

# 河川護岸へのゼロショット異常検知の応用

○七海 友康 †, 都築 幸乃 †, 大久保 順一 †, 藤井 純一郎 †

†: 八千代エンジニアリング株式会社 技術創発研究所

tm-nanaumi@yachiyo-eng.co.jp

概要: 本研究は河川護岸ブロックの外観検査に対してゼロショット異常検知の枠組みを導入し、最適化に用いるデータの違いが検知性能に与える影響を分析する。評価対象外のデータを用いて学習可能ベクトルを最適化するゼロショット異常検知手法を用いて、他の河川護岸のデータで最適化を行った後、追加の訓練なしで対象の河川護岸に適用する。このアプローチにより、2つの河川のデータを用いたクロスドメイン評価では、対象河川の護岸データを用いずに教師なし異常検知手法 PatchCore を上回る AUROC および AUPRO を達成した。また、護岸データと MVTec AD を併用する実験では、全異常タイプで一貫した改善は見られなかったが、異常タイプレベルでは性能が向上する場合があることを示した。

<キーワード>ゼロショット異常検知, 外観検査, 河川護岸

## 1. はじめに

河川護岸のような社会インフラの維持管理において、AI 技術を用いた高度化・効率化が期待されている。中でも異常検知は点検作業の自動化・省力化の鍵となる技術である。既存研究では護岸全体の画像からブロック領域を切り出し、ブロックごとに異常検知を行う枠組みが提案されている [1, 2, 3].

教師あり異常検知は高い精度が期待できる一方、データ収集やアノテーションコストの高さが実務適用に向けたボトルネックとなる。教師なし異常検知はアノテーションを必要としない利点があるが、依然として現場ごとに正常データの収集とモデルの構築が必要となる欠点がある。そこで本研究は、対象ドメインの画像をモデルの最適化に用いないゼロショット異常検知に着目する。近年では、別のドメインでファインチューニングした基盤モデルを、追加の訓練なしで対象ドメインに適用する手法が多く提案されている [4, 5, 6, 7]. しかし、工業分野のベンチマークデータセット MVTec AD [8] 全体でファインチューニングしたモデルを護岸ブロックデータに適用した場合、教師なし異常検知手法と比較して性能が大きく劣る [3]. 一方で、実務上は他の河川のデータが既に蓄積され、利用可能な場面も想定される。そのような場面では、他の河川のデータを活用したゼロショット異常検知が

有効である可能性がある。さらに、MVTec AD 内にはひび割れや欠損などの護岸に生じる異常と意味的に類似した異常タイプも含まれており、データセット全体ではなく一部を補助的に活用することで性能がさらに向上する可能性がある。

本研究では、河川護岸に対して、訓練に用いるデータの違いがゼロショット異常検知性能に与える影響を分析する。具体的には、一方の河川の護岸データで最適化したモデルを他方の河川に適用するクロスドメイン評価を行い、新規河川に追加データなしで適用した場合の性能を評価する。また、MVTec AD を併用する際に含める異常タイプを変更し、性能への影響を探索的に評価する。

## 2. ゼロショット異常検知

本研究におけるゼロショットとは、対象ドメインのデータを訓練に一切用いず、異なるドメインのデータで訓練したモデルをそのまま適用する設定を指す。近年の研究では、CLIP [9] を対象外の教師ありデータでファインチューニングする手法 [4, 5, 6] が多く提案されている。これらの手法は、入力画像の特徴量と、正常・異常を示すテキストの特徴量とのコサイン類似度に基づいて異常スコアを算出する。画像全体の特徴を用いれば画像レベルの異常スコア、中間層の局所特徴を用いればピクセルレベルの異常マップが得られる。CLIP を用いる

手法はテキストによる柔軟な表現を活用することができ一方、プロンプトの違いが性能に影響するため、訓練データの違いがどのように性能に寄与しているかを厳密に分析することが難しい。

