

## 実運用を目指した二時期の画像検索の検証

石井明† 内山英昭‡

†八千代エンジニアリング株式会社 ‡奈良先端科学技術大学院大学

E-mail: akri-ishii@yachiyo-eng.co.jp

### 1 はじめに

#### 1.1 研究背景

我が国では、高度経済成長期に集中整備された大量の社会資本施設が建設後 50 年以上経過し老朽化が進んでいる[1]。社会資本施設の安全性と機能を今後も維持・継続させていくためには点検結果に基づいて適切にメンテナンスや更新を実施していく必要がある。しかしながら、点検は技術者による目視により実施されており、人口減少等による社会構造変化も想定される中において、建設分野ではデジタル技術を活用したインフラ DX の様々な取り組みにより、現場の生産性向上を目指している[2]。

#### 1.2 研究目的

ダムでは日々の日常点検の他、3年毎の定期検査、30年毎の総合点検等を実施している。点検項目は多岐にわたり、また点検時期によっても内容も変わるが、前述の背景のもと筆者らは定期点検(個別ダムのみ)や総合点検で実施するダムの堤体の効率的な点検・診断の実現を目指して、UAVの自律航法による空撮[3]と深層学習AIを活用した画像認識[4][5]による劣化調査手法について研究を進めてきた。しかしながら、この定期的な堤体点検作業は点検技術者の作業代替であり、管理者の直接の作業効率化やDXにはつながっていない。また、堤体を定期的に等距離正対で空撮した高解像度画像データも一時利用に留まっており有用に活用されていない問題もある。管理者が行う日常点検や緊急時の点検にて、いち早く予兆や劣化を察知・評価して対策を検討できることが必要と考えられる。

そこで本稿では、管理者が日々現場で使える提案している想定ツールのフレームワークを示し、そのフレームワークを実現するために必要となる画像検索技術について、実際のダム堤体の画像を対象としてモデルパフォーマンスの検証実験を実施し、知見の整理を行ったものである。

### 2 実運用上の問題点を踏まえた検証事項の整理(問題設定)

管理者が使用する想定ツールのフレームワークを図-1に示す。定期点検ではダム堤体全体に対してUAVに搭載したデジタルカメラを活用して堤体に近接した空撮を行い、例えば2mm/pixelのような高解像度のオルソ画像を作成して画像解析により劣化変化を抽出したり、二時期比較を実施する。一方、日常点検や地震時の緊急点検ではUAVによる自動飛行技術的、制度的に導入が難しいため、スマホ等による手軽な機器を用いた撮影を想定する。撮影画像はUAVの近接画像と比べて画像分解能が低く、画角も異なるが、スマホ等の撮影画像を入力画像とし、経年的実施されている定期点検等の高解像度の画像と比較ができ、要監視対象の事象について簡単に経過観察を可能にする仕組みである。

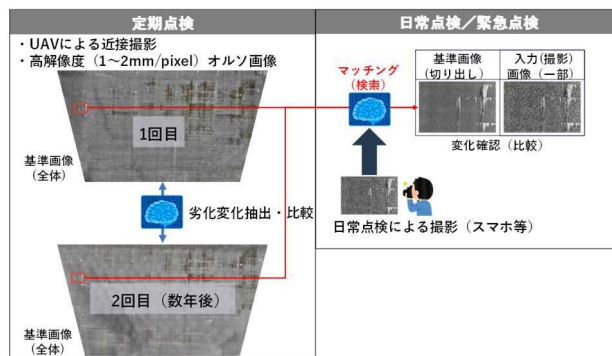


図-1 提案ツールのフレームワークのイメージ図

このフレームワークを実現するためには、1)カメラ性能・分解能、カメラ設定や撮影位置(視点)が違う条件でも画像検索ができることが必要である。また、ダムは屋外空間にある巨大構造物であり、撮影は日照・気象条件にも左右されるため、2)同一箇所でも色合いが異なる画像検索ができることも必要である。さらには、画像検索後は3)異なる解像度の劣化状況を比較・診断できる技術も必要となる。

