画像特性の変化に頑健な
護岸のひび割れ検出モデル作成手法の検討

吉田　龍人[1]・藤井　純一郎[1]・大久保　順一[2]・天方　匡純[1]

[1]正会員　八千代エンジニヤリング株式会社（〒111-8648 東京都台東区浅草橋 5-20-8）
E-mail: ry-yoshida@yachiyo-eng.co.jp (Corresponding Author)

[2]非会員　八千代エンジニヤリング株式会社（〒111-8648 東京都台東区浅草橋 5-20-8）

道路橋定期点検要領の改訂など，AI を活用したインフラ維持管理の運用に向けた気運が高まっている．一方で，現場で撮影される画像は天候などの影響によって特性が大きく変化するため，多様なシーンで撮影された画像から AI が安定して変状を検出できないといった課題がある．

これを受けて本稿では河川護岸のひび割れ Segmentation をテーマに，河川の左右岸別に画像の特徴量を比較することで，異なる状況で撮影された画像特性の違いを定量評価した．さらに画像の見かけの違いに頑健になると言われている Instance Normlization をひび割れ検出モデルに導入し，特性の異なる画像間で安定してひび割れを検出できるモデル構造を検討した．

Key Words:*Deep Learnig, Maintenance, Segmentation, Instance Normalization*

1.　はじめに

令和 2 年版国土交通白書[1]においてインフラの老朽化や技術者不足が提言されている．同書ではインフラ老朽化が進行する一方で，維持管理を行う建設就業者数は減少し，2023 年には 3 万人もの人員が減少すると予測されている．こうした現状は既に実体化しており，富山市では点検・補修を行う対象を重要な橋梁のみに絞る橋梁トリアージ[2]と称する取り組みが実施されるなど，既設のインフラを維持しきれない時代が近づいている．

この問題を改善するために，AI を活用したインフラ維持管理の効率化が検討されている．2019 年に策定された道路橋定期点検要領[3]では，橋梁の健全性の診断の根拠となる状態の把握は近接目視により行うことを基本としつつも，「自らの近接目視によるときと同等の健全性の診断を行うことができる情報が得られると判断した方法」である場合，目視以外の手法を用いても良いという記述が新たに示された．この記述は点検における機械の活用を促すものであり，AI を活用した維持管理（以下，AI 維持管理とする）の一般化に向けた気運の高まりを表している．同様に，他の構造物においても AI 維持管理が次第に容認されることが予測されるため，点検要領等のルール改訂に先立って技術の研究・開発に取り組む必要がある．

河川の護岸は現行の点検要領において目視点検を必須としている．2019 年に策定された点検・評価要領[4]では，点検時に発見した変状はサイズの計測と画像の撮影を実施することが定められており，長距離移動を伴う非効率な点検が実施されている．点検に AI を活用できれば，従来点検で必須とした計測と撮影の工程のうち，計測が機械化できるため，点検の効率化が実現できる．さらにドローンなどを活用することで，撮影効率の向上も期待される．一方で護岸の変状検出 AI の研究事例は少ない．既往研究[5]ではひび割れ検出 AI の作成手順や評価の手法が示されたが，実用化には至っていない．AI 実用化に向けた課題には，汎化性と精度の向上がある．現場で撮影される画像は，天候などの外的要因によってデータ特性が変化するため，その変化に頑健でなければならない．

そこで本稿では，護岸のひび割れ Segmentation を題材に，画像から変状を検出する AI を作成する上での課題を整理しつつ，画像の特性の変化に頑健なモデルの作成方法について検討する．なお本研究で検討する AI 技術は，単に人が行う既存の点検プロセスを代替するものでなく，天方[6]らが示すように維持管理フローを変革し，効率化を達成することを目標としている．したがって現在の護岸の管理基準となる 2mm ひび割れの検出は追求せず，あくまで効率的な維持管理を念頭に置いたツール開発のための研究であることに留意されたい．

2. AI 維持管理に要する撮影および AI 解析技術

護岸の AI 維持管理では，護岸の変状を撮影した画像から変状を検出する AI のモデルを構築し，そのモデルによって定期的に撮影した画像から健全度を評価するといったフローが想定される．したがって，適切に画像を撮影する技術と，画像から変状を精度よく検出する技術を確立しなければならない．そこで本章では護岸の代表的な変状であるひび割れをテーマに，画像撮影とひび割れ検出モデル作成において必要となる技術や課題についてまとめる．

(1) 画像の撮影

撮影時は適度な空間解像度の設定やレンズの焦点設定，ブレ防止など最低限守るべき条件はあるが，後の AI 解析の工程で最も課題となるのが，撮影時の画像の特性を可能な限り揃えることである．撮影対象はいずれも屋外にあるため，撮影された画像の特性は天候など外的要因によって変化する．

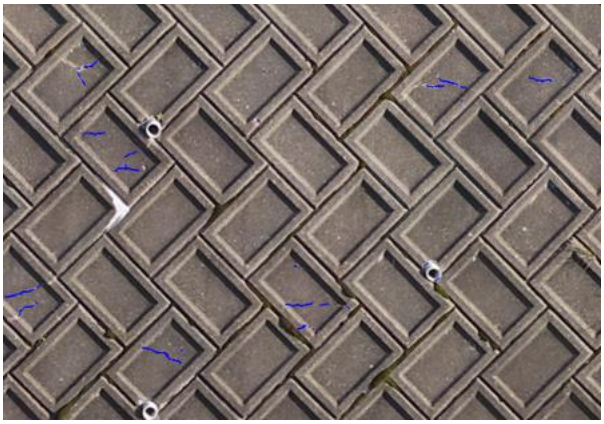
図-1 は単独測位の GPS を使って自動航行するドローンから異なる 2 時期の同一護岸を撮影した結果を示している．いずれも護岸から約 8m 離れた位置を飛行ルートとして設定し，撮影機体やカメラの種類，設定の条件等も統一した．撮影時期は**図-1** (1)が 2018 年 11 月で，**図-1** (2)が 2020 年 11 月である．2 枚の画像より護岸の色合いに違いが見られる．2018 年の撮影では連日晴れが続いていたため護岸が乾燥した状態であったが，2020 年の撮影では雨が上がった翌日であったため護岸が湿潤な状態であった．加えて撮影時の天候や撮影時間帯などの違いも相まって，画像の見かけに大きな変化が生じた．見かけ以外にも，GPS の測位誤差や現場で吹く風の影響等を受け，撮影位置に違いが生じた．400m ほどの区間を連続的に撮影しているため地点によって護岸との距離はさまざまであるが，この地点の撮影結果においては，2020 年の画像の方が護岸に近い位置で撮影された．