そこで本研究では、画像エンコーダのみを用いた手法 [7] を採用する。具体的には、重みを固定した基盤モデルの画像エンコーダを用いて、特徴空間上で正常・異常クラスを代表する学習可能ベクトルを対象外のデータで最適化する。本研究では、この正常・異常ベクトルの最適化を単にベクトル最適化と呼ぶ。推論時には、入力画像の特徴量と正常・異常ベクトルとのコサイン類似度に基づいて異常スコアを算出する。これにより、プロンプトの影響を除外した上で訓練データの違いが検知性能に与える効果を分析できる。

### 3. 河川間クロスドメイン評価

ベクトル最適化において、対象河川に対して他の河川のデータを利用することの有効性を検証する。

#### 3.1. 護岸ブロックデータセット

本研究では、既存研究 [1, 2, 3] と同じ護岸ブロックデータセットを用いる。このデータセットは、河川 A の水平方向に積んだ（布積み）ブロック群と、河川 B の斜め方向に積んだ（谷積み）ブロック群から構成されている。したがって、本研究における河川の違いはブロック様式の違いと対応する。本研究で扱う異常タイプは、護岸で典型的に観察されるひび割れ・目地開き、剥離・欠損・抜け、エフロレッセンス、および溶出物である。以降、これらをそれぞれ「ひび割れ」、「剥離」、「エフロレッセンス」、「溶出物」と表記する。ラベルは二層構造で付与されており、画像レベルで正常・異常を判別するラベルと、ピクセルレベルでの異常領域のマスクが与えられている。各河川のデータの詳細な構成は表 1 に示す。

#### 3.2. 実験設定

ベクトル最適化では、バックボーンには SigLIP-so400m-patch14-384<sup>1</sup> の画像エンコーダを用いる。既存の教師なし異常検知手法と性能を比較するため、ベースラインとして PatchCore [10] を採用し、

<sup>1</sup><https://huggingface.co/google/siglip-so400m-patch14-384>

表 1 護岸ブロックデータセットの画像枚数。異常画像については異常タイプ別の内訳を示す。

		河川 A	河川 B
訓練	正常画像枚数	2031	2608
	正常画像枚数	69	129
	異常画像枚数	48	36
テスト	ひび割れ・目地開き	12	12
	剥離・欠損・抜け	12	12
	エフロレッセンス	12	12
	溶出物	12	-

バックボーンには WideResNet-101 [11] を用いる。

最適化および評価には、河川護岸ブロックデータセットを使用する。ベクトル最適化には対象でない河川のテストデータのみを用いる。（例：河川 B で性能評価する場合は河川 A のテストデータで最適化）。最終的な性能は対象河川のテストデータのみに基づいて算出する。PatchCore では、訓練データで学習し、対象河川のテストデータで評価する。

実験では、以下の 3 つのケースを設定する。

**ケース 1-1.** PatchCore を用いて、対象河川の訓練用正常画像から学習して評価する。ただし、PatchCore は教師なし異常検知を前提としており、対象ドメインの正常画像を必要とするため、ゼロショット設定とは前提条件が異なる点に注意する。

**ケース 1-2.** ゼロショット設定における PatchCore の性能を評価する。一方の河川の護岸データの訓練用正常画像で学習し、他方の河川に適用するクロスドメイン評価を行う。

**ケース 1-3.** ベクトル最適化によるゼロショット異常検知手法を用いる。一方の河川の護岸データでベクトル最適化を行い、他方の河川に適用するクロスドメイン評価を行う。

#### 3.3. 評価指標

画像レベルとピクセルレベルの双方で性能評価を行う。画像レベルには AUROC (Area Under the Receiver Operating Characteristic)、ピクセルレベルには AUPRO (Area Under the Per-Region Overlap) を用いる。

**Image-AUROC:** 各画像に対して異常スコアを算

表 2 画像レベルの AUROC [%] とピクセルレベルの AUPRO [%] による性能の比較. ゼロショット設定では, 河川 B の画像で最適化したモデルを河川 A で評価する (逆も同様). 最高性能を太字で示す.