上記 1),2)は画像特徴量のマッチング技術であり、

resnet[6]など様々な既往技術が提案されている。しかしながら、これらの技術が実用上有用かは不明であり、段階的に以下①～③に示す問題設定の検証実験を実施した。また、各問題設定における画像サイズと解像度の関係を表-1 に示す。なお、用語の定義として、全体画像はダム堤体全体のオルソ画像、分割画像は基準画像を一定の範囲で分割した画像で画像のサイズを意味する。また、基準画像は入力画像によって検索される画像であり、入力画像は基準画像とマッチングを検索する画像を意味している。

#### 【問題設定】

- 問題設定①:異なる二時期の画像マッチング問題(同じ解像度の分割画像同士のマッチング)
- 問題設定②:基準画像と分割画像の画像マッチング問題(画角が異なる画像のマッチング)
- 問題設定③:基準画像と解像度が異なる分割画像のマッチング問題

表-1 各問題設定における画像サイズ・解像度

問題設定	基準画像	入力画像	基準画像と入力画像の解像度
①	分割画像	分割画像	同一
②	全体画像	分割画像	同一
③	全体画像	分割画像	異なる

## 3 検証実験と考察

### 3.1 使用データ

実験に使用するダムの堤体画像は、東北地方にあるダムで令和元年8月(以下、「R1画像」という。)、令和4年9月(以下、「R4画像」という。)にそれぞれUAVで空撮し作成した二時期のオルソ画像(解像度2mm/pixel)を使用した(図-2)。撮影時の天候や日照条件の違いにより色合いの違いを確認できる。なお、本稿では、画像解像度を変更したデータを準備するため、基準画像はこのオルソ画像を6,000 pixel×4,000 pixel(12m×8m)に一旦分割し、その分割画像を調整してから、連結して全体の基準画像を再構成している。

### 3.2 問題設定①に対する実験

問題設定①の実験は、異なる二時期の画像(R1画像とR4画像)の分割画像同士が検索可能であるかの検証である。二時期の分割画像の解像度は同じ

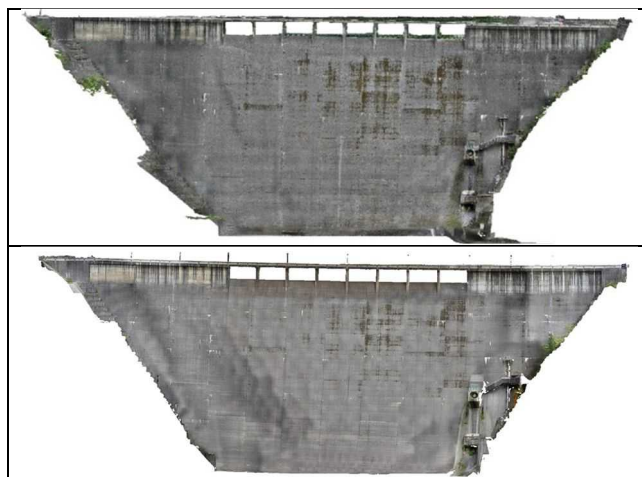


図-2 堤体画像全体(上段:R1画像、下段:R4画像)

とし、分割画像の解像度は元々6,000 pixel×4,000 pixelの画角であるが、ここではLarge(3,000 pixel×2,000 pixel)、Middle(1,500 pixel×1,000 pixel)、Small(900 pixel×600 pixel)、Tiny(600 pixel×400 pixel)の4種類のサイズに変更して実施する。

検証方法は、R1画像の各分割画像に対してR4画像の分割画像を入力し、類似度に基づいてランキングを作成する。R1画像とR4画像を入れ替えたバージョンも実施する。具体には、まず各分割画像をpre-trained modelを用いて特徴ベクトルを抽出する。なお、ここで使用するpre-trained modelは、resnet50(2,048次元)、resnet101(2,048次元)、densenet201(1,920次元)[7]、efficientnet\_v2\_1(1,280次元)[8]、ConvNeXt Large(1,536次元)[9]の5つとした。次に、類似度は全てのサイズの分割画像に対して距離を算出(brute force matching)しランキングする。評価はRecall@k[10]で各入力画像に対して上位k個の予測に含まれる正解数が総正解数のうちの程度の割合が含まれているか計算し、同じファイル名が検索された場合(ファイル名が一致した場合)に正解と判定した。