図-1 (1)と同日に撮影した同河川の護岸画像 120 枚で学習させたひび割れ検出モデルによって，2018 年と 2020 年に撮影した同地点の画像に対してひび割れ検出を行った結果が**図-2** である．なおモデルは斎藤[7]らの手法に準拠して作成したものを用いている．これにより 2018 年画像で検出されたひび割れが 2020 年画像ではほとんど検出されていないことが分かる．2 枚の画像の間でひび割れ形状の変化は特になく，画像の色合いだけが異なるため，2018 年の画像から 2020 年の画像の間に発生した何らかの画像特性の変化がひび割れ検出精度に影響したことが示唆される．

(1) 2018 年 10 月

(2) 2020 年 11 月

図-1 日付別の同地点空撮結果



(1) 2018 年画像のひび割れ検出結果

(2) 2020 年画像のひび割れ検出結果

図-2 撮影年が異なる同地点画像のひび割れ検出結果

(2) CNN を使った画像の AI 解析

一般に CNN のモデルは，着目対象の形状情報よりもテクスチャ情報を重視する傾向にあると言われている．例えば Goodfellow [8]らは，定性的に見て変化がない程に微弱なノイズを画像に与えるだけで，分類器が誤分類を起こす Adversarial Examples と呼ばれる現象を示した．さらに Geirhos [9]らは，テクスチャ変換を与えたある形状の画像を分類器に入力すると，形状ではなくテクスチャが属するカテゴリに分類される傾向にあるという実験結果を示した．ここで本稿では画像に写る護岸のブロックやひび割れなど，画像全体の画素値の勾配によって認識される物体自体を形状とし，個々の画素値の集合によって表現される画像の色合いや質感などを総称してテクスチャとする．

既往研究を踏まえて，同じ形状の物体を写した**図-2** の画像においてひび割れ検出精度が低下したのは，テクスチャに違いがあったためだと推察した．テクスチャによってひび割れ検出精度が低下する現象は，モデル作成に用いる教師画像にさまざまなテクスチャの画像を与えることで防止・低減できると考える．ただ，教師画像のテクスチャのバリエーションを担保するには，テクスチャの分布を評価する必要がある．そこで，テクスチャの分布を評価する手法を 3 章にて検討する．具体的にはある河川の護岸画像の特徴量を左右岸別にさまざまな観点から整理し，分布の違いを定量的に評価可能な指標を検討する．

AI のモデル構造の選定は高精度なひび割れ検出実現には欠かせない．これまでに精緻な Segmentation を実現するためのモデルに関する研究は多方で実施されてきた．U-Net [10]では Encoder-Decoder 構造のモデルに Encoder と Decoder を結ぶ skip connection が導入された．これによって浅い層の細かな特徴を持つ出力が深い層へと伝達され，精度の高いローカライズが可能となり，従来のモデルより鮮明な出力を可能とした．さらに FusionNet [11]では，U-Net で導入された skip connection に加えて，ResNet [12]で用いられた残差ブロックが short skip として導入された．ResNet では残差ブロックを用いることで，勾配消失や勾配爆発の問題が解消され，層数の多いネットワークを作成することが可能となった．この構造を Encoder-Decoder 構造の Segmentation モデルに導入することで出力がより精細となり，神経細胞の Segmentation タスクである ISBI EM Segmentation challenge のデータセットにおいて 2016 年時点での SOTA を達成した．斎藤 [7]らはこの FusionNet の構造をアレンジして，護岸ひび割れ検出用のモデルを新たに構築し，高精度なひび割れ検出を可能とした．そのモデルが**図-3** である．このモデルではベースとした FusionNet に比べてプーリング層が 1 層追加され，long skip が ResNet で用いられる残差結合から U-net で用いられるスキップ結合へと変更された．

モデル構造の変化以外にもひび割れ検出精度が向上した要因に Normalization 層の導入が挙げられる．FusionNet や斎藤らのモデルにも導入された Batch Normalization [13]（以下，BN とする）は，ミニバッチのデータを平均が 0，分散が 1 となるように正規化するもので，内部共変量シフトを防止する役割を果たしている．これにより勾配消失や勾配爆発の影響を低減することができ，モデルの層を深くすることを可能とした．Normalization は，BN 以外にも正規化を行うデータの範囲に合わせて Layer Normalization [14]や Instance Normalization [15]（以下，IN とする），Group Normalization [16]などが提案されている．ここで IN とはミニバッチ内の各データを ch ごとに平均が 0，分散が 1 となるよう正規化を行う手法である．

画像のテクスチャの違いとは，元の画像の RGB の 3ch の画素の分布，ひいては RGB の 3ch を基に畳み込みを行った中間層の ch ごとの分布の違いであると仮定すると，ch ごとに正規化を行う IN を導入することで，入力データの分布の偏りが減り，テクスチャの違いに起因するひび割れ検出精度の低下を防止することが期待される．実際に IN と BN を組み合わせたモデルによって，色合いの異なる画像における CNN の精度が向上するという研究結果が Pan [17]らによって示されている．そこで本稿では，**図-3** に示すモデルに対して，IN を取り入れたモデルを作成し，異なるテクスチャの画像間での IN モデルの推論精度を検証する．実験の概要と結果は 4 章に示す．

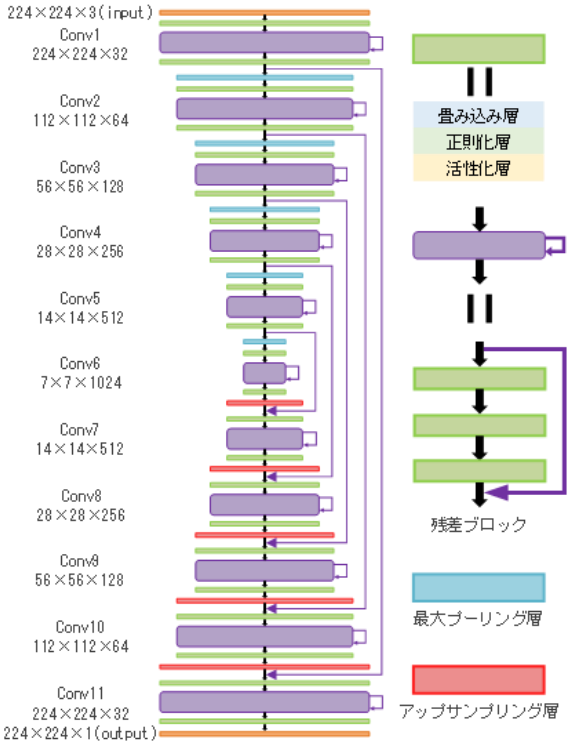


図-3 ひび割れ検出モデル [7]