	河川 A		河川 B	
	Image-AUROC	Pixel-AUPRO	Image-AUROC	Pixel-AUPRO
ケース 1-1	94.7	64.5	97.8	74.9
ケース 1-2	86.4	56.0	82.6	61.2
ケース 1-3	<b>96.1</b>	<b>68.8</b>	<b>98.0</b>	<b>77.6</b>

出し, 縦軸に真陽性率 (True Positive Rate; TPR), 横軸に偽陽性率 (False Positive Rate; FPR) をとり, 閾値を連続的に変化させたときの曲線 (ROC 曲線) の下の面積である. 閾値に依存しない指標で, 「異常画像を画像単位で取りこぼさないか」の能力を表す. ランダムな予測に対しては 50%を示し, 100%に近いほど正確な識別ができることを示す.

**Pixel-AUPRO**: 予測されたスコアマップと真の異常領域のマスクとの領域単位での一致度を評価する指標である. ある閾値において, 各異常領域ごとに, 予測スコアマップに対して閾値を用いて 2 値化し, 領域内の再現率を算出する. すべての異常領域の再現率の平均を PRO (Per-Region Overlap) とする. この時, 閾値を連続的に変化させたときの曲線 (PRO 曲線) の下の面積が AUPRO と定義される. AUPRO は, 領域単位で異常をどの程度網羅的に検出できているかを示し, 小さな異常や部分的検出の取りこぼしを考慮した検出性能を評価する. 完全に正しいマスクが得られる場合は 100%, ほとんど覆えない場合は 0%に近づく. 本研究では, [8] に従い, FPR が 30%以下の範囲での AUPRO を報告する.

### 3.4. 実験結果と考察

河川間でのクロスドメイン評価の結果を表 2 に示す. PatchCore は対象河川の正常画像が利用できる設定 (ケース 1-1) では高い性能を示す一方で, ゼロショット設定を想定して他の河川の正常画像で学習した場合 (ケース 1-2), Image-AUROC で最大 15.2%, Pixel-AUPRO で最大 13.7%低下する. これに対し, 他方の河川の護岸データでのベクトル最適化 (ケース 1-3) は, 対象河川の正常画像を利用できる PatchCore と比較しても, Image-AUROC で最大 1.4%, Pixel-AUPRO で最大 4.3%上回る性能を達成している. また, ゼロショット設定で統

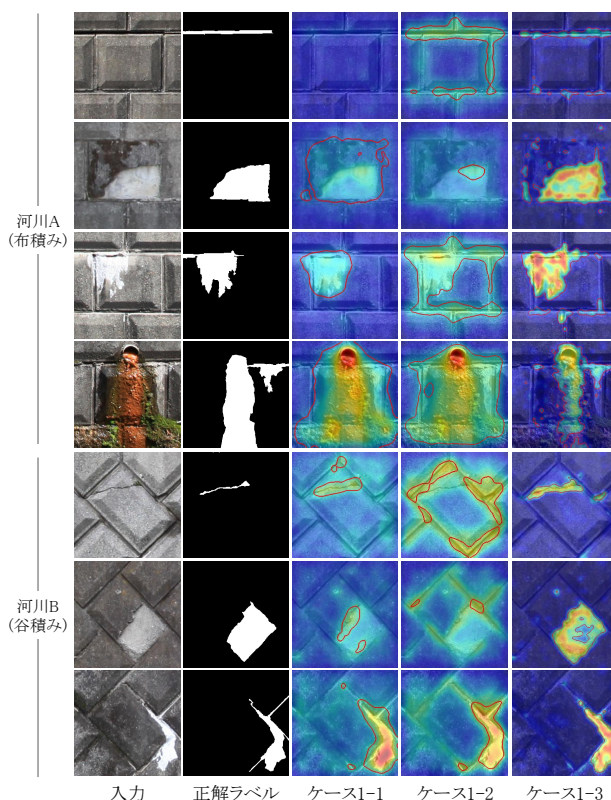


図 1 護岸ブロックデータにおける異常マップの可視化. 赤に近い領域ほど異常スコアを高く割り当てていることを示す. また, 赤線で適合率と再現率の調和平均 (F1 スコア) が最大となる時の閾値による正常と異常の境界を示す.