上位5個、10個、20個、40個を対象として整理したRecall@kを表-2に示す。また、図-3には上位5個、10個、20個、40個毎に、各モデルの画像サイズRecall@kを整理した。なお、DirectionのR1toR4はR1画像をR4画像に入力していることを示す(R4toR1はその逆を示す)。また、入力画像検索の一例を図-4に示す。efficientnet\_v2\_1を除くすべてのpre-trained modelでは上位5個だけでは7割弱でしか検索できないが、対象画像数を多くするほどほぼ確実に対象

表-2 Recall@k 一覧表

Model	Size	R1toR4				R4toR1			
		Top5	Top10	Top20	Top40	Top5	Top10	Top20	Top40
resnet50	Tiny	61.8	73.4	89.6	94.8	57.6	70.0	81.8	92.9
	Small	68.2	81.5	91.9	97.1	61.8	74.7	88.8	95.3
	Middle	63.6	76.9	91.9	96.0	51.2	70.0	85.3	95.3
	Large	48.6	67.1	88.4	93.6	40.0	51.2	75.9	92.4
resnet101	Tiny	60.1	73.4	85.0	93.1	58.2	67.6	84.7	92.9
	Small	68.2	80.9	91.3	96.5	62.4	78.8	92.9	98.2
	Middle	56.6	76.9	87.9	96.5	54.7	72.4	88.2	97.1
	Large	49.7	65.9	83.8	94.2	45.3	59.4	76.5	88.8
efficientnet_v2_l	Tiny	35.8	46.8	57.8	66.5	32.9	46.5	62.4	68.8
	Small	37.6	47.4	58.4	63.6	28.8	36.5	51.2	65.9
	Middle	32.9	44.5	55.5	60.7	38.2	52.4	64.1	74.7
	Large	34.1	45.7	54.9	63.0	37.6	54.7	64.7	70.0
densenet201	Tiny	63.0	72.8	83.8	89.6	63.5	75.3	86.5	88.8
	Small	67.1	79.8	90.2	95.4	61.8	74.7	87.6	92.9
	Middle	53.8	69.9	89.6	96.0	52.9	67.6	80.6	93.5
	Large	39.3	54.9	79.2	92.5	38.8	51.2	72.9	87.1
convnext_large	Tiny	59.0	72.3	84.4	90.2	52.9	66.5	80.0	89.4
	Small	68.8	77.5	94.2	94.2	61.2	70.6	87.6	94.1
	Middle	63.6	76.3	90.8	96.0	58.8	70.6	85.9	95.3
	Large	45.7	64.2	83.8	91.3	40.0	56.5	76.5	91.8