3. 護岸画像特性の定量化

前章で示した通り，Segmentation によるひび割れ検出では画像のテクスチャによって精度が低下することが想定されるが，具体的にテクスチャの違いを定量評価する手法は確立されていない．そこで，本章では布積みの 1 種類の護岸を対象に，左右岸別で画像の特徴量を整理し，テクスチャの違いを評価する手法を検討する．合わせて護岸のひび割れ Segmentation の対象となる護岸ひび割れの分布特性を明らかにし，モデルの学習における課題を明らかにする．

(1) 対象画像の概要

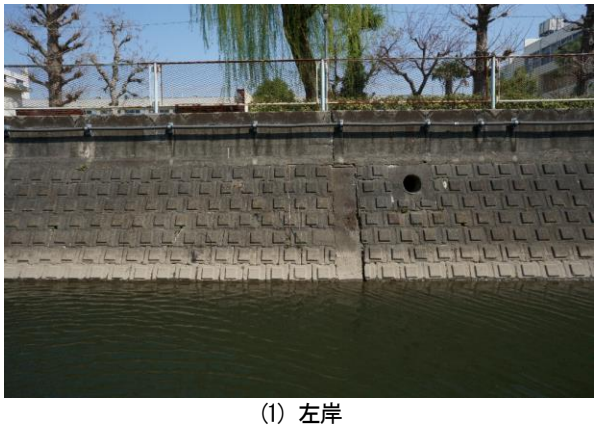
特徴量算出を行う対象画像の例を**図-4** に示す．この図は石井[18]らが visual SLAM で自律飛行するドローンによって撮影した 256 階調の RGB 画像である．画像の横幅が 6000px，高さが 4000px で，空間解像度が 2mm/px 以下を満たしつつ，概ね護岸と正対するよう撮影されている．この画像は西から東へと流れる河川の両岸を撮影したものであるため，**図-4** (1)に示す左岸画像は日が差しており，**図-4** (2)に示す右岸画像は陰になっている．これにより左岸の画像は暖色，右岸の画像は寒色らしい配色となっていることが定性的に判断され，左右岸でテクスチャの変化が生じていると考えられる．

空撮画像サイズに対して，ひび割れ検出モデルに入力するサイズは縦横 224px であることから，学習用のデータセットの作成時や推論時には画像の分割を行う．本稿では斎藤[7]らの手法に従って，Faster R-CNN[19]で護岸領域の上端と下端の y 座標を取得したのちに，上下端内をメッシュ状に区切り，縦横 224px の画像に分割する（以下，分割前の画像を元画像，分割後の画像を crop 画像とする）．個別の crop 画像の特徴量をさまざまな観点から算出し，元画像全体での分布を明らかにする．特徴量算出の対象となる画像は，右岸 46 枚，左岸 21 枚の計 67 枚の元画像を分割した 16380 枚の crop 画像である．**表-1** に 16380 枚の crop 画像の内訳を示す．

(2) 特徴量の算出結果

a) ひび割れ分布

図-5 は crop 画像のひび割れ面積の累積和を示したグラフである．なお，本稿で取り扱うひび割れは河川技術者の精査の下で画像に付与したラベルによって評価しており，ラベルには個人の判断誤差が含まれることに注意されたい．人目でも明らかに劣化が進行していると分かる**図-4** の護岸においても，ひび割れラベルが付与されている crop 画像の割合は画像全体の 25%程度と少ない．



(1) 左岸

(2) 右岸

図-4 解析対象の護岸画像

表-1 crop 画像内訳

	左右岸	画像枚数	ひび割れ ラベル px 数
全 crop 画像	R	11622	5522895
	L	4758	1283992
train&val	R	1179	1556933
	L	1179	928396
test	R	284	369544
	L	278	355596

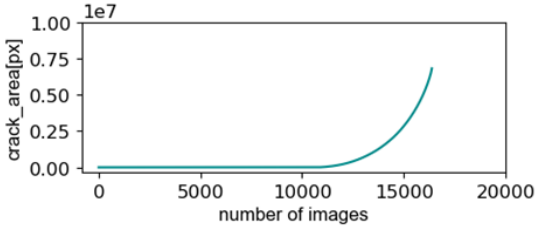


図-5 ひび割れ面積の累積和

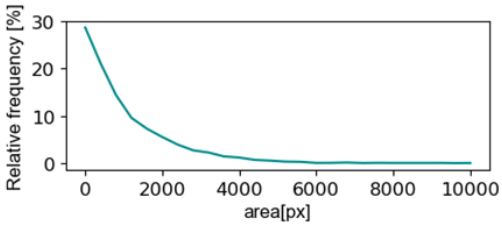


図-6 ひび割れ面積の相対度数分布

図-6 はひび割れラベルの付いた crop 画像から計測したひび割れ面積の相対度数分布を示している．crop 画像はラベルの付いた画像の割合が小さいだけでなく，画像中に占めるラベルの割合も小さく，分布の多くが 2000px 以下に集中している．crop 画像は面積が 50176px であるため 2000px は画像全体の 4%程度に相当する．

以上 2 つのグラフより，対象画像がひび割れラベルの少ない，極めて不均衡なデータであることが示された．

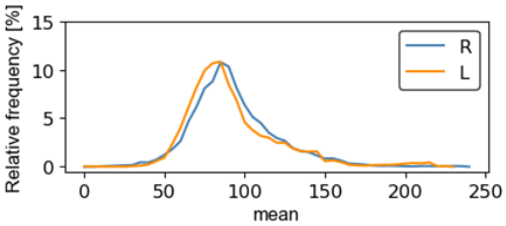
b) 画像の基礎的特徴量

図-7 は crop 画像の左右岸別の平均画素値の相対度数分布である．凡例の R が右岸，L が左岸の分布を示している．ここでは RGB の 3ch の平均を取った結果と，RGB の各 ch の平均を取った結果を示している．定性的に寒色と判断される右岸の画像群において B の ch の平均値が，左岸に比べて若干高く分布しているものの，この分布から左岸と右岸を区別するのは困難である．また RGB の 3ch の平均と各 ch の平均の分布がほぼ同等であり，色合いがグレー主体である護岸画像の特徴が反映されている．crop 画像のある ch の平均値と残り 2ch の平均値の差が大きくなった場合，植物などの護岸以外のオブジェクトが画像内に含まれることが推察される．

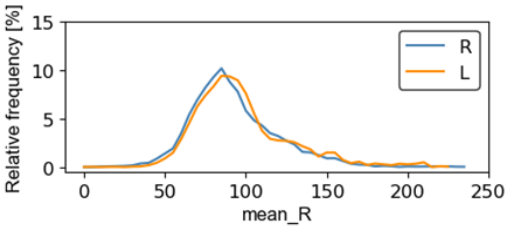
図-8 は crop 画像の 3ch の画素値から算出した左右岸別の標準偏差の相対度数分布である．crop 画像に写る物体が護岸のみであった場合，画像全体の色がグレーとなり画素値のばらつきが小さくなることから，偏差が比較的小さい値に集中した．さらに日が差す左岸の画像の方がひび割れや陰影とのコントラストが大きくなるためか，偏差がやや大きな値に分布した．

図-9 は平均と偏差の分布において，発生頻度の低い極値の画像の例を示している．RGB の平均画素値の小さい画像は橋梁下部の画像が多く，平均値の大きい画像は白色コンクリートで作成したパラペットが多くを占めていた．一方で標準偏差の小さい画像はエッジの少ない護岸が多く，標準偏差の大きい画像はパラペットと空が同時に写り込む画像が多く見られた．

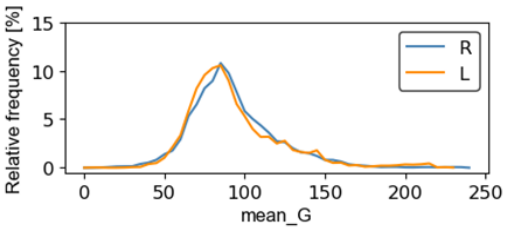
図-10 は crop 画像を HSV 色空間に変換し，左右岸別に色相（Hue），彩度（Saturation），明度（Value）の平均値を示した相対度数分布である．色相のグラフの横軸の文字色は値が示す色に対応している．3 つのグラフより左右岸で違いが現れたのは色相と彩度の分布であった．色相において左岸は橙に分布が集中し，右岸は青に分布が集中するという顕著な違いがあり，**図-5** の画像を見たときに人が受ける印象とデータの分布が概ね一致した．さらに彩度は日が差す左岸がやや高い値に分布した．一方で明度の分布の違いはほとんどなく，定性的に明るさが異なるように見える本検証区間での左右岸の差は見られなかった．



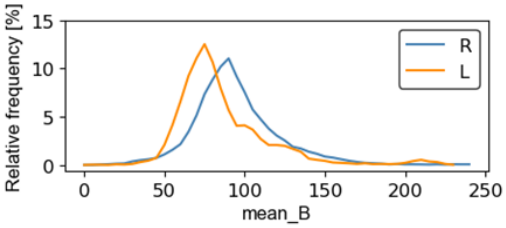
(1) RGB の平均画素



(2) R の平均画素



(3) G の平均画素



(4) B の平均画素

図-7 平均画素値の相対度数分布

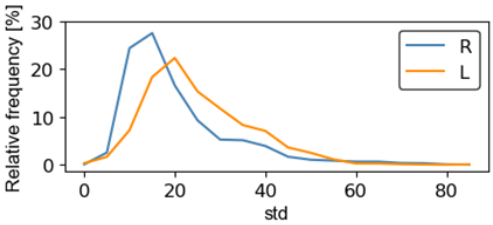


図-8 画素標準偏差の相対度数分布

(1) 平均:小

(2) 平均:大

(3) 偏差:小

(4) 偏差:大

図-9 顕著な特徴量を持つ crop 画像例

(3) 護岸画像の特性定量化結果の総論

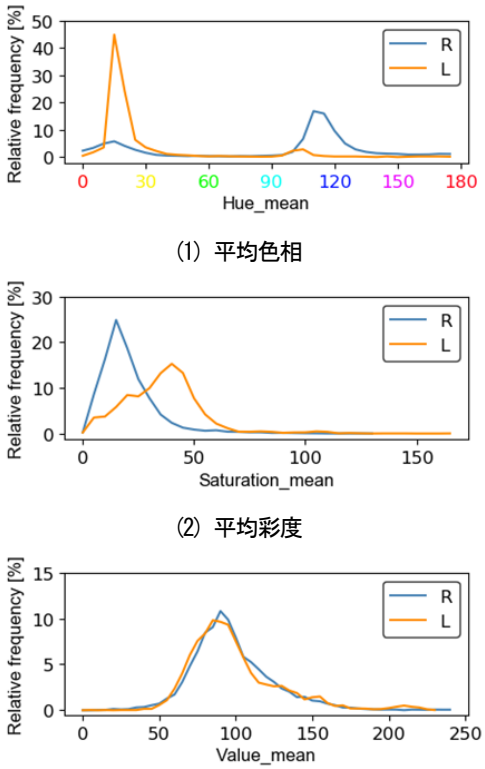
前節より，教師ラベルの不均衡の問題が明らかになった．さらに画像の持つ色情報の分布をさまざまな観点から整理することで，左右岸の画像間で色相と彩度の分布に差異があることが明らかとなり，テクスチャの構成要素の一つである色合いを HSV 色空間の色相と彩度によって評価できる可能性が示唆された．以上を踏まえると，汎化性の高いモデルを作成するには，教師ラベルと色相・彩度の不均衡問題を解消する必要があると考える．色相・彩度の不均衡については，数の少ないデータを拡張するオーバーサンプリング，多いデータを間引くアンダーサンプリングなどの対策が検討される．ラベルの不均衡には上記のサンプリング手法に加えて，少ないラベルに対する学習率を変更することなどが検討される．

4. IN 導入モデルの性能評価実験

(1) 実験内容

本章では，図-3 に示す正規化層として BN のみを取り入れた既往のモデルと既往のモデルに IN を導入したモデルを作成し，テクスチャの異なる画像間でのモデルの性能を比較する．なおここで示す性能は，3 章で示した左右岸の画像を混ぜた教師によって学習させたモデルの左右岸別の推論精度によって評価する．

