常検知の適用とその技術検証を行った。得られた主な知見を以下に述べる。

- (1) VAE により護岸ブロックが再構成できることを確認した。ただし、植生の詳細部分まで再構成することはできない。
- (2) 正常データの多様性が高いほど教師なし深層学習による異常検知性能は低下することが示唆された。
- (3) 一つの護岸ブロックに対して正常/異常を判断する

指標として、入力画像と AE-Grad 画像の MSE を異常スコアに用いることが有効だと示された。

- (4) 護岸全体に異常検知を適用し定性評価を行った結果、異常検知によって多様な変状を検出できていることを確認した。また、教師あり学習の適用が困難なデータ数の少ない変状についても異常検知によって検出が可能であることを確認した。

今後の課題について述べる。本研究では、多様な変状が検出される一方で、変状以外の植生や水抜口等を異常として検出する事例が多く見られた。しかし VAE で複雑な特徴を再構成することは困難であるため、入力と再構成または AE-Grad 画像の差分が大きくなることは避けられない。この対応として、異常検知の適用後、異常と判定された要因が植生に依るものか判別するためのクラス分類を適用することを検討している。また、本研究では異常検知のベースモデルとして DIPVAE を採用したが、その他の VAE モデルや VAE 以外の深層生成モデル (GAN や拡散モデル等) の検討が必要である。さらに、本研究では一つ一つのブロックを手作業で切り出して画像データを用意したが、効率的な切り出し手法の検討が必要である。

付録 1

本研究で使用する VAE モデルを決定するにあたり、オリジナル VAE、 β -VAE、disentangled β -VAE³¹⁾、SWAE³²⁾、DIPVAE、WassersteinVAE (IMQ Kernel)³³⁾ の 6 つのモデルについて検討した。エポック数を 50、それ以外の計算条件を表-2 の通りに設定し、データセット A を用いて学習を行った。各モデルの再構成結果を図-19 に示す。

付録 2

VAE において、Data Augmentation がすべてのオブジェクトに適しているわけではないことがしばしば指摘されている³⁹。そこで本研究では、Data Augmentation の適用前後での再構成精度の比較を定性的に行った(図-20)。その結果、Data Augmentation の適用前は非常に似たブロック画像が再構成されたが、適用後は再構成結果に多様性が見られるようになった。これより、本研究では Data Augmentation の適用は有効であると考えられる。