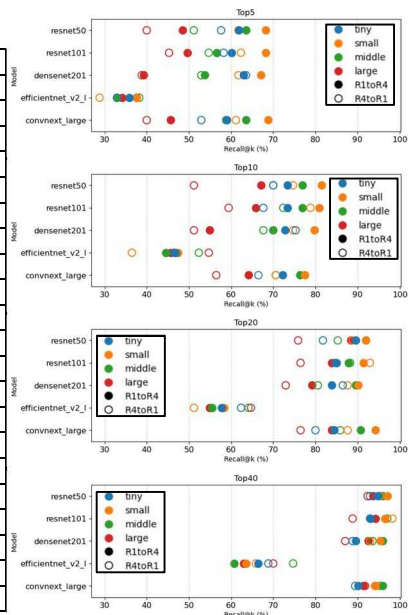


図-3 Recall@k 分布図

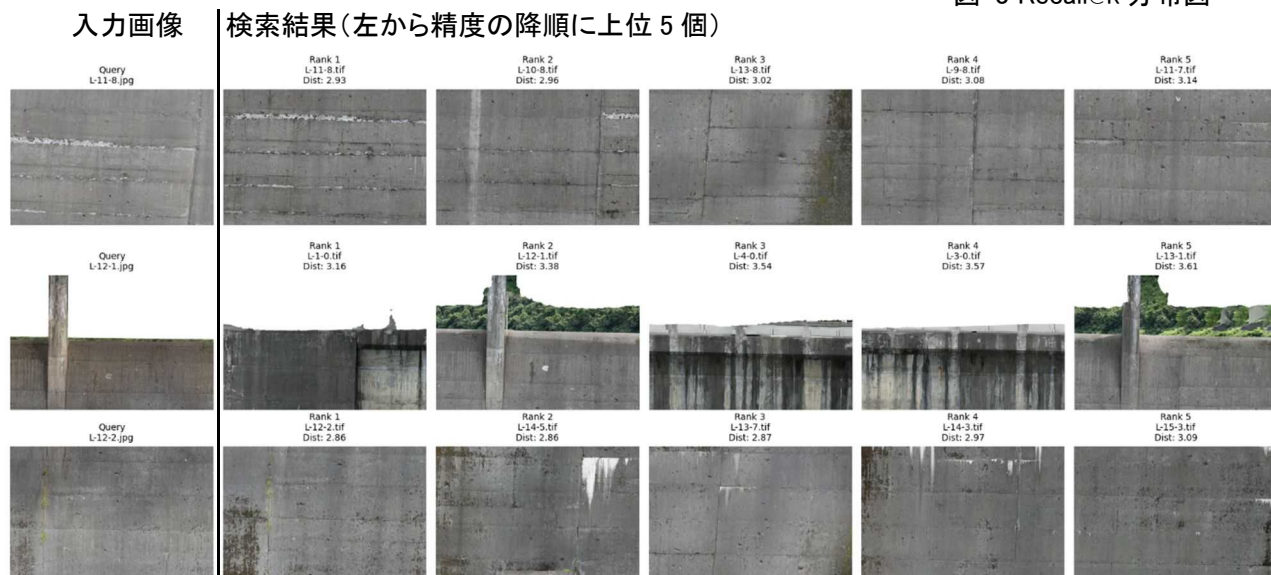


図-4 問題設定①の画像検索結果の一例

画像範囲が含まれる結果となった。また、画像分割サイズは小さい方が検索される割合が高く、Small や Tiny の画像サイズで十分に検索できている。ダム堤体の定期点検における劣化抽出では、検知する劣化サイズの縛りもあるため 2mm/pixel のような高解像度の画像を必要としているが、画像検索ではそこまで高い解像度は必要なく十分に堤体表面のテクスチャの特徴を捉えて画像検索できることが分かり、新たに画像検索用にデータ取得する必要もないと考えられる。なお、本稿の以降の問題設定②および問題設定③の検討においては、運用時の計算負荷を軽減するためにも画像サイズが小さいことが好ましいため、

Resnet50, 分割画像サイズ Tiny (600×400) の組合せにて検討を行うことにする。

### 3.3 問題設定②に対する実験

問題設定②の実験は、入力画像の分割画像について、基準画像の位置を探索可能かの検証であり、基準画像と分割画像は同一解像度とし、画角が異なる画像を検索する。基準画像は全体のサイズが 120,000 pixel×48,000 pixel であるが、問題設定①の実験結果を踏まえて基準画像のサイズを 1/10 の 12,000 pixel×4,800 pixel にリサイズし、入力画像はそのリサイズした全体画像から後述する窓画像枠を

使って作成した分割画像(600 pixel×400 pixel)の中からランダムに20枚抽出を行った。

検証方法は、以下のアルゴリズムで実施した。

#### 【検索アルゴリズム】

- 1) 分割画像の横と縦それぞれ半分の幅(横:300 pixel, 縦:200 pixel)をスライド幅とし、基準画像全体を左上から右下に向かって分割画像のサイズの窓画像枠の分割画像を作成する(図-5)。