図-11 に IN を導入したモデルの構造図を示す．本実験では Pan [17]らの論文を参考に，3 つの構造のモデルを作成した．いずれも図-3 に示す残差ブロックとその前後のブロック単位で変更を加えた．ケース A では残差ブロック内の最初の畳み込み層の直後に，BN 層と正規化する ch を二分するように IN 層を加えた．ケース B では Pan らの論文に準拠して，残差結合直前の Relu 層を取り除き，残差結合を経た直後に IN 層と Relu 層を加えた．ケース C では図-3 のモデルに準拠して，残差ブロック内は変更せず残差結合を経た直後に IN 層のみを加えた．Pan らの論文より，IN 層はモデル全体の中でも浅い層に導入することが推奨されているため，A～C のケースのモデルを図-3 の Conv1 に導入したケースと Conv1 および Conv2 に導入したケースで学習させた．

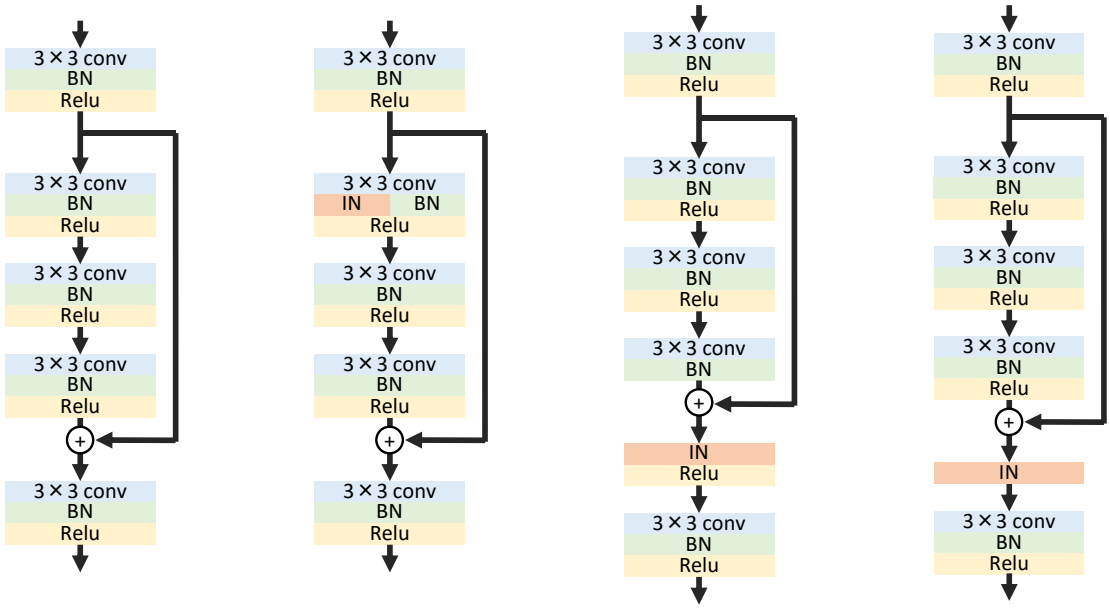


(1) 平均色相

(2) 平均彩度

(3) 平均明度

図-10 HSV 色空間の相対度数分布



(1) original (BN)　(2) ケース A　(3) ケース B　(4) ケース C

図-11 Instance Normalization 導入時のモデル構造

学習および評価に使用する画像は，3章で示した16380枚のcrop画像のうち，ひび割れラベルの付いた画像のみとした．ラベル付き画像を対象とするのは，ラベル量の不均衡の解消および学習の効率化を目的としている．さらにラベル付きcrop画像を，モデルのパラメータ更新に用いるtrain画像，過学習の判断に用いるvalidation（以下，valとする）画像，モデルの性能評価に用いるtest画像に分割する．分割にあたって，同じ元画像から作成したcrop画像がtrain画像とtest画像内に含まれると，train画像でtest画像のテクスチャを学習する恐れがあることから，test画像に用いる元画像を予め選定し，残った元画像をtrain，valに用いた．

test画像は左岸から3枚，右岸から3枚の計6枚を選定した．この時，画像枚数やひび割れラベルの量が左右岸で同程度となり，かつ色相の分布がばらつくようにした．**表-2**にtest画像の内訳を示す．左岸側の3枚の画像が持つひび割れラベルの偏差はやや大きいが，左右岸全体で合計するとほぼ同程度の数であった．さらに色相の平均値は，**図-10** (1)に示される分布と比較し，ある程度ばらつきを持った状態で画像が抜粋されていることを確認した．以上より左右岸で公平な評価を可能とするtest画像を抽出できたと判断した．

trainおよびvalの画像は，test画像を抜いた残りの61枚から作成したcrop画像を用いるが，testと同様に左右岸で枚数の偏りが無くなるよう，画像枚数の少ない左岸側に合わせて右岸の画像をランダムにアンダーサンプリングした．サンプリングで集まった左右岸合計2358枚の画像からランダムに9対1で画像を分割し，9割をtrain，1割をvalとした．最終的に得られたtrain，val，test画像の内訳は**表-1**に示す通りである．なおtrainとvalは学習の度にシードを固定した疑似乱数によって分割した．

モデルの学習は0～2の3つのシードによってtrainとvalの分割方法を組み替えて，3回×7モデルの計21回実施した．学習時はエポックごとにval画像によってLossを評価し，最もval Lossが低くなったモデルを記録した．これにより7種のモデル構造に対して各3つのモデルが生成されるが，testデータを使って算出したIoUの平均値が最も高くなったモデルをベストモデルとし，本実験で性能を比較する対象のモデルとした．上記の学習で得られたモデル構造の異なる7つのベストモデルは，train，val，testデータを対象とした左右岸別のPrecision，Recall，IoUの平均値によって評価した．

学習条件はモデル以外同一とし，中間層の活性化関数はRelu，バッチサイズは16，エポック数は40，最適化関数は学習率0.01のMomentum SGD，LossはBinary Cross Entoropy Loss（以下，BCELossと呼ぶ）とした．学習フレームワークにはPyTorch 1.4.0を用いた．

