

GEOMORPHOLOGICAL MAPPING OF INTERTIDAL AREAS

M. Lu^{1*}, L. Groeneveld¹, D. Karssenberg¹, S. Ji³, R. Jentink², E. Paree², E. Addink¹

¹ Department of Physical Geography, Utrecht University, the Netherlands lumeng0312@gmail.com

² Rijkswaterstaat CIV, the Netherlands

³ Department of remote sensing and information engineering, Wuhan University, China.

Commission III, WG III/1

KEY WORDS: Geomorphology, Inter-tidal flats, semantic segmentation, multi-dimensional, feature interpretation, OBIA, XG-Boost

ABSTRACT:

Spatiotemporal geomorphological mapping of intertidal areas is essential for understanding system dynamics and provides information for ecological conservation and management. Mapping the geomorphology of intertidal areas is very challenging mainly because spectral differences are oftentimes relatively small while transitions between geomorphological units are oftentimes gradual. Also, the intertidal areas are highly dynamic. Considerable challenges are to distinguish between different types of tidal flats, specifically, low and high dynamic shoal flats, sandy and silty low dynamic flats, and mega-ripple areas. In this study, we harness machine learning methods and compare between machine learning methods using features calculated in classical Object-Based Image Analysis (OBIA) vs. end-to-end deep convolutional neural networks that derive features directly from imagery, in automated geomorphological mapