

技術動向レポート

# 機械学習による河川画像の土地分類手法の技術開発

情報通信研究部  
コンサルタント 加田 匠

本レポートでは、治水対策の検討に必要となる水域と水域以外を分類する土地分類の課題に対し、河川周辺の航空写真から機械学習により自動で水域を分類することを試みた結果について報告する。

## 1. はじめに

日本の国土の約1割は洪水氾濫区域であり、洪水時の河川の水位より地盤の高さが低く、河川からの洪水氾濫によって浸水する可能性がある。この洪水氾濫区域に人口の約5割、資産の約4分の3が集中し、さらに、都市の大部分が洪水時の河川水位よりも低い等、河川の氾濫による被害は甚大であるため、治水対策は極めて重要である<sup>(1)</sup>。また、近年、局地的な豪雨が増加し、河川の水が堤防を越えてあふれ出す越水や堤防の決壊等の甚大な被害が増加する等、治水対策の重要性は増している<sup>(2)</sup>。「令和2年7月豪雨」では、西日本から東日本、東北地方の広い範囲で大雨が発生し、特に九州地方では熊本県を中心に記録的な大雨となり、球磨川等の一級河川が相次いで氾濫する事態となった<sup>(3)</sup>。

治水対策には、新たに堤防を築く築堤や、川幅を広げる引堤、洪水を一時的にためる遊水地等があるが、これらを際限無くおこなうことはできない。現実的には、氾濫の危険度や浸水の被害度が大きいと予想される重点区域に優先して治水対策を施すこととなり、そのためには、河道変遷をもとにした蛇行流路の分析や、河川周辺の土地環境等の把握等、洪水や侵食によ

て絶えず変化する河川の状態をモニタリングすることが望ましい。現状は、現地に出向いての現地調査や航空写真からヒトによる目視判読により河川状況のモニタリングをおこなっているものの、ともに多大な時間と労力を要し、時間と共に変化する多くの河川を網羅的、即時的に把握することは困難である。

本レポートでは、河川に主眼を置いた地形判読の自動化に向けた基礎検討として、機械学習を用いて航空写真から水域と水域以外を分類(以降、水域分類)する手法について検討をおこなった。定期的に航空写真を撮影する必要はあるものの、コンピュータ等を用いて航空写真から自動的に河川周辺の地形を判読することができれば、網羅性、即時性を大きく向上させることが期待できる。筆者の調べでは、機械学習を用いて航空写真からの地形の自動判読を目的とした文献<sup>(4)</sup>はあるものの学習データの作成に留まっており、地形の自動分類には至っていなかった。

次章以降、複数の手法によって水域分類を試行した結果について述べる。第2章では、画像の色味成分のみを用いて水域分類を試行した結果について説明する。第3章では、深層学習モデルである U-Net を水域分類に適用した結果に

ついて説明する。最後の第4章において、上記の試行結果を総括した上で、今後の展開可能性や、他分野への適用可能性について展望する。

## 2. 色味による水域分類の試行

水域の自動判読に関する初期検討として、航空写真の各ピクセルの色味成分について、①水域(河道内)、②植生、③岩・砂州等、④堤防の4つの土地条件への分類をおこなった。予め、4つの土地条件の代表色を指定し、航空写真の各ピクセルについて4つの代表色との(R、G、B)の3次元距離を比較、最も近い土地条件であると判定した。

分類結果を図表1に示す。各土地条件の代表色は右の凡例に示す通りとし、上段が入力とした航空写真、下段が4つの土地条件に分類した結果である。河川中心付近のみを対象として分類をおこない、黒色の部分は分類の対象外とした。

図表1左の事例では、4つの土地条件の分類に概ね成功した。中央の事例では、水域については正しく分類されたものの、水域に近い色味の植生部分が水域と誤判定された。さらに右の事例では、河川が濁っていたことにより完全に誤っ

