

光学衛星画像を用いた機械学習による河道状態把握の検討

日本工営株式会社 大阪支店流域水管理部 ○片岡 健吾
日本工営株式会社 大阪支店流域水管理部 佐藤 京介
日本工営株式会社 仙台支店流域水管理部 西俣 淳一
日本工営株式会社 河川水資源事業部水工インフラマネジメント部 能美 享

論文要旨

揖保川本川における河道内の樹木繁茂が顕著な区間(計3.0km)を対象に、光学衛星画像を用いた教師あり機械学習による河道内の地被分類の自動判読を試行した。教師データに揖保川全区間のデータを用いた場合は約95%の精度、一部区間を用いた場合は60%精度で地被状態の判読が可能であることが確認された。また河道掘削・樹木伐採後のモニタリングの一手法として、本手法を用いた河道内状態把握の高度化・省力化の可能性を検証した。

キーワード：光学衛星画像，機械学習，ランダムフォレスト法，河道管理，河道点検，高度化・省力化

1. 研究背景と目的

近年日本全国各地の河川では毎年のように、豪雨に伴う河川災害が激甚化・頻発化している。河川管理者は、被害の拡大を最小限に抑えるため適切に河川を維持管理し、これまで河川改修等で確保された河道流下能力を維持することが重要である。河道流下能力の把握は、河川定期横断測量、空中写真測量、河床材料調査、植生調査等の結果を用いた水位計算により行うことが基本である。一方で、水位計算により流下能力が懸念される箇所、河道掘削や樹木伐採等の河川工事の区域においては、流下能力及び河川管理施設の機能・安定性への影響が懸念されることから、河道の変化を継続的に把握することが重要である。

こうした背景から、「堤防等河川管理施設及び河道の点検・評価要領，平成31年4月，国土交通省（以下「点検要領」という）」には、流下能力に関する点検事項¹⁾が示されており、河道のモニタリングは目視点検を基本としている。目視点検による河道点検は、現在 RiMaDIS（河川維持管理データベース）内に点検記録を蓄積しているが、現場条件（樹木繁茂、流れが速い等）によっては近接目視が困難であるため、橋梁等の定点からの遠方目視が基本であり、面的に河道を把握することは難しい。また出水により面的に変化する土砂や樹木の状況を目視により把握することは人手不足の観点から現実的ではない。現場での負担を考慮し点検の省力化を図りつつ、面的な河道状態を把握することが求められている。

一方でこれまで我々は、河道内の地被状態を①面的に、②頻度高く、③安価に把握する技術として、衛星画像を用いた機械学習による面的な地被分類の自動判読技術の開発を行ってきた。例えば、数10kmの河川の河道内地被分類の自動判

読技術を開発し、その判読精度・有効性を確認している²⁾。

本検討では、今後の河道維持管理への現場実装を見据え、対象河川における衛星画像を用いた自動判読技術について試行を実施する。また、光学衛星画像を活用した河道管理の高度化・省力化の可能性について検討した。

2. 検討手法

(1) 検討概要

河道内地被分類の自動判読には、機械学習のアルゴリズムであるランダムフォレストを用いた。ランダムフォレストは、決定木と呼ばれるある特徴量を閾値で分岐させる分類木を大量に組み合わせる手法である。本研究で用いた解析フローを図-1に示す。本解析では、既往の河川水辺の国勢調査（植物調査）結果を基に地被状態分類の正解データ、トレーニングデータを作成し、機械学習による自動判読を実施した。自動判読結果は、既往の河川水辺の国勢調査（植物調査）結果より作成した正解データを用いて判読精度の確認を行った。