REFERENCES

- 1) 国土交通省: 水防法及び河川法の一部を改正する法律案新旧対照条文, p. 10, 2012.
- 2) 国土交通省 水管理・国土保全局 河川環境課: 堤防等河川管理施設及び河道の点検・評価要領, 2019.
- 3) 国土交通省 水管理・国土保全局: 河川データブック, p.197, 2021.
- 4) 国土交通省 道路局: 道路橋点検要領, 2019.
- 5) 国土交通省 道路局 国土・技術課: 道路トンネル点検要領, 2019.
- 6) アナログ規制撤廃, 閣議決定 目視・対面など 4000 条項, 日本経済新聞, 2022.6.7.
- 7) 全邦釘, 井後敦史: Random forest によるコンクリート表面ひび割れの検出, 土木学会論文集 F3 (土木情報学), 71 巻 2 号, pp. I_1-I_8, 2015. [Chun, P. and Igo, A.: CRACK DETECTION FROM IMAGE USING RANDOM FOREST, Journal of Japan Society of Civil Engineers, ser. F3 (Civil Engineering Informatics), Vol. 71, No. 2, pp. I_1-I_8, 2015.]
- 8) 全邦釘, 嶋本ゆり, 大窪和明, 三輪知寛, 大賀水田生: ディープラーニングおよび Random Forest によるコンクリートのひび割れ自動検出法, 土木学会論文集 F3 (土木情報学), 73 巻 2 号, pp. I_297-I_307, 2017. [Chun, P., Shimamoto, Y., Okubo K., Miwa C. and Ohga, M.: DEEP LEARNING AND RANDOM FOREST BASED CRACK DETECTION FROM AN IMAGE OF CONCRETE SURFACE, Journal of Japan Society of Civil Engineers, ser. F3 (Civil Engineering Informatics), Vol. 73, No. 2, pp. I_297-I_307, 2017.]
- 9) 横山傑, 松本高志: Deep Learning によるコンクリートの変状自動検出器の開発と Web システムの実装, 土木学会論文集 A2 (応用力学), 73 巻 2 号, pp. I_781-789, 2017. [Yokoyama, A. and Matsumoto, T.: DEVELOPMENT OF AN AUTOMATIC DETECTOR OF CONCRETE SURFACE DETERIORATIONS USING DEEP LEARNING AND IMPLEMENTATION OF WEB SYSTEM, Journal of Japan Society of Civil Engineers, ser. A2 (Applied Mechanics), Vol. 73, No. 2, pp. I_781-I_789, 2017.]
- 10) 山根達郎, 全邦釘: Deep learning による Semantic Segmentation を用いたコンクリート表面ひび割れの検出, 構造工学論文集, 65A 巻, pp. 130-138, 2019. [Yamane, T. and Chun, P.: Crack detection from an image of concrete surface based on semantic segmentation by deep learning, Journal of Structural Engineering, A, Vol. 65A, pp. 130-138, 2019.]
- 11) 野村泰稔, 松尾彩希, 阪口幸広, 古田均: 深層畳み込みニューラルネットワークに基づくコンクリート表面のひび割れ検出システム, 土木学会論文集 F6 (安全問題), 71 巻 2 号, pp. I_189-I_198, 2017. [Nomura, Y., Murao, S., Sakaguchi, Y. and Furuta, H.: CRACK DETECTION SYSTEM FOR CONCRETE SURFACE BASED ON DEEP CONVOLUTION NEURAL NETWORK, Journal of Japan Society of Civil Engineers, Ser. F6 (Safety Problem), Vol. 71, No. 2, pp. I_189-I_198, 2017.]
- 12) 青山亘佐, 河村伸哉, 中村聡, 中村秀明: 深層学習による画像認識を用いたコンクリート構造物の変状検出に関する研究, 土木学会論文集 E2 (材料・コンクリート構造), 74 巻 4 号, pp. 293-305, 2018. [Aoshima, K., Kawamura, S., Nakano, S. and Nakamura, H.: STUDY ON VARIANT EXTRACTION OF CONCRETE STRUCTURES USING IMAGE RECOGNITION BY DEEP LEARNING, Journal of Japan Society of Civil Engineers, Ser. E2 (Materials and Concrete Structures), Vol. 74, No. 4, pp. 293-305, 2018.]
- 13) Goodfellow, J. I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative Adversarial Nets., Advances in Neural Information Processing Systems, pp. 2672-2680, 2014.
- 14) 田村晃一郎, 原田隆郎: 敵対的生成ネットワークを用いた対候性鋼材のさび画像生成に関する基礎的研究, AI・データサイエンス論文集, 2 巻 J2 号, pp. 792-800, 2021. [Tamura, K. and Harada, T.: BASIC STUDY ON RUST IMAGE GENERATION OF WEATHERING STEEL BY GENERATIVE ADVERSARIAL NETWORKS, Intelligence, Informatics and Infrastructure, Vol. 2, No. J2, pp. 792-800, 2021.]
- 15) 泉翔太, 全邦釘: Attention 機構を用いた Deep Learning モデルによるひび割れ自動検出, AI・データサイエンス論文集, 2 巻 J2 号, pp. 545-555, 2021. [Izumi, S. and Chun, P.: CRACK DETECTION USING DEEP LEARNING WITH ATTENTION MECHANISMS, Intelligence, Informatics and Infrastructure, Vol. 2, No. J2, pp. 545-555, 2021.]
- 16) 齊藤彰儀, 上総虎智, 平木悠太, 天方匡純, 吉田武司: 深層学習によるコンクリート護岸劣化領域検出システムの開発, 情報処理学会デジタルプラクティス, 10 巻 2 号, pp. 288-303, 2019. [Saito, A., Kazusa, T., Hiraki, Y., Amakata, M. and Yoshida, T.: Development of a concrete revetment deterioration area detection system using deep learning, IPSJ digital practice, Vol. 10, No. 2, pp. 288-303, 2019.]
- 17) Kingma, P. D. and Welling, M.: Auto-Encoding Variational Bayes, ICLR, 2014.
- 18) 立花亮介, 松原崇, 上原邦昭: 深層生成モデルによる非正規化異常度を用いた工業製品の異常検知, 人工知能学会全国大会論文集, 2018. [Tachibana, R., Matsubara, T. and Uehara, K.: Anomaly Manufacturing Product Detection using Unregularized Anomaly Score on Deep Generative Models, Proceedings of the Annual Conference of JSAI, 2018.]