一して比較した場合には, Image-AUROC で最大 15.4%, Pixel-AUPRO で最大 16.4%上回る性能を示している.

図 1 に異常マップの可視化例を示す. 他の河川のデータを用いたベクトル最適化では PatchCore よりも正確に異常領域を特定できており, 定量評価と整合する結果となった.

これらの結果は, 他の河川のデータでベクトル最適化することで, 対象河川のデータを利用できない条件でも高い異常検知性能を達成することを示している. すなわち, 代表的な河川護岸データ

表 3 異常タイプ別 Image-AUROC. ケース 2-3 については, 性能改善が期待される異常タイプでの評価のみ黒字で, 対象としていない他の異常タイプに対しては灰字で示す. 最高性能を太字で示す. エフロレッスンは「エフロ」と略記している.

	河川 A				河川 B				
	ひび割れ	剥離	エフロ	溶出物	全体	ひび割れ	剥離	エフロ	全体
ケース 1-3 (ベースライン)	85.1	99.9	99.5	99.8	<b>96.1</b>	95.2	<b>98.8</b>	<b>100.0</b>	98.0
ケース 2-1	70.0	99.8	98.6	99.1	91.9	92.9	93.4	<b>100.0</b>	95.5
ケース 2-2	69.9	<b>100.0</b>	<b>99.8</b>	99.8	92.1	94.2	94.8	<b>100.0</b>	96.3
ケース 2-3 (ひび割れ)	<b>86.0</b>	99.9	99.4	99.8	96.3	<b>97.8</b>	98.3	100.0	98.7
ケース 2-3 (剥離)	83.3	<b>100.0</b>	99.8	100.0	95.8	96.7	98.5	100.0	98.4
ケース 2-3 (エフロ)	83.7	99.9	<b>99.8</b>	99.8	95.8	96.9	98.2	<b>100.0</b>	98.4
ケース 2-3 (溶出物)	84.7	99.9	99.5	<b>100.0</b>	96.0	—	—	—	—
ケース 2-4	83.0	99.8	99.6	<b>100.0</b>	95.6	97.4	98.0	<b>100.0</b>	<b>98.6</b>

で最適化したモデルを事前に用意することで, すべての河川護岸に対して汎用的に適用し得る可能性を示唆している.

#### 4. MVTec AD の併用

本章では, ベクトル最適化において, 護岸データに加えて MVTec AD をどのように組み合わせた場合に性能が向上するか検証する.

##### 4.1. 実験設定

ケース 1-3 をベースラインとして, 訓練データの構成のみ変化させた 4 つの実験ケースを設定する. 河川 A の護岸データをベクトル最適化に用い, 河川 B を評価対象とする設定で各実験ケースを説明する (逆の設定についても同様に行う).

**ケース 2-1.** MVTec AD 単体での汎化性を確認するため, MVTec AD のみを訓練データとする.

**ケース 2-2.** MVTec AD を単純に併用するだけで性能が改善するか検証するため, 河川 A の護岸データと MVTec AD の全カテゴリを併せて訓練データとする.

**ケース 2-3.** 護岸の各異常タイプに特化して MVTec AD の異常タイプを試行的に選抜し, 性能が改善するか検証する. まず, 河川 B の護岸データおよび MVTec AD の各異常タイプ  $t$  について, ベクトル最適化により異常タイプを代表する異常ベクトル  $\mathbf{v}_t$  を事前に獲得する. この代表ベクトルは, 各異常タイプにおける異常の方向を表すベクトルとして扱う. 河川 B の護岸の異常タイプ  $g_i$  と MVTec AD の異常タイプ  $m_j$  の類似度  $s(g_i, m_j)$  を