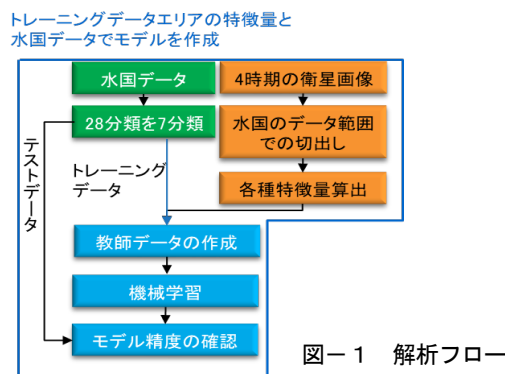


図-1 解析フロー

(2) トレーニングデータ

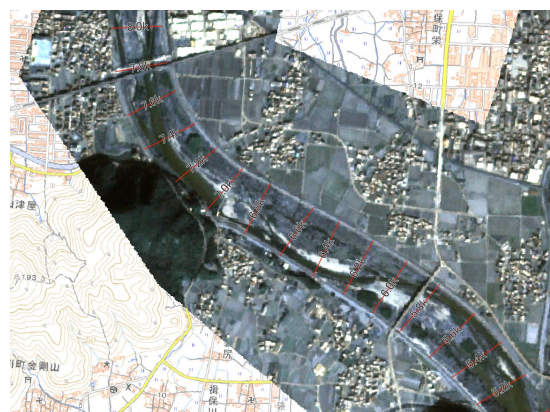
本解析では衛星画像の各メッシュに対し、既往の河川水辺の国勢調査(植物調査)における調査結果(平成31年度)を正解データとして用いた。河川水辺の国勢調査の植物調査は、河川における植物の生育状況を把握することを目的とし、河川内の植生基本分類として28分類(その他、不明を除く)に区分している。植生について細区分されており、例えば、単子葉草本群落については、ヨシ群落、ツルヨシ群落、オギ群落、その他といった詳細に分類している。一方、河川の水利解析では、植生は粗度や樹木群(死水域)として表現されるため、詳細な植物の種類までの判読は必要としない。そのため本検討では、水国の28分類の群落を①草本、②低木、③高木、④竹林、⑤その他植生、⑥裸地、⑦水面の7分類に変換(表-1)した。

表-1 分類項目

水国 28分類	本検討 7分類	水国	7分類
1 沈水植物群落	①草本	15 落葉針葉樹林	③高木
2 浮葉植物群落	①草本	16 常緑広葉樹林	③高木
3 塩沼植物群落	①草本	17 常緑針葉樹林	③高木
4 砂丘植物群落	①草本	18 植林地(竹林)	④竹林
5 一年生草本群落	①草本	19 植林地(スギ・ヒノキ)	③高木
6 多年生草本群落	①草本	20 植林地(その他)	③高木
7 単子葉草本群落 ヨシ群落	①草本	21 果樹園	③高木
8 単子葉草本群落 ツルヨシ群落	①草本	22 畑	⑤その他植生
9 単子葉草本群落 オギ群落	①草本	23 水田	⑤その他植生
10 単子葉草本群落 その他	①草本	24 人工草地	⑤その他植生
11 ヤナ半低木林	②低木	25 グラウンド等(ゴルフ場)	⑤その他植生
12 ヤナ半高木林	②高木	26 人工構造物	⑥裸地
13 その他の低木林	②低木	27 自然裸地	⑥裸地
14 落葉広葉樹林	③高木	28 開放水面	⑦水面

(3) 衛星画像

衛星技術は近年大きく進化しており、解像度、撮影頻度が異なる多種多様な衛星が宇宙空間を飛び交っている状態である。本検討では、より撮影頻度が高く、安価に入手可能な光学衛星画像Planet Doveを用いた。Planet Doveの概要及び撮影画像を表-2、図-2に示す。



Includes material c[2024]Planet. All rights reserved.