た分類結果となり、河道を岩・砂州等と誤判定し、植生部分を水域と誤判定した。

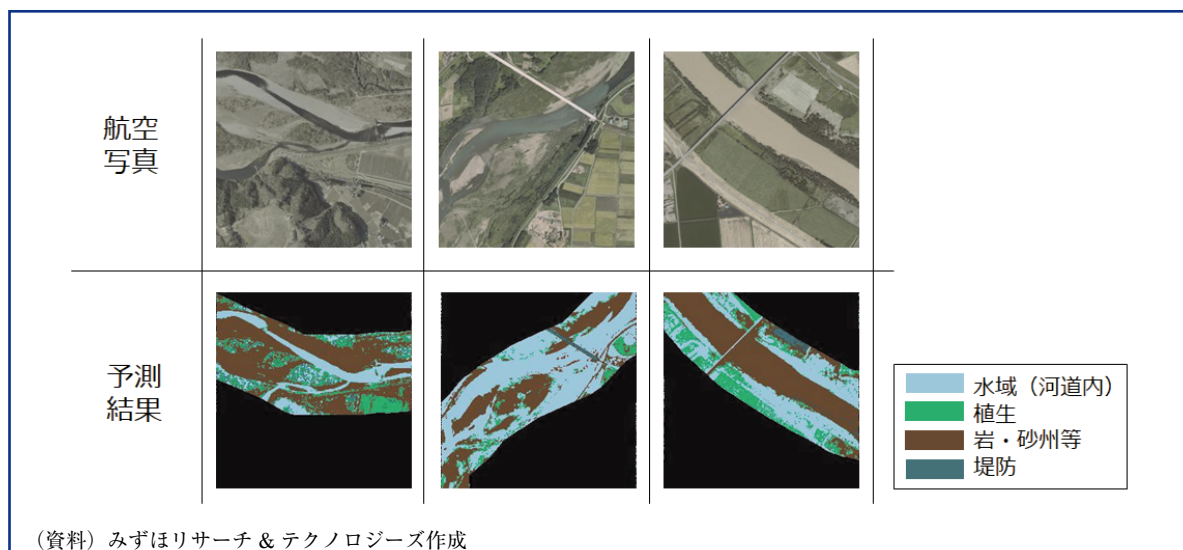
画像全体の色合いを調整する等の方法により精度を向上させる余地はあるものの、各ピクセルの単独の色味だけで水域を判定することは困難であり、周囲のピクセルの色味との関係性や形状を考慮することが必要であると予想された。

## 3. 深層学習による航空写真の水域分類

周囲のピクセルの色味情報を考慮することを目的として、深層学習を取り入れた手法を用い、水域と水域以外の2つに分類することを試みた。

深層学習は、画像認識・音声認識等の分野で急速に発展している機械学習の方法論の1つである。特に、CNN(畳み込みニューラルネットワーク)は、周辺との関係性を含めての判断が可能となる手法であり、その高い認識精度から、画像全体を分類(例えば、犬の画像か、猫の画像か)する画像分類の分野でのデファクトスタンダードと言える。また、CNNは画像認識におけるセマンティックセグメンテーションというタスクに対しても強みを発揮する。セマンティックセグメンテーションとは、画像内の各ピクセ

図表1 色味による予測結果



ルがどのカテゴリに属するか(例えば、あるピクセルは犬に含まれるか)を分類するものである。

本レポートでは、セマンティックセグメンテーションにより、水域と水域以外の2つのカテゴリに分類することを試みた。深層学習のモデルとしてU-Net<sup>(5)</sup>を使用した。U-Netは、医療用画像を解析するために開発された、セマンティックセグメンテーション用の比較的シンプルなモデルであり、その分類性能の高さから、現在では医療用画像以外に対しても応用が進んでいる。