(2) 学習結果

a) ベストモデルの選定

3つのシードによって得た計21のモデルより，ベストモデルを選定した．全てのモデルでtestデータを推論し，IoUの平均値を算出した結果を**図-12**に示す．ここでBNはoriginalモデルを示し，ABCのアルファベットは3つのケース，末尾の数字の1はConv1，2はConv1およびConv2にINを導入した結果を示している．なお推論結果はモデルがひび割れと予測する確率がピクセル毎に0から1の値で表示されるため，0.5以上をひび割れ，0.5未満をひび割れ以外とした．

図-12よりBN，A2，B1，B2ではseed0のモデル，A1，C1，C2ではseed1のモデルをベストモデルとした．最も高い性能を示したのがC2，次いでA1であったが，平均して高い性能を示したのはB2であった．

b) 学習曲線

図-13に7つのベストモデルの学習曲線を示す．学習曲線上の菱形のプロットは，val Lossが最低値となり，その学習においてモデルが記録されたときのEpochとLossを示している．図よりいずれのモデルにおいても，train，valともにLossがほぼ横ばいに推移するまで学習していることが確認された．trainにおいてLossが最低値を示したのはB2の0.001403で，B1は0.001407と，僅差でB2に及ばなかった．valにおいてLossが最低値を示したのはA2であったが，C以外は概ね同程度の値となった．ここで，BN，A2，B1，B2とA1，C1，C2ではモデル採択時のシードが異なるため，trainとvalに用いた画像が異なるという点に注意されたい．trainとvalの組み合わせの違いが起因したためか，seed1で作成したモデルはseed0で作成したモデルに比べてLossが大きくなる傾向にあった．

表-2 test画像の内訳

画像	左右岸	crop 画像枚数	ひび割れ ラベル px	色相平均値
①	R	99	144085	110
②	R	87	112519	156
③	R	98	112940	133
④	L	72	82176	19
⑤	L	103	179767	29
⑥	L	103	93653	25

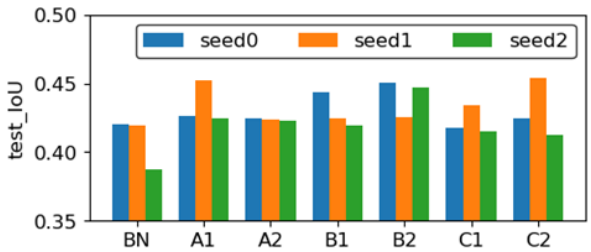


図-12 全モデルのtestデータでのIoU

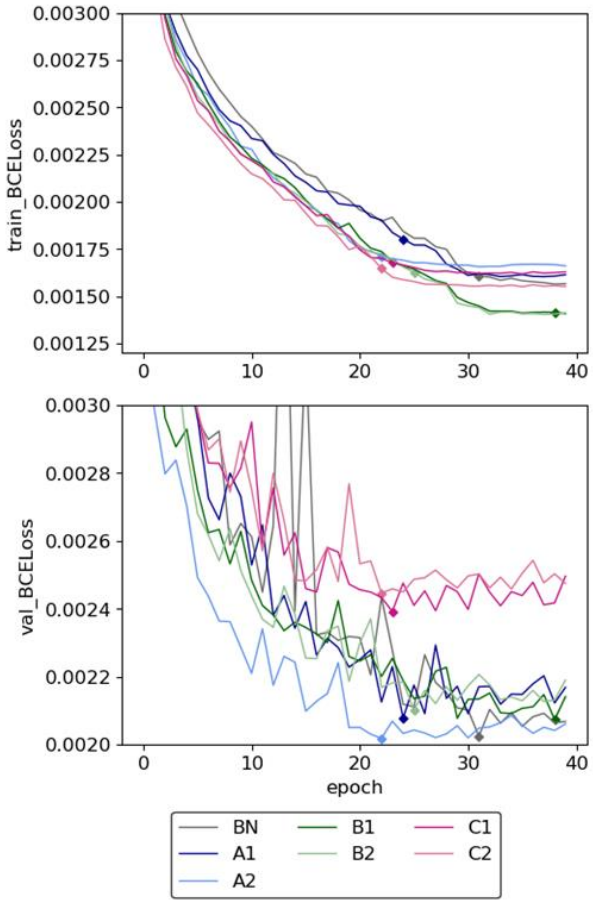


図-13 学習曲線

c) 画像への推論結果

全ベストモデルで推論を行い，train，val，test ごとに Precision，Recall，IoU の平均値を算出した．その結果が図-14 である．全てのモデルにおいて，Recall が Precision を上回った．さらに，全てのモデル・評価指標において左岸が低い値を示した．特に左岸の train・val の評価結果は低く，同じ左岸の test と比較しても明らかに低い値となった．train と val の Precision および Recall を左右岸で比べると，Precision に比べ，Recall の差は小さくなった．

図-13 に示す学習曲線において最低あるいは最高の Loss 値を示したモデルの BCELoss と，Precision，Recall，IoU の評価結果に相関性は見られなかった．特に B1 は他のモデルと比較して遅い 38 エポック目のモデルが採択されていたため，train 画像に対して高い性能を示すことが予測されたが，B1 以上の性能を持つモデルは他に多数存在した．

IN を導入したモデルと BN のみのモデルの IoU を比較すると，左右岸に関係なく IN モデルの方が高い IoU を示す傾向にあった．一方で左右岸別の Precision，Recall，IoU の評価結果が全てが最良となったモデルはなく，個々のモデルに一長一短があった．IoU という指標で評価するのであれば，C2 が最良の結果であった．

(3) 考察

図-14 に示すように，左岸の train，val において性能が全体的に低下した原因は，左右岸でのタスク難度の違いおよび左岸の教師の量の不足が影響していると判断する．タスク難度の違いは日光の有無が要因にあると考える．図-4 のような凹凸のある護岸ブロックでは日光が差すことで生じた護岸上の陰影が誤検出を誘引する．実際に日光が差す左岸は Recall に比べて Precision の低下が激しいため，陰影の誤検出の増加が IoU 低下につながったと判断した．図-15 に本実験で用いた crop 画像とその推論結果のサンプルを示す．図-15 (5)や(6)のように陰影が写った画像において，誤検出が発生していることが分かる．また表-1 に示す通り，本実験で用いたデータは右岸の方が枚数が多く，左右岸で枚数が等しくなるようにランダムサンプリングしても，右岸の方がひび割れラベル数が多かった．よって，左右岸で教師画像の不均衡が生じ，右岸の方が学習しやすい状況にあったと判断した．本実験において，test 画像は右岸と左岸を可能な限り等しい条件で評価できるよう，定量的かつ定性的に選んだ 6 枚の画像を用いたため，左右岸で乖離の少ない結果が得られたが，train および val の右岸画像はランダムにサンプリングをしたものを用いたため，左岸の特徴量やひび割れ分布に合わせて右岸画像をサンプリングする必要があったと考える．図-14 の結果において，seed2 のモデルが総じて低い IoU となったのも，train と val の組み合わせが影響していたと推察した．