図-2 衛星画像(Planet Dove)の例

表-2 Planet Doveの概要

使用した衛星	Planet Dove
特徴	130機以上の小型衛星群で全地球を常時撮影
撮影頻度	概ね1日1回 ※光学衛星のため、曇天、夜間の撮影は不可
解像度	約3m
アーカイブ	2017年以降のデータを蓄積 ※新規のリクエスト撮影も可能
価格	約2,000円/km ² ※最低購入価格設定あり。 円安等の社会情勢により多少変化

(4) 特徴量

ランダムフォレストの分類に用いる特徴量は、各衛星画像のRGB情報(可視域赤、緑、青の波長)に加え、NIR(Near-infrared:近赤外線値)、NDVI(Normalized Difference Vegetation Index:正規化植生指標)、NDWI((Normalized Difference Water Index:正規化水指数)、BSI(Bare Soil Index:裸地指数)の計7特徴量を4時期(春夏秋冬)、用意し計28個とする。NDVI、NDWI、BSIは以下に示す数式により算出を行う。

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

$$NDWI = \frac{(Green - NIR)}{(Green + NIR)}$$

$$BSI = \frac{((Red + Blue) - Green)}{((Red + Blue) + Green)}$$

3. 解析

(1) 対象河川

対象河川と自動判読対象区間は、揖保川本川の林田川合流点(5.0k)~8.0kとした。本区間は、河道内の樹木繁茂や堤防植生が顕著であり、イノシシ等の出没が多く確認されている。また、再繁茂抑制対策の取組みや地元住民との地域連携を進めている区間である。

(2) 解析条件

本検討では、対象河川における自動判読技術の適用に加え、現場実装を見据えた省力化の観点から一部区間で教師データを作成し自動判読した場合、判読精度へどのように影響するか検証した。具体的な解析条件を表-3に示す。解析ケースは大きく2ケースを用意し、教師データを全川(本川)と設定したもの(CASE1)、一部区間で設定したもの(CASE2)である。CASE2については、自動判読検証区間と同様の地被状態分類を示す4箇所(clip1~clip4)を選定した(図-3、図-4)。

表-3 解析条件

解析条件	自動判読検証区間	教師データ			モデル精度確認
		作成区間	面積[km ²]	特徴量	
CASE1		全川(本川)	-	RED、GREEN、BLUE、NIR(近赤)	4時期(平成31年度) 正解データ「河川水辺の国勢調査結果」(H31年度)
CASE2	5.0~8.0km 【0.9km ² 】	clip1_1.8~2.8km	0.31	NDVI、NDWI、BSI(計7種)	
		clip2_4.0~5.0km	0.30		
		clip3_9.0~10.0km	0.36		
		clip4_10.0~12.0km	0.56		