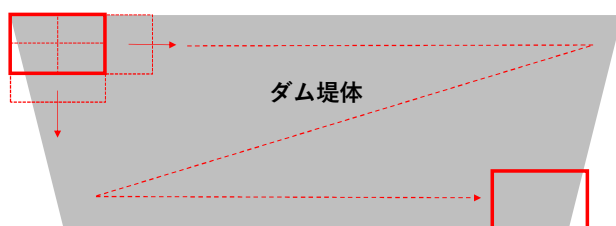


図-5 窓画像枠の作成のイメージ

- 2) 問題設定①の決定モデルにて画像検索を行い、窓画像で分割された画像と入力画像の特徴量を比較し、類似度上位N枚の画像群を検索する。なお、今回はN=10で実験を実施した。
- 3) 上位10位の画像についてLoFTR[11]で特徴点マッチングを行い、特徴点のマッチング数(Inliers)が多い順に並び替えを実施する。
- 4) 入力画像と分割された画像の2枚の画像間でRANSACを用いたホモグラフィにより基準画像上での入力画像の位置を算出し、分割画像の位置

は既知であることを利用し、画像の左上座標の誤差を算出する。

問題設定②における画像検索結果の一例を図-7に示す。類似度(dict)の順位と特徴点マッチング数(Inliers)の順位は必ずしも一致はしておらず、入力画像に対してより確度の高い検索画像に並び替えられている。これは使用した画像検索モデルに依存するが、どのモデルでも必ず1番目に正解画像が検索されることはなく、上位の対象枚数を絞れば絞るほど正解画像が検索される割合は低くなるため、特徴点マッチングによる並び替えは有用であると言える。また、図-7の最上図の画像検索結果について、入力画像の全体画像における位置検索結果およびLoFTRを使った特徴点マッチング状況の一例を図-8に、20枚の全ての入力画像について、基準画像における位置を推定した際の誤差を図-6にそれぞれ整理した。この検証では、1pixel以下で推定することができており入力画像の基準画像全体における位置を正しく推定できていることが確認できた。

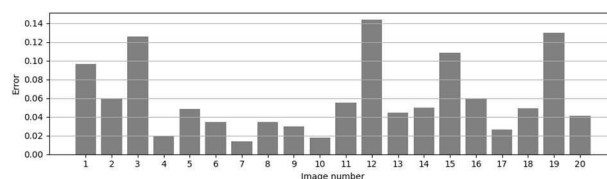


図-6 問題設定②の推定位置誤差

#### 入力画像 窓画像による分割画像の検索結果(左から特徴点マッチング点数の降順に上位10個)



図-7 問題設定②の画像検索結果の一例



図-8 問題設定②の位置検索結果の一例(左:入力画像, 中:入力画像の位置検索, 右:特徴点マッチング)

### 3.4 問題設定③に対する実験

問題設定③の実験は、基準画像とする全体画像と入力画像の分割画像の解像度が異なる場合において、分割画像の入力画像の位置について基準画像の位置を検索可能か検証する。基準画像は問題設定②と同様に 12,000 pixel×4,800 pixel とし、入力する分割画像は 6,000 pixel×2,400 pixel, 10,000 pixel×4,000 pixel, 20,000 pixel×8,000 pixel, 24,000 pixel×9,600 pixel の 4 種類のサイズの全体画像から問題設定②と同様に窓画像枠を使って分割画像 (600 pixel×400 pixel) をランダムに 20 枚抽出を行った。なお、検証および評価方法は問題設定②と同じである。

表-3 問題設定③の推定位置誤差

検索画像番号	入力画像のサイズ			
	6,000 × 2,400	10,000 × 4,000	20,000 × 8,000	24,000 × 9,600
1	3.846	0.940	0.068	0.284
2	13.588	0.558	0.369	5877.833
3	5.525	0.751	0.397	0.828
4	27.532	0.536	0.200	0.085
5	14.226	0.760	0.105	0.964
6	20.158	0.625	0.037	3244.828
7	4.510	0.678	0.091	0.143
8	7.675	0.754	0.201	2134.623
9	9.321	0.713	0.071	0.031
10	13.282	0.756	0.135	1693.667
11	13.490	0.931	0.262	0.630
12	6.230	0.799	2120.995	0.259
13	7.431	1.253	0.042	0.051
14	2668.326	1.232	0.445	0.278
15	9.666	1.167	0.094	42.244
16	10.282	0.900	0.163	0.222
17	10.373	1.139	1334.160	0.129
18	10.314	1.296	1657.230	0.250
19	11.359	1.274	0.134	0.210
20	12.741	1.295	0.166	0.079