- 19) Bergmann, P., Löwe, S., Fauser, M., Sattlegger, D. and Steger, C.: Improving Unsupervised Defect Segmentation by Applying Structural Similarity To Autoencoders, VISPP 2019 - 14th International Conference on Computer Vision Theory and Applications, 2019.
- 20) 上原雅俊, 味曾野雅史, 中山浩太郎, 山口亮平, 河添悦昌, 大江和彦, 松尾豊: 生成モデルを利用した X 線写真に対する異常検知の試行と考察, 人口知能学会第二種研究会資料, 2016 巻, 2016. [Uehara, M., Misono, M., Nakayama, K., Yamaguchi, R., Kawazoe, Y., Ohe, K. and Matsuo, Y.: A study chest X-ray image anomaly detection with deep generative models, JSAI Technical Report, Type 2 SIG, Vol. 2016, 2016.]
- 21) 山本健生, 橋本敦史, 岡本大和: 平均画像に対する VAE 異常検知の適用による道路落下物検出, 人工知能学会論文集, 2021. [Yamamoto, Y., Hashimoto, A. and Okamoto, Y.: Fallen Object Detection on Road by Using VAE Anomaly Detection with Average Image, Proceedings of the Annual Conference of JSAI, 2021.]
- 22) Yasuno, T., Fujii, J., Ogata, R. and Okano, M.: VAE-iForest: Auto-encoding Reconstruction and Isolation-based Anomalies Detecting Fallen Objects on Road Surface, arXiv:2203.01193, 2022.
- 23) 青島亘佐, 中野聡, 徳永皓平, 中村秀明: 深層学習による異常検知手法を用いたコンクリート表面の変状検出, 土木学会論文集 A2 (応用力学), 75 巻 2 号, pp. I_559-I_570, 2019. [Aoshima, K., Nakano, S., Tokunaga, K. and Nakamura, H.: DAMAGE DETECTION OF CONCRETE SURFACE USING ANOMALY DETECTION METHOD BY DEEP LEARNING, Journal of Japan Society of Civil Engineers, ser. A2 (Applied Mechanics), Vol. 75, No. 2, pp. I_559-I_570, 2019.]
- 24) Razavi, A., Oord, A. V. D. and Vinyals, O.: Generating Diverse High-Fidelity Images with VQ-VAE-2, Advances in Neural Information Processing Systems, 2019.
- 25) Kumar, A., Sattigeri, P. and Balakrishnan, A.: VARIATIONAL INFERENCE OF DISENTANGLED LATENT CONCEPTS FROM UNLABELED OBSERVATIONS, ICLR, 2018.
- 26) Perarnau, G., Weijer, J. V. D., Raducanu, B. and Álvarez, J. M.: Invertible Conditional GANs for image editing, arXiv:1611.06355, 2016.
- 27) Dehaene, D., Frigo, O., Combrexelle, S. and Eline, P.: ITERATIVE ENERGY-BASED PROJECTION ON A NORMAL DATA MANIFOLD FOR ANOMALY LOCALIZATION, ICLR, 2020.
- 28) Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., Mohamed, S. and Lerchner, A.: Beta-vae: Learning basic visual concepts with a constrained variational framework, ICLR, 2017.
- 29) Wang, Z., Bovil, A. C., Sheikh, H. R. and Simoncelli, E. P.: Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, Vol. 13, No. 4, 2004.
- 30) 荒井敏, 長尾智晴: 深層ゼロ平均正規化相互相関ネットワーク, 情報処理学会研究報告, 2017.[Arai, A. and Nagou T.: Deep Zero-mean Normalized Cross Correlation Networks, IPSJ Technical Report, 2017.]
- 31) Burgess, C. P., Higgins, I., Pal, A., Matthey, L., Watters, N., Desjardins, G. and Lerchner, A.: Understanding disentangling in β -VAE, arXiv:1804.03599, 2018.
- 32) Kolouri, S., Pope, P. E., Martin, C. E. and Rohde, G. K.: Sliced Wasserstein Auto-Encoders, ICLR, 2019.
- 33) Tolstikhin, I., Bousquet, O., Gelly, S. and Schoelkopf, B.: Wasserstein Auto-Encoders, ICLR, 2018.
- 34) Bergmann, P., Batzner K., Fauser, M., Sattlegger, D. and Steger, C.: Beyond Dents and Scratches: Logical Constraints in Unsupervised Anomaly Detection and Localization, International Journal of Computer Vision, Vol. 130, No. 4, pp. 947-969, 2022.

(Received June 24, 2022)

(Accepted November 30, 2022)

ANOMALY DETECTION OF REVETMENT BY UNSUPERVISED DEEP LEARNING USING VARIATIONAL AUTO-ENCODER

Yukino TSUZUKI, Ryuto YOSHIDA, Junichi OKUBO,
Junichiro FUJII and Takayoshi YAMASHITA

Anomaly detection for infrastructure by deep learning using supervised learning has problems that it is difficult to prepare data set on abnormal data, consequently unlearned deterioration cannot be detected. On the other hand, several prior works have shown the effective for the anomaly detection based on generative models training using only normal data in image recognition tasks with little training data and diversity of the deteriorations to be detected. In particular, VAE-based methods have been used for anomaly detection on uniform parts. For the purpose of building the anomaly detection method of revetment that is independent of the type of deterioration, we examined and considered for the anomaly detection of revetment in block units using VAE. This paper empirically clarifies the guideline for preparation of data set and proposes an anomaly calculation method suitable for revetment. We also discussed t detecting various types of deteriorations.