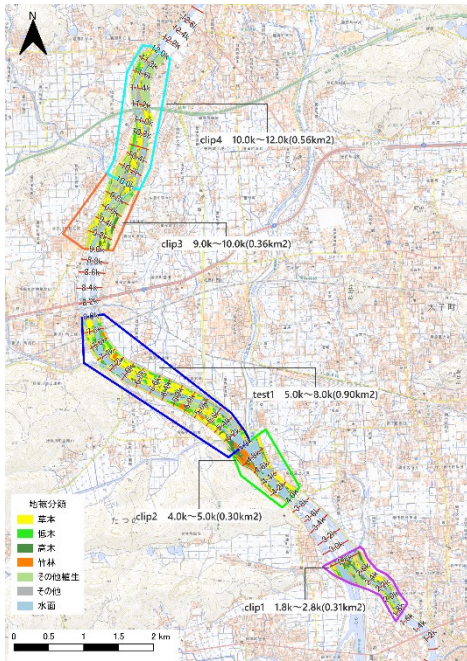


図-3 教師データと自動判読検証区間

表-4 全川と教師データ及び検証範囲の地被分類別面積割合

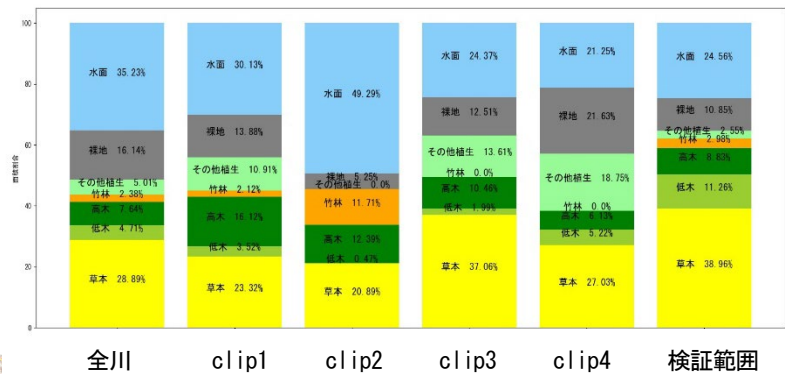


表-5 地被状態の自動判読結果 (CASE1)

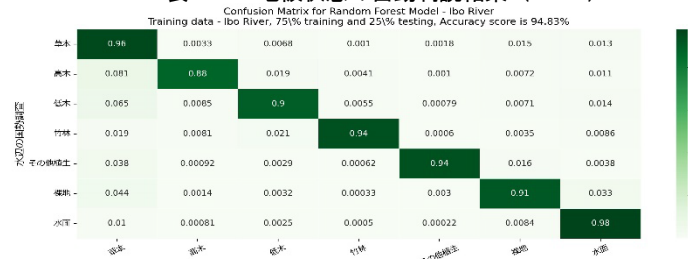


表-6 地被分類ごとの正答率 (CASE2)

地被分類	CASE2: 教師データ (一部区間)			
	clip1 (1.8~2.8km)	clip2 (4.0~5.0km)	clip3 (9.0~10.0km)	clip4 (10.0~12.0km)
草本	0.67	0.62	0.71	0.49
低木	0.12	0.00	0.12	0.49
高木	0.60	0.47	0.57	0.50
竹林	0.22	0.46	0.00	0.00
その他植生	0.09	0.00	0.01	0.24
裸地	0.38	0.51	0.23	0.42
水面	0.75	0.94	0.92	0.82
全体	0.56	0.58	0.59	0.55

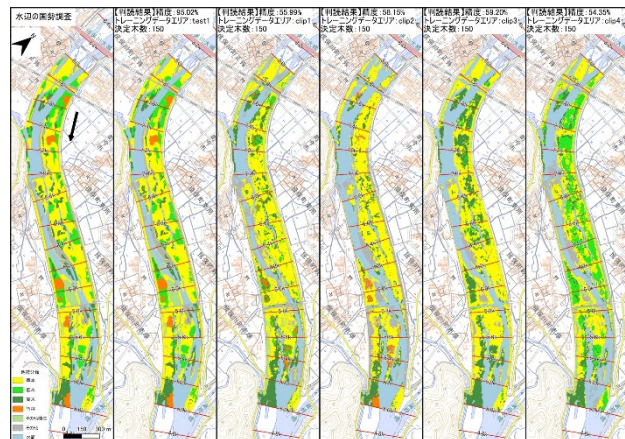


図-4 一部区間教師データによる自動判読結果 (CASE2)

4. 解析結果と課題

1) CASE1 (教師データを全川区間対象)

表-5には、CASE1における解析結果を示す。縦軸に正解データである水辺の国勢調査結果、横軸に自動判読結果を示し、交差箇所を示す数字は正解率を示す。正解率が高いほど(すなわち判読精度が高い)、濃い緑色ハッチングで示す。解析結果より、本解析による判読精度は94.8%と精度よく再現できていることが確認された。地被分類ごとの解析精度に着目すると、水面、裸地は98%、91%と精度高く判読できており、植生域(草本、高木、低木、竹林、その他植生)はもっとも精度が悪い高木でも88%、植生域では平均92%と精度高く判読できることが確認された。なお、誤判読(判読結果と正解データが異なる)については、植生域同士や水面との誤判読が確認されているが、これは、水面から近い植生域が多くを占め、反対に水面より離れて繁茂している高木や竹林との明らかな誤判読はほとんど確認されない。

2) CASE2 (教師データを一部区間対象)