推定位置誤差を表-3 に示す。表-3 中のハッチ箇所は推定位置誤差 10pixel 以上を示す。推定位置誤差が 1pixel 以下は、入力画像サイズが 6,000 pixel×2,400 pixel の場合は 8 枚, 10,000 pixel×4,000 pixel の場合は 13 枚, 20,000 pixel×8,000 pixel の場合は 17 枚, 24,000 pixel×9,600 pixel の場合は 15 枚となっており、入力画像の解像度が基準画像の解像度の同程度から 2 倍程度であれば、正しく位置を推定できることが示唆される結果となった。また、画像検索結果および位置検索結果の一例として、表-3 の検索画像番号 9 の結果を入力画像サイズ毎に図-9～図-12 に示す。なお、図-9～図-12 は上段に位置検索結果、下段左に入力画像、下段中央に入力画像の基準画像における位置検索結果、下段右に入力画像と基準画像の特徴点マッチングを示す。入力画像の画像サイズが小さいほど、より広範囲の情報を含む画像になっていることから、それに応じた入力画像の基準画像における位置検索範囲も推定できていることが確認できた。

## 4 検証実験と考察

本稿では、ダム管理者の DX を実現するフレームワークを実運用する際に必要な技術について、問題設定を行いモデルのパフォーマンス実験を実施した。カメラ性能や分解能、カメラ設定、撮影位置(視点)が違う条件での画像検索や、日照・気象条件によって色合いが変わる画像検索において、テクスチャーの特徴量をロストしない基準画像の解像度が求められること、また入力画像に対して画像検索の精度が高めるための基準画像の解像度設定が重要であること

入力画像 | 窓画像による分割画像の検索結果 (左から特徴点マッチング点数の降順に上位 10 個)

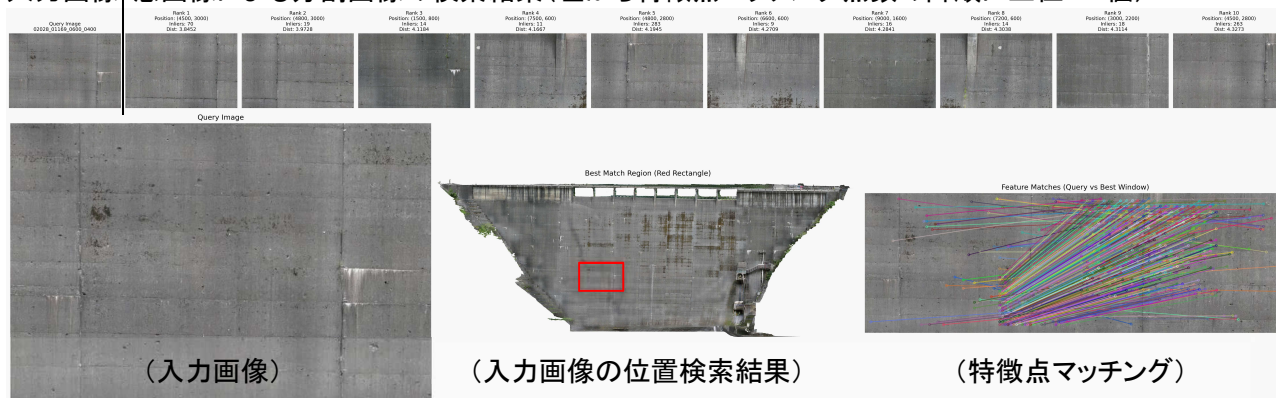


図-9 問題設定③の画像検索結果および位置検索結果の一例(入力画像サイズ 6,000×2,400)

入力画像 | 窓画像による分割画像の検索結果(左から特徴点マッチング点数の降順に上位 10 個)