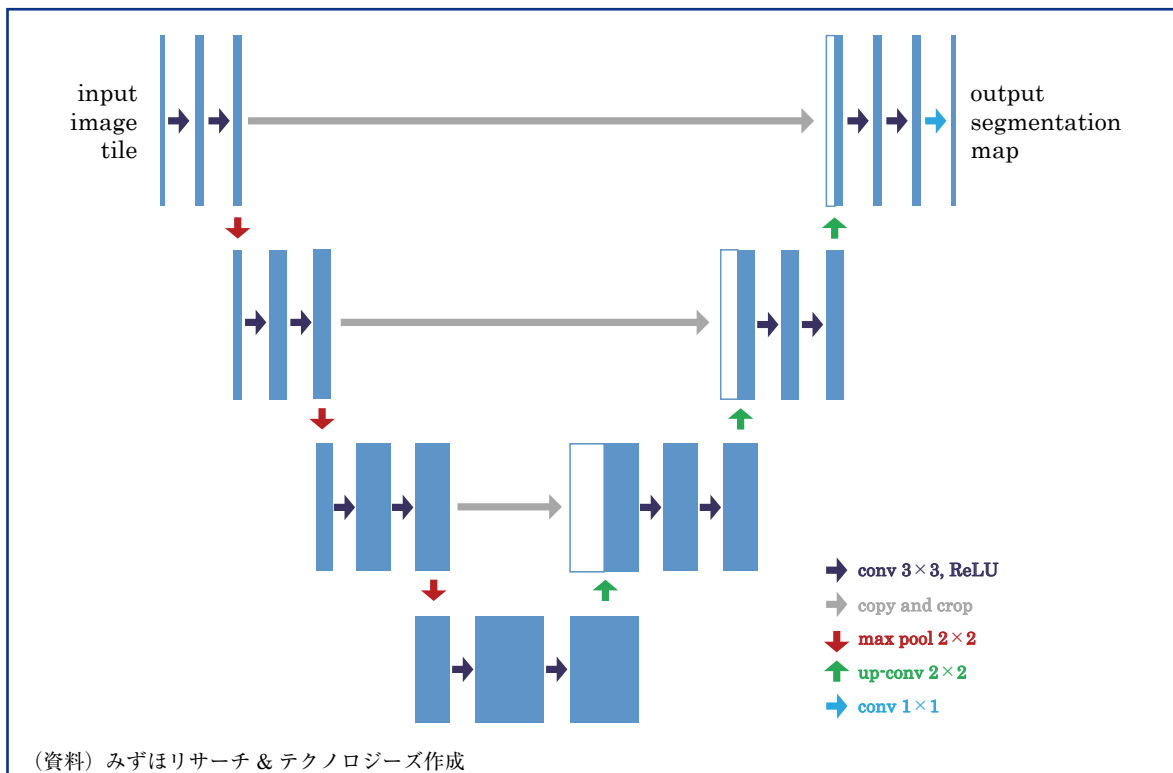
U-Netの一般的なネットワーク構造を図表2に示す。畳み込み層のフィルタの大きさやネットワークの深さ等は本レポートでの試行と一致するものではないものの、本レポートにおいても同様のネットワーク構造を用いた。

入力された画像に対して畳み込みとプーリング(図表2の conv 3x3, ReLU, max pool 2x2)を繰り返し適用し、周辺のピクセルの色味情報が

考慮された画像の特徴を抽出する。畳み込みは、フィルタを用いて画像の特徴を抽出するものであり、周辺のピクセルの色味情報も考慮して水域の特徴と類似した形状(フィルタ形状)の箇所を検出し、特徴マップが作成される。プーリングは、特徴を保持しつつ特徴マップのサイズを小さくするものであり、フィルタ形状の位置ずれを吸収する機能を持つ。その後、逆畳み込みと呼ばれる処理(図表2の conv 3x3, ReLU, up-conv 2x2)を繰り返しおこなって入力画像と同じ画像サイズまで戻し、各カテゴリに属する確率を各ピクセルについて出力する。U-Netの特徴は、畳み込み層の各層の結果を、対応する逆畳み込み層に伝播させ(図表2の copy and crop)、畳み込みの際に失われた解像度を回復することであり、画像中の物体位置検出の精度を上げている。

学習用データとして、800枚の航空写真、お

図表2 U-Netの一般的なネットワーク構造



よび、航空写真に対応して水域部分をマスキングした正解画像を使用した。さらに、航空写真と正解画像の双方を90度、180度、270度回転させた画像を学習用データに加え、データ数を4倍に拡張した。