表-6には、CASE2における各地被分類ごとの正答率を、図-4には正解データと各解析結果の地被分類結果を示す。解析結果より、選定した一部区間(clip1~clip4)における解析精度は、約55%~59%程度であることが確認された。地被分類ごとの解析精度に着目すると、草本、高木、裸地、水面は比較的高い判読精度はあるが、低木、竹林、その他植生の地被分類の判読正答率は低い。表-4より、検証範囲には竹林が含まれるため、教師データに竹林が含まれるclip1及びclip2は竹林の判読精度が比較的高く、竹林を含まないclip3及びclip4については竹林の判読精度が低いことが確認される。このため、教師データに含まれない地被分類がある場合は、判読が困難である。

5. 課題への対応

前述の地被分類(表-4)からも確認されるように、河道内植生の地被分類割合は水面、裸地、草本が大部分を占める一方で、河道内樹木のモニタリングに必要な低木、高木、竹林等は割合が低く、特に竹林は揖保川全域のデータにおいても約2%程度となっている。今回用いたランダムフォレストは、正解データの地被分類に合うようモデルパラメータを自動的に最適化する。そのため、面積割合を大きく有する(サンプルが多い)地被分類が重視され、面積割合が小さい(サンプル数が少ない)地被分類の判読精度が低下する傾向がある。

そこで本検討では、ランダムフォレストの課題であるサンプル数の大小について、データ補完手法であるオーバーサンプリング手法を用いて判読精度への向上に寄与するか検証を行う。オーバーサンプリング手法とは、サンプル数データに対し、疑似的にデータを増やす手法でありデータ数が少ない一部区間におけるデータ補完手法として検証を実施する。

(1) 解析条件

本検討では、オーバーサンプリング手法である、SMOTE³⁾(Synthetic Minority Over-sampling Technique)とADASYN⁴⁾(Adaptive Synthetic sampling)の2つの手法を用いた。本論文では、これら2手法の詳細については割愛させて頂くが、SMOTEは、少数派データの中で近傍に位置するデータの中からランダムに選定したデータとの間に内挿的にデータを作成することでデータ数を増やす方法である。一方で、ADASYNはSMOTEの考え方をベースとした拡張版と考えることができ、少数派データと多数派データの情報から、少数派と多数派の境界付近に重点的にデータを増やす方法である。いずれのオーバーサンプリング方法もpythonのライブラリであるimbalanced-learn v0.12.0⁵⁾を用いた。

なお、オーバーサンプリングのパラメータについては、判読精度が最もよくなるようモデルパラメータを最適化した。

(2) 解析結果

表-7に、オーバーサンプリング手法を用いた地被分類ごとの正答率を、図-6には正解データと各判読結果の地被分類結果を示す。判読精度は、一部55~60%程度と、若干の上昇が確認され、clip3を教師データに用いた場合では、精度の向上が確認された。これは、データ補完により少数派データであった低木のデータ数が増えたことで低木と判読する結果が増え、正解データに近くなっていることが確認できる(図-5)。また、地被分類ごとに河川水辺の国勢調査と比較すると、草本、高木、裸地、水面は比較的高い判読精度はあるが、低木、竹林、その他植生の地被分類の判読正答率は低い。ただしclip2の竹林の判読率が高いことから、教師デ

表-7 データ補完ありの地被分類ごとの正答率

地被分類	CASE2: 教師データ(一部区間) オーバーサンプリング手法適用			
	clip1 (1.8~2.8km)	clip2 (4.0~5.0km)	clip3 (9.0~10.0km)	clip4 (10.0~12.0km)
草本	0.67	0.66	0.74	0.65
低木	0.15	0.00	0.13	0.15
高木	0.60	0.47	0.55	0.60
竹林	0.02	0.50	0.00	0.02
その他植生	0.09	0.00	0.00	0.08
裸地	0.36	0.46	0.24	0.36
水面	0.75	0.94	0.91	0.75
全体	0.56	0.59	0.60	0.55