本実験では BCELoss と推論結果の評価結果の違いが明らかとなり，学習時の Loss を変更する必要性が示された．BCELoss であれば AI の判定の揺らぎが評価に反映されるが，IoU ではある閾値を持って推論結果が 2 値化されるため，判定の揺らぎは評価に反映されない．護岸のひび割れ管理において AI のひび割れ検出結果からひび割れの面積等を計測する場合，ひび割れか否かを示した 2 値画像で評価することが想定されるため，IoU が高くなるモデルを用いることが望ましい．したがって IoU や Dice といった推論結果とラベルの重なり具合によって評価する Loss を取り入れることが課題であると考えた．

上記に示す課題解決の必要性は残されているものの，本実験の結果より IN がテクスチャの異なる画像間での推論精度向上に効果があったと判断した．一方で，ひび割れ検出に適したモデル構造は一意には定められず，対象データおよび重視する評価指標によって変わることが明らかとなった．IN 層の最適な深さはモデルによって異なり，A では 1 層，B および C では 2 層の方が良くなる結果となった．ただし本実験の結果からは Pan らの論文と同様に層数に関しての直接的な関連は見受けられず，ひび割れ検出に最適なモデルは解析対象データやモデル作成方針に依存するものであると考えた．

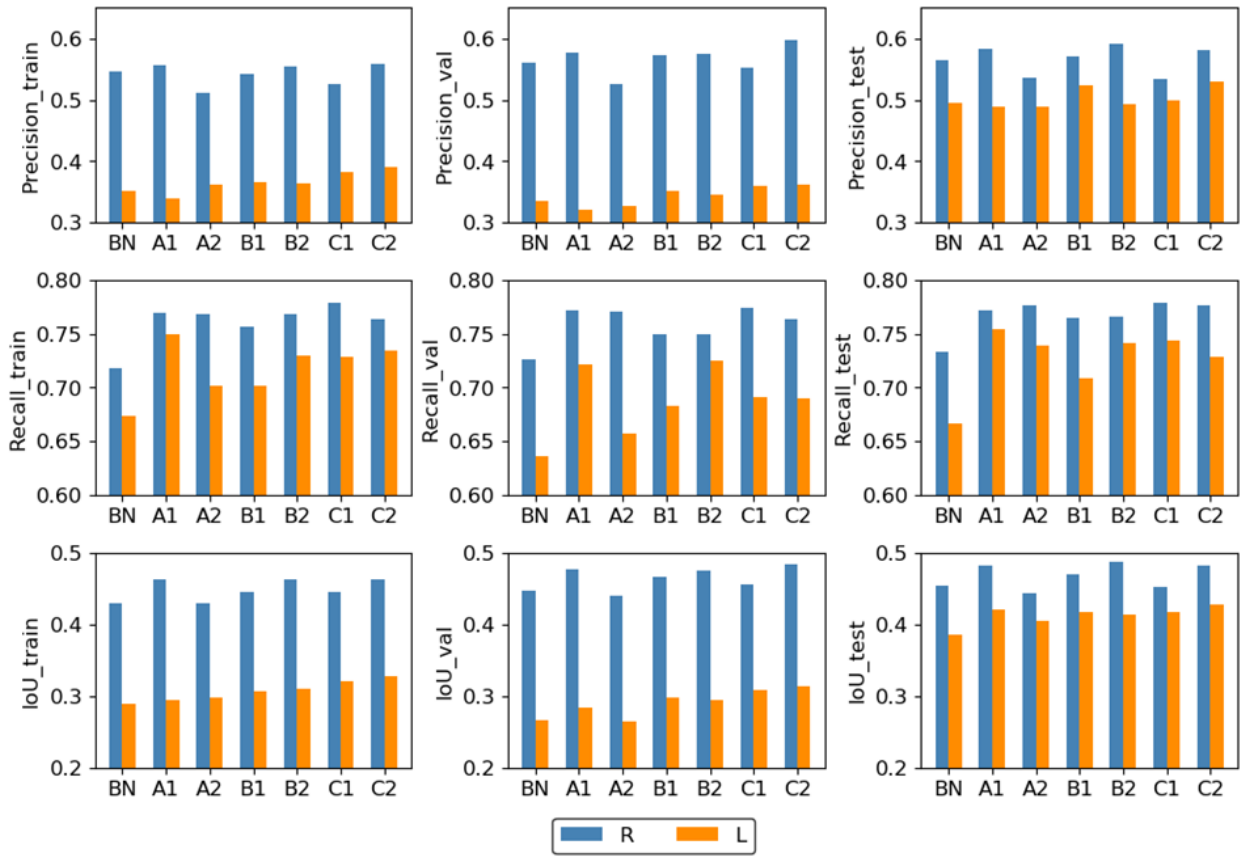


図-14 モデルの性能評価結果

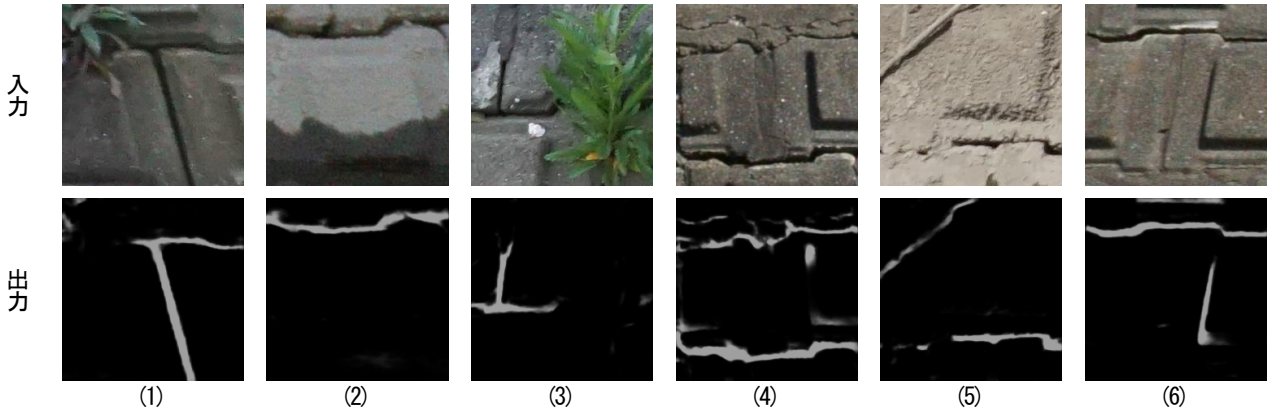


図-15 crop 画像と C2 モデルでの推論結果例

5. 最後に

(1) まとめ

本稿では，護岸のひび割れ Segmentation において，異なる 2 時期に撮影した同じ現場の画像において，護岸のひび割れ検出精度が低下するといった事例を課題として示した．さらにこの精度低下は画像のテクスチャに起因していると仮定し，テクスチャ変化に頑健となるモデルの作成手法を検討した．

3 章では画像の特性の違いを定量評価するべく，護岸画像の特徴量を左右岸に分けて整理した．その結果 HSV 色空間の色相と彩度において，左右岸の特徴量の分布に顕著な違いがあることが分かった．

4 章では画像のテクスチャ変化に頑健になるとされる Instance Normalization を導入したモデルをさまざまな構造で構築し，従来の Batch Normalization のみモデルとの性能を比較した．その結果，ひび割れ検出において最も精度が向上するモデル構造こそ明らかにはならなかったが，特性の異なる左右岸の護岸のひび割れ検出において，Instance Normalization 導入モデルが高い精度を示す傾向にあることを確認した．

(2) 今後の課題

本稿ではモデルの入力画像の特性と出力結果の関係から Instance Normalization の効果を検証したが，この効果をより詳細に把握するためには，モデルの中間層の出力の違いを明らかにする必要がある．具体的には Pan [17]らが行ったように，中間層の出力を取り出して，さまざまな画像間の KL ダイバージェンスを算出することで，Batch Normalization 層と Instance Normalization 層を通過した出力の分布特性の違いを明らかにすべきである．中間層に存在する畳み込み層やプーリング層の出力値も同様に解析することで，既往研究の慣例にならって定めていた層の数の意味を再考し，より最適なモデル構造の検討を行うべきと考える．また本稿の結果より train, validation, test に用いる画像の選定によって学習結果が大きく変わることが示された．上記に示す中間層出力の分布特性の把握を通じて，AI がひび割れを認識する上で重視している特徴を明らかにし，学習に必要な画像の特性を把握したい．

参考文献

1) 国土交通省：令和 2 年版国土交通白書, pp.114-115, p.121, 2020.
2) 富山市：持続的かつ適正な橋梁マネジメントの実現に向けて, p.7, 2020.
3) 国土交通省道路局：道路橋定期点検要領, p.2, 2019.
4) 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, p.5, p.9, p.19, 2019.
5) 吉田龍人, 藤井純一郎, 大久保順一, 天方匡純：護岸のひび割れ検出 AI の作成とオルソ画像を活用したひび割れ幅の評価, AI・データサイエンス論文集 1 巻 J1 号, pp.508-513, 2020.
6) 天方匡純, 藤井純一郎, 吉田龍人：深層学習技術のインフラ維持管理への導入に際しての留意点, AI・データサイエンス論文集 1 巻 J1 号, pp.35-40, 2020.
7) 齋藤彰儀, 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,デジタルプラクティス Vol.10 No.2, 情報処理学会, 2019.
8) Ian J. Goodfellow, Jonathon Shlens and Christian Szegedy.: Explaining and Harnessing Adversarial Examples, arXiv: 1412.6572, 2015.