評価用データとして石狩川周辺の147枚の航空写真を用い、学習したモデルに入力して水域分類をおこない、正解画像との比較により性能を評価した。評価用データの航空写真は国土地理院が提供する「電子国土基本図(オルソ画像)(2007年～)」<sup>(6)</sup>を、評価用データの正解画像は「川だけ地形地図」<sup>(7)</sup>を使用した。ともにタイル画像としてデータが公開されており、タイル座標とズームレベルをURLに組み込むことで、指定した領域のデータをダウンロードすることが可能である。本レポートでは、タイル画像の拡大率を表すズームレベルを14に設定(地図を2の14乗四方に分割)し、石狩川周辺の航空写真および正解画像147枚をそれぞれダウンロードした。評価用データの例を図表3に示す。左が航空写真であり、右の正解画像の水色の部分が航空写真に対応した水域を表す。

U-Net に評価用データの航空写真を入力し、各ピクセルが水域である確率を出力した後に、

確率が0.5以上のものを水域、0.5未満を水域以外と判定した。この判定結果と正解画像とをピクセル単位で比較し、正解・不正解を求め、その正解率を性能評価の指標とした。なお、本レポートでは正解率を指標としたが、評価用データの水域の面積比率が高い(低い)場合に、入力画像に関わらず全ピクセルを水域(水域以外)と予測する誤ったモデルであっても正解率が高くなってしまうため、実際には、再現率、特異度等の指標も考慮する必要がある。











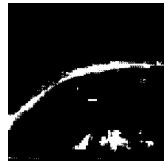
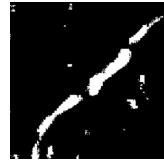
性能評価の結果、147枚の正解率は平均で84.4%を示した。図表4に正解率90%以上の事例を、図表5に正解率90%未満の事例を紹介する。図表4、図表5とも、上段が評価用データの航空写真、中段が正解画像、下段がU-Netによる予測結果と正解率である。

図表4の右2つの事例では、2章の色味による水域自動判読では誤判定となった河川の色味が異なる場合においても、河川の形状が正しく認識され、正解率が90%以上となった。一方、正解率の低い図表5の事例では、評価用データの中で相対的に川幅が狭い航空写真に対して水域を正しく認識できなかった。一般的に、河川の上流・下流では川幅が異なる他、航空写真のズー

図表3 評価用データの例









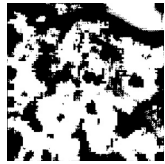
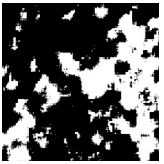
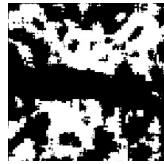
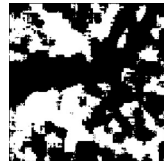


図表 4 U-Net の予測結果 (正解率90%以上の事例)

航空 写真				
正解 画像				
予測 結果				
	90.5%	94.6%	90.8%	93.6%

(資料) みずほリサーチ & テクノロジーズ作成

図表 5 U-Net の予測結果 (正解率90%未満の事例)

航空 写真				
正解 画像				
予測 結果				
	44.2%	67.0%	59.3%	60.6%

(資料) みずほリサーチ & テクノロジーズ作成

ムレベルによっても川幅は異なるため、精度の向上に向けては、このようなスケールの違いに対応することが必要と考えられる。

本レポートで紹介する事例は一定の時間的制

約の中でおこなった試行であり、精度向上に向けての検討が充分とは言えない。実運用に向けては、機械学習や画像処理に関する工夫を加えることにより、精度を向上させる必要があると

考えている。機械学習の観点では、

- ・河川の様々な形状や色味に対して正しく分類できるよう、学習に用いる航空写真の質と量を増やす。
- ・分類精度が川幅の太さに依存しないよう、拡大・縮小に関するデータ拡張を適用する。
- ・同じく、分類精度が川幅の太さに依存しないよう、CNNのモデルを検討する。例えば、DeepLabv3+<sup>(8)</sup>、AWMF-CNN<sup>(9)</sup>等のモデルが候補となる。
- ・河川特有の時間変化の特性を利用することで精度が向上する可能性があるため、時系列のデータを学習に用いる。その際には、モデルの検討も必要となる。