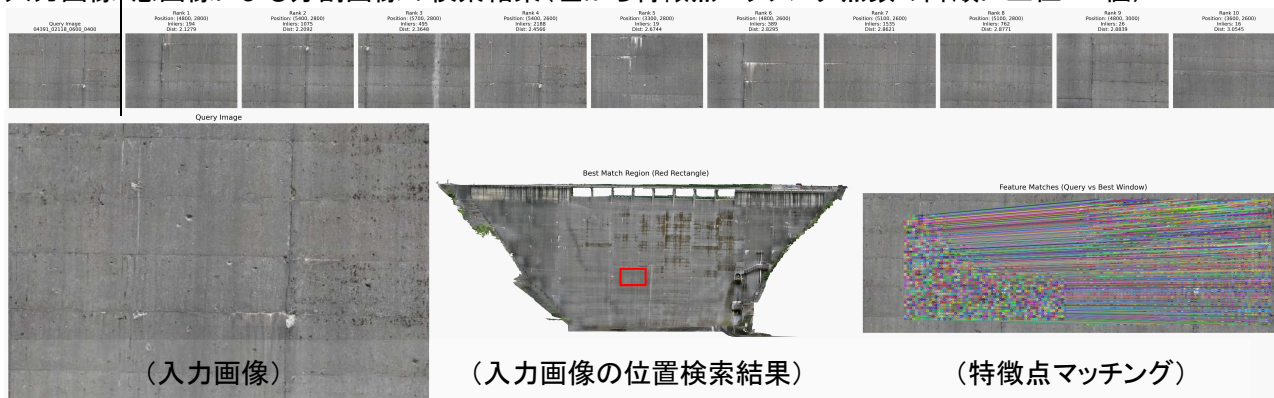


図-10 問題設定③の画像検索結果および位置検索結果の一例(入力画像サイズ 10,000 × 4,000)

入力画像 | 窓画像による分割画像の検索結果(左から特徴点マッチング点数の降順に上位 10 個)

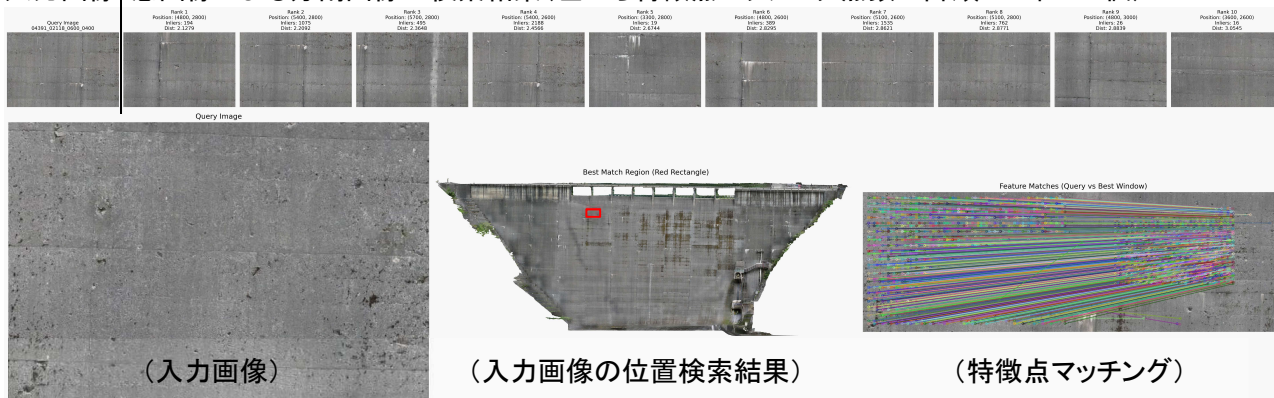


図-11 問題設定③の画像検索結果および位置検索結果の一例(入力画像サイズ 20,000 × 8,000)

入力画像 | 窓画像による分割画像の検索結果(左から特徴点マッチング点数の降順に上位 10 個)