9) Robert Geirhos, Patricia Rubisch, Claudio Michaelis, Matthias Bethge, Felix A. Wichmann and Wieland Brendel.: ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness, arXiv: 1811.12231, 2019.
10) Olaf Ronneberger, Philipp Fischer and Thomas Brox.: U-Net: Convolutional Networks for Biomedical Image Segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp.234-241, 2015.
11) Tran Minh Quan., David G. C. Hildebrand and Won-Ki Jeong.: FusionNet : A Deep Fully Residual Convolutional Neural Network for Image Segmentation in Connectomics, arXiv : 1612. 05360, 2016.
12) Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun.: Deep residual learning for image recognition, IEEE conference on computer vision and pattern recognition, pp.770-778, 2016.
13) Sergey Ioffe and Christian Szegedy.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, the 32nd International Conference on Machine Learning, pp.448-456, 2015.
14) Jimmy Lei Ba, Jamie Ryan Kiros and Geoffrey E. Hinton.: Layer Normalization, arXiv: 1607.06450, 2016.
15) Dmitry Ulyanov, Andrea Vedaldi and Victor Lempitsky.: Instance Normalization: The Missing Ingredient for Fast Stylization, arXiv:1607.08022, 2017.
16) Yuxin Wu and Kaiming He.: Group Normalization, arXiv: 1803.08494, 2018.
17) Xingang Pan, Ping Luo, Jianping Shi and Xiaoou Tang.: Two at once: Enhancing learning and generalization capacities via ibn-net, European Conference on Computer Vision, pp. 464-479, 2018.
18) 石井明, 天方匡純, 菅原宏明, 藤井純一郎, 小篠耕平, 六門直哉：非 GPS 環境下における UAV の自律飛行実験 -AI によるコンクリート護岸点検・診断用画像の効率的な撮影のために-, 建設ロボットシンポジウム論文集 19 巻, 2019.
19) Shaoqing Ren, Kaiming He, Ross Girshick and Jian Sun.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, arXiv:1506.01497, 2016.

(Received June 30, 2021)
(Accepted August 31, 2021)

A STUDY FOR STABLE CRACK SEGMENTATION REGARDLESS TEXTURE IN REVETMENT IMAGE.

Ryuto YOSHIDA, Junichiro FUJII and Junichi OKUBO and Masazumi AMAKATA

Convolutional Neural Networks are biased towards recognising textures. Thus, the accuracy is also reduced between images with different textures in crack Segmentation. In this study, We evaluated the difference of texture by calculating features of crack images. To make a model that can detect cracks more stably, we compared the performance of a model with only batch normalization and models contained instance normalization.