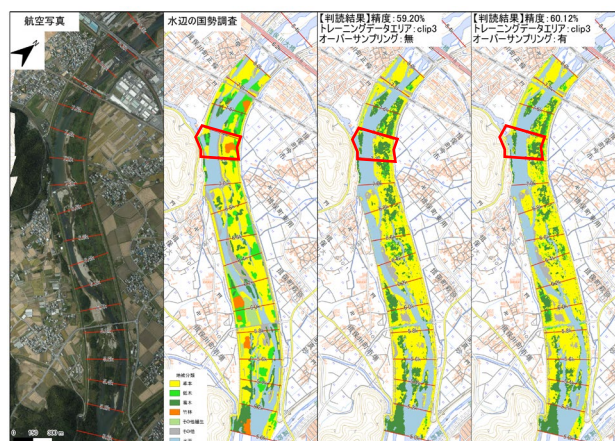


図-5 データ補完あり・なしによる自動判読結果(clip3)

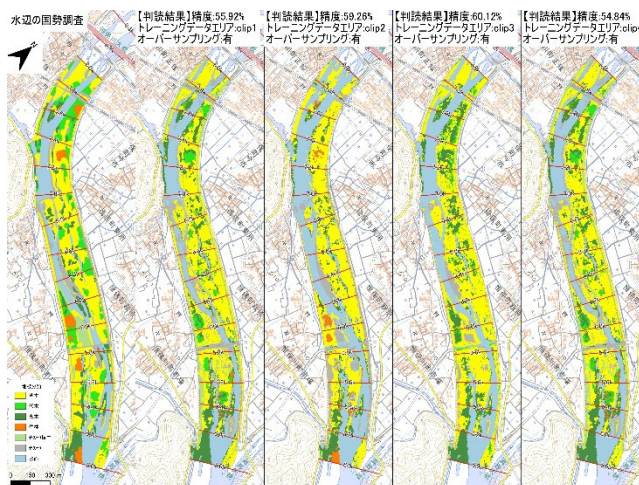


図-6 データ補完ありによる自動判読結果

ータに選定には留意が必要でありその手法については今後の課題である。

一方で、航空写真と比較すると右岸高木敷上の樹木密生状況と自動判読結果の差異は小さく、流下能力の算出に用いる樹木繁茂範囲の把握は可能であるように判断できる。したがって、求める自動判読の精度については流下能力の算定にあたり実施する水位計算の結果に大きく影響が出ないこと、また河道掘削や樹木伐採等の河川工事による河道変化をモニタリングできることが重要であり、この点については今後検証を行っていく。

6. 河道管理の高度化・省力化への可能性

(1) 河道内モニタリングへの活用

河川水辺の国勢調査は、5年に1度の頻度で実施されているが、本技術は衛星画像の購入のみで河道内のモニタリングが可能となるため、年に1度のように高頻度な状況把握の実施が可能となる。図-7には、河川整備事業による河道掘削前後の堆積土砂の範囲を本手法によってモニタリングした事例を示す。図より、施工前後で堆積土砂の範囲を自動判読結果からGIS等で堆積範囲を計測することで、堆積の進行性を定量的に把握することが可能である。これにより、モニタリング箇所において堆積の進行性を踏まえた形で記録することが可能であり、河道の変遷を踏まえた植生侵入などが想定される旨も合わせて記載することで、河道管理における留意事項を共有することが可能となる。

(2) 河道点検への活用

近年、河道点検の重要性は認識されつつある中で、河道点検は堤防や河川管理施設と異なり出水により変化が多いことが特徴である。その一方で、河道点検記録のほとんどは目視調査による定点観察が多く記録されており、橋梁等からの定点観察のため定性的な記録のみで、定量的な記録はほとんど確認されてない。よって、目視による定点観察では土砂堆積・樹木繁茂状況の進行性を把握することは難しいことが問題である。そこで、図-8に示すように衛星画像を活用した自動判読技術を活用することで、現地作業を省力化かつ定量的に把握することが期待される。