それぞれの代表ベクトルのコサイン類似度として定義する:

$$s(g_i, m_j) = \frac{\mathbf{v}_{g_i}^\top \mathbf{v}_{m_j}}{|\mathbf{v}_{g_i}| |\mathbf{v}_{m_j}|}. \quad (1)$$

異常タイプ  $g_i$  に対して, 類似度  $s(g_i, m_j)$  が高い上位  $k$  種類 (本実験では  $k = 4$ ) の MVTec AD の異常タイプのサブセット  $S_i$  を選抜し, 河川 A の護岸データと併せて訓練データとする. この設定により, 異常タイプ  $g_i$  ごとに個別にベクトル最適化を行い, 異常タイプ  $g_i$  で異常検知性能が改善するか検証する.

**ケース 2-4.** ケース 2-3 で得た各サブセットを統合してベクトル最適化を行っても, 異常タイプ別の性能向上が両立可能か検証する. 河川 B の護岸データを参照して得た MVTec AD のサブセットの和集合  $S_{\text{union}} = \bigcup_i S_i$  を河川 A の護岸データと併せて訓練データとする.

ケース 2-3 と 2-4 については, 対象河川の異常タイプを参照して MVTec AD の異常タイプを事後的に選抜しているため, 実運用の前提を満たさない点に注意する.

##### 4.2. 実験結果と考察

各ケースでの異常タイプごとの Image-AUROC を表 3 に示す. まず, MVTec AD の全データを追加した場合 (ケース 2-2), 全体での性能は MVTec AD のみの場合 (ケース 2-1) と同程度であり, 護岸データのみの場合 (ベースライン) と比べて性能は最大 4.0% 低下した. すなわち, データセット全体

を単に統合するだけでは、工業製品に由来する異常パターンの情報を護岸特有の異常の検知に十分活かすことは難しいことが示唆された。

これに対して、対象河川の護岸における各異常タイプに類似する MVTEC AD の異常タイプに絞って追加した場合 (ケース 2-3) は、河川 B の剥離を除いて対応する異常タイプに対しては性能が向上した。これは、MVTEC AD のような異なる分野のデータであっても、対象の異常タイプと類似度が高いデータであれば、性能向上に寄与し得ることを示している。一方で、河川 B の剥離で改善しなかった要因としては、MVTEC AD 内に河川 B の剥離に類似する異常タイプが存在しなかった、あるいは十分に意味的に類似した異常タイプが選抜されなかった可能性が考えられる。実際に、図 2 に示すように、本実験での異常タイプの選抜方法では護岸のひび割れに対して wood / hole や hazelnut / crack といったように意味的に類似する異常を選抜できている一方で、河川 B の剥離に対しては grid や carpet といった護岸ブロックと近い色調の画像が上位に選ばれる傾向が見られた。すなわち、代表ベクトル同士のコサイン類似度が、異常の意味的な近さだけでなく色調のような視覚的外観の近さも反映している可能性がある。

タイプ別に選抜された MVTEC の異常タイプを同時に用いてベクトル最適化した場合 (ケース 2-4) は、河川 A ではエフロレッセンス・溶出物、河川 B のひび割れおよび全体に対して性能向上が見られた一方で、他のタイプでは逆に性能低下が生じ、異常タイプ間で一貫した向上は見られなかった。すなわち、タイプごとに追加した場合に得られた効果を両立することは難しいことが確認された。

これらの結果は、MVTEC AD 内には護岸の異常タイプレベルでの異常検知性能を改善し得るデータが存在するが、それらを同時に活用して効果を維持することが容易ではないことを示している。異なる分野のデータを活用して性能向上を図る実用的な方法としては、異常タイプごとにベクトル最適化を行い、推論時に対象の異常タイプに応じたベクトルを参照するといった工夫が必要であると考えられる。