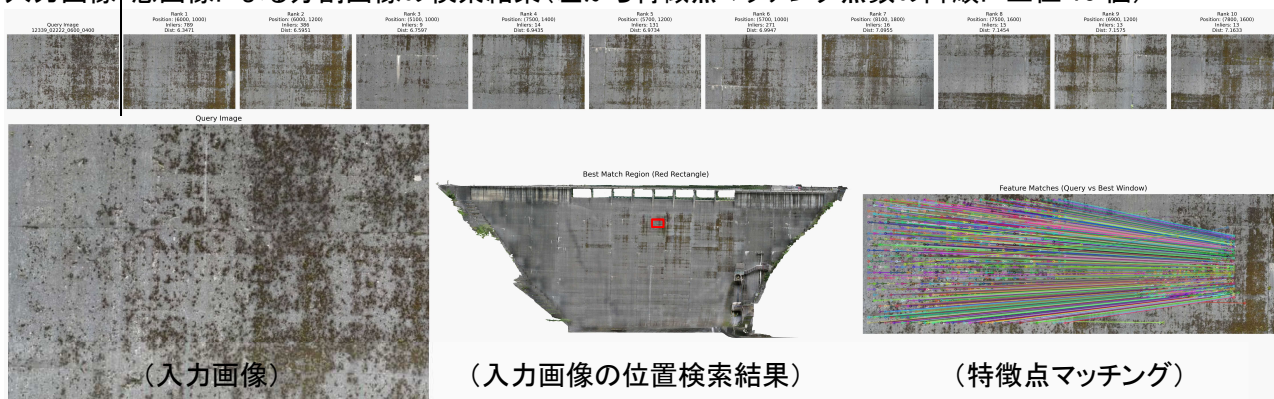


図-12 問題設定③の画像検索結果および位置検索結果の一例(入力画像サイズ 24,000 × 9,600)

が明らかにすることができた。

今回は pre-trained model を使ってモデルの精度についてパフォーマンス検証を実施した。より精度を追求していくためには fine-tuning することも考えられる。一方、実運用する現場では高性能の計算スペックが必ずしも確保されていない。そのため、推論速度と精

度のトレードオフも分析して最適なモデルを現場に投入できるように検証を進めていくことが今後の課題である。加えて、異なる解像度上での劣化状況の定量的な比較・診断を実現するために、まずは粗い入力画像を超解像度化して劣化可能性を表示できる方法も今後検討していく必要がある。

## 参考文献

- [1] 国土交通省 社会資本の老朽化対策情報ポータルサイト [https://www.mlit.go.jp/sogoseisaku/maintenance/02research/02\\_01.html](https://www.mlit.go.jp/sogoseisaku/maintenance/02research/02_01.html)
- [2] 国土交通省:i-Construction2.0～建設現場のオートメーション化～, 2024年4月.
- [3] 石井明, 菅原宏明, 小篠耕平, 天方匡純: UAVの自律航行と空撮画像を活用したダム堤体点検の効率化・高度化に関する研究, AI・データサイエンス論文集, vol.1 J1, pp613-622, 2020.
- [4] 嶋本ゆり, 安野貴人, 栗飯原稔, 藤井純一郎, 大久保順一, 天方匡純: ディープラーニングによるポップアウトの自動検出手法の提案, 第33回人工知能学会全国大会, 4C3-J-13-05, pp.79-82, 2019
- [5] 藤井純一郎, 平嶋智希, 栗飯原稔: ロックフィルダムの堤体点検への画像認識AIの適用, 人工知能学会全国大会論文集, JSAI2022巻, p. 4Yin231, 2022.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition, arXiv:1512.03385v1, 2015.
- [7] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger: Densely Connected Convolutional Networks, arXiv:1608.06993v5, 2018.
- [8] Mingxing Tan, Quoc V. Le: EfficientNetV2: Smaller Models and Faster Training, arXiv:2104.00298v3, 2021.
- [9] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie: A ConvNet for the 2020s, arXiv:2201.03545v2, 2022.
- [10] Recall@k <https://zenn.dev/hellorusk/articles/7e336fd3c6be20a8f8d1>
- [11] Jiaming Sun, Zehong Shen, Yu'ang Wang, Hujun Bao, Xiaowei Zhou: LoFTR: Detector-Free Local Feature Matching with Transformers, arXiv:2104.00680v1, 2021.