7. 結論

本検討では、揖保川本川における河道内の樹木繁茂が顕著な区間(計3.0k)を対象に、光学衛星画像を用いた教師あり機械学習による河道内の地べ分類の自動判読を試行した。また光学衛星画像を用いた今後の河道管理への高度化・省力化の可能性を検討した。得られた知見は以下の通りである。

- 1) 教師データを揖保川全川(Case1)とした場合、約9割以上の精度で自動判読可能であること。
- 2) 教師データを一部区間(Case2)とした場合、解析上の課題としてサンプル数不足が挙げられたため、データ補完手法であるオーバーサンプリング手法を適用した結果、若干であるが判読精度の向上が確認された。
- 3) 衛星画像による自動判読技術により、河道掘削前後の堆積土砂の範囲を把握することで掘削後のモニタリングについて省力化を図れる可能性がある。流下能力の維持管理の観点において自動判読結果を用いて、GIS上で定量的に河道の変化を把握することで、河道管理の高度化・省力化の可能性はある。今後は使用目的に応じた自動判読精度の向上について検証を進めていく。

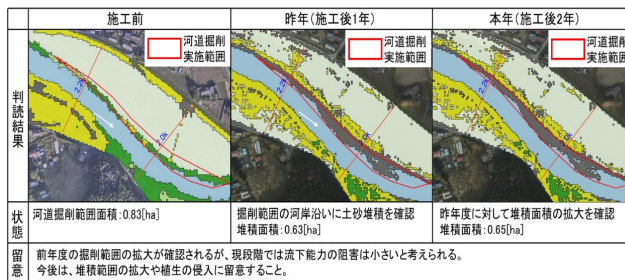


図-7 土砂掘削後のモニタリングの事例

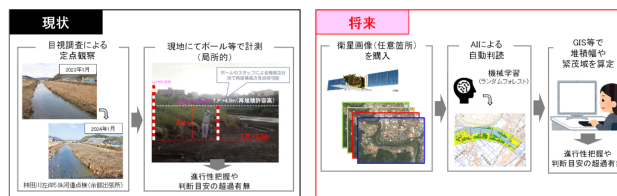
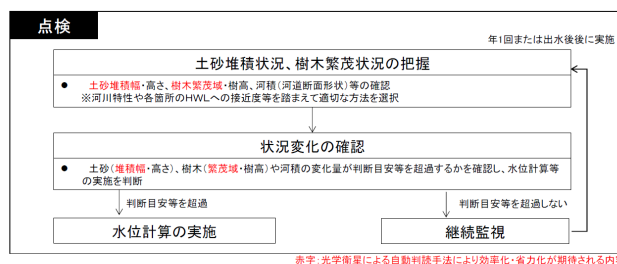


図-8 光学衛星画像を活用した河道点検への適用(案)

8. 今後の展望

本手法では、衛星画像の購入と解析費用のみで地べ分類の判読が可能となり、判読素図の作成や現地調査、地べ分類等の技術者による作業に係る費用が大きく縮減可能であると考えられる。今後は、河道点検に係る現場作業の省力化を見据え、自動判読精度の向上に努めていく。

謝辞

本検討にあたり国土交通省近畿地方整備局姫路河川国道事務所より、最新年度の河川水辺の国勢調査結果を頂きました。上記して、謝意を表します。

9. 参考文献

- 1) 「堤防等河川管理施設及び河道の点検・評価要領,国土交通省,平成31年4月,pp. 39」
- 2) 西俣ら,衛星画像データを用いた機械学習による河道内地べ分類の自動判読技術の開発,令和5年度建設コンサルタント業務研究発表会
- 3) Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic over-sampling technique. Journal of artificial intelligence research, 16, pp. 321-357
- 4) Han, H., Wang, W. Y., & Mao, B. H. (2005). Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In International conference on intelligent computing (pp. 878-887). Springer, Berlin, Heidelberg.
- 5) imbalanced-learn(<https://imbalanced-learn.org/stable/>)