## 5. まとめ

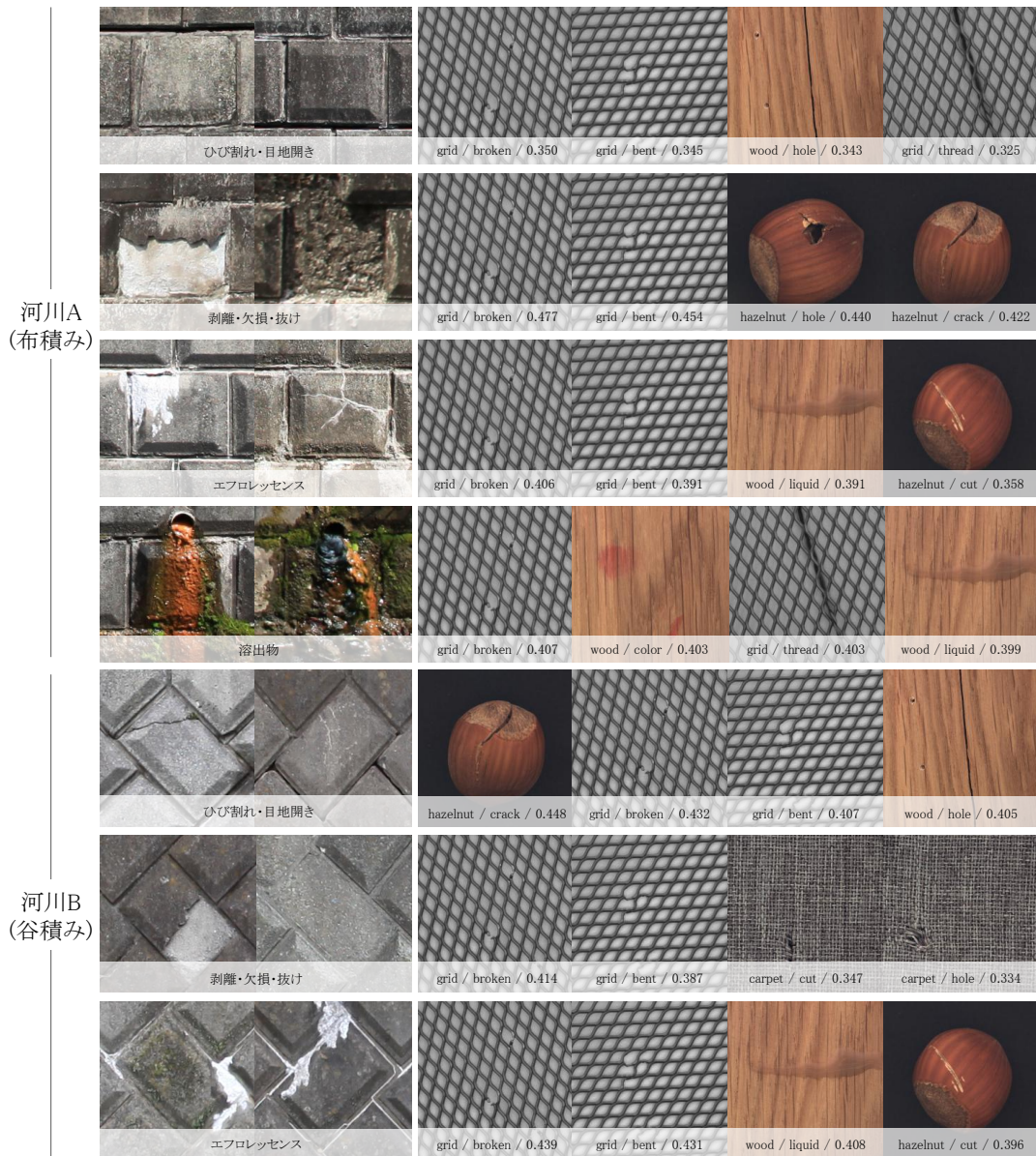
本研究では、河川護岸ブロックを対象に、訓練データの違いがゼロショット異常検知性能に与える影響を検証した。ベクトル最適化手法を用いたクロスドメイン評価により、他の河川のデータで最適化することで、対象河川のデータを用いずに教師なし異常検知手法 PatchCore を上回る性能を達成した。また、MVTEC AD を併用する実験では、一貫して性能を向上させることはできなかった一方で、異常タイプレベルでは性能が向上するような MVTEC AD のデータが存在することを示した。今後の課題としては、より多様な河川での検証や、異なる分野のデータを用いて一貫して性能を改善できるような手法の検討が挙げられる。