等が考えられ、古典的な画像処理技術との組み合わせの観点では、

- ・時系列の過去の判定結果を参照し、各ピクセルの水域の確率計算に盛り込む。
- ・エッジや平滑度等の特徴量を機械学習の入力に追加する、または、画像処理と機械学習の判定結果とを組み合わせで判定する。
- ・各ピクセルの判定結果に対し、微小領域や孤立領域を削除する。

等が考えられる。

本章では、機械学習を用い航空写真からの水域と水域以外の2分類を試みた。まずは、この2分類の精度向上が前提とはなるが、その先として、①水域(河道内)、②植生、③岩・砂州等、④堤防の4つのクラスター(土地条件)への分類や、河川周辺以外を対象とした汎用的な分類、さらには、時系列のデータ解析による河道変遷の分析をおこなう等、より有用性の高い情報を自動で取得するための技術開発に取り組んでいきたい。衛星画像に対する分類を自動化することができれば、地球規模で有用性の高い情報を得ることも可能となる。

#### 4. まとめ

本レポートでは、治水対策の前提となる河川周辺の現状把握の効率化を目的として、河川周辺の航空写真から、機械学習による自動での水域分類を試行した事例を紹介した。

第2章では、各ピクセルの色味のみで判定をおこなう手法での判定精度を確認した。第3章では、各ピクセルの色味だけではなく周辺のピクセルの色味も考慮可能な、深層学習モデルであるU-Netを用い、判定精度を確認した。河川の形状や周囲との関係性を含めて学習・推定することで、第2章で誤判定となった河川等の色味の変化に対する頑健性が向上した。

本レポートで紹介した事例は、あくまでも基礎的なものであり、決して高い分類精度とは言えない結果であったが、精度向上に向けて考慮すべき事項のヒントが得られた。また、衛星やドローン等、航空写真以外の様々な種類の画像を利用することで、目的に応じた様々な範囲、様々な分解能での土地分類が可能になっていくものと期待される。

画像認識における深層学習の大きな強みは、学習用データを用意することができれば、水域分類のような特定の目的を意識することなく、学習用データに適応して分類ロジックが自動的に学習される点にある。実際、防犯、医療、商業、観光など、様々な分野において深層学習を用いての画像からの自動分類に関する研究や実用化が進んでいる。ただし、いずれの分野においても、実用に耐えうる高い分類精度を得るためには、一般的に十分な質と量の学習データが必要である。また、適切な前処理やネットワーク構造を選択し、パラメータチューニングが必要となることも多い。

当社では、人の助けとなるべく、画像処理や深層学習等のAI技術の研究・開発に、引き続

き取り組んでいきたい。

#### 謝辞

2章で紹介した事例は、国立研究開発法人土木研究所寒地土木研究所から提供されたデータを使用しております。同研究所に深い謝意を表します。

#### 注

- (1) [https://www.mlit.go.jp/river/pamphlet\\_jirei/kasen/gaiyou/panf/gaiyou2007/pdf/c1-1.pdf](https://www.mlit.go.jp/river/pamphlet_jirei/kasen/gaiyou/panf/gaiyou2007/pdf/c1-1.pdf)
- (2) <https://www.jma.go.jp/jma/kishou/books/hakusho/2020/index1.html>
- (3) <https://www.data.jma.go.jp/obd/stats/data/bosai/report/2020/20200811/20200811.html>
- (4) <https://www.gsi.go.jp/common/000228103.pdf>
- (5) Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
- (6) <https://maps.gsi.go.jp/development/ichiran.html#ort>
- (7) <https://www.gridscapes.net/>
- (8) Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." Proceedings of the European conference on computer vision (ECCV). 2018.
- (9) Tokunaga, Hiroki, et al. "Adaptive weighting multi-field-of-view CNN for semantic segmentation in pathology." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.