## 謝辞

本研究を進めるにあたり、中部大学の山下隆義先生に貴重なご助言を賜りました。ここに深く感謝の意を表します。

## 参考文献

- [1] 都築幸乃, 吉田龍人, 大久保順一, 藤井純一郎, 山下隆義. VAE を用いたブロック単位での護岸の異常検知. 土木学会 第 25 回応用力学シンポジウム, 2022.
- [2] 都築幸乃, 吉田龍人, 大久保順一, 藤井純一郎. VAE を用いた河川護岸の異常検知における連続潜在変数モデルと離散潜在変数モデルの比較. 人工知能学会全国大会, 2023.
- [3] 藤井純一郎, 都築幸乃, 七海友康, 大久保順一. 撮像環境が統一できない異常検知タスクに関する一考察. 人工知能学会全国大会, 2025.
- [4] Xuhai Chen, Yue Han, and Jiangning Zhang. A zero-/few-shot anomaly classification and segmentation method for cvpr 2023 vand workshop challenge tracks 1&2: 1st place on zero-shot ad and 4th place on few-shot ad. In *CVPRW*, 2023.
- [5] Qihang Zhou, Guansong Pang, Yu Tian, Shibo He, and Jiming Chen. AnomalyCLIP: Object-agnostic prompt learning for zero-shot anomaly detection. In *ICLR*, 2024.
- [6] Yunkang Cao, Jiangning Zhang, Luca Frittoli, Yuqi Cheng, Weiming Shen, and Giacomo Boracchi. Ada-CLIP: Adapting CLIP with hybrid learnable prompts for zero-shot anomaly detection. In *ECCV*, 2024.
- [7] 七海友康, 都築幸乃, 大久保順一, 藤井純一郎, 山下隆義. ゼロショット異常検知のための学習可能ベクトルと階層構造への適応によるファインチューニング. 第 28 回画像の認識・理解シンポジウム, 2025.
- [8] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. MVTEC AD – a comprehensive real-world dataset for unsupervised anomaly detection. In *CVPR*, 2019.



(a) 護岸ブロックの異常タイプ (b) 各異常タイプに対応して選抜したMVTec ADの異常タイプ(上位4種)

図 2 各河川護岸の異常タイプと、それに対応して選抜された MVTec AD のクラス/異常タイプ/コサイン類似度(上位 4 種類)。

- [9] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastri, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. *ICML*, 2021.
- [10] Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *CVPR*, 2022.
- [11] Sergey Zagoruyko and Nikos Komodakis. Wide residual networks. In *BMVC*, 2016.

七海友康：八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 研究員。深層学習，特に基盤モデルを用いた画像異常検知とその応用の研究に従事。

都築幸乃：八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 研究員。インフラを対象とした画像異常検知の研究に従事。

大久保順一：八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 上席研究員。自然言語処理，画像処理の研究に従事。

藤井純一郎：八千代エンジニアリング株式会社 技術創発研究所 AI 解析研究室 室長。建設分野への AI 適用の研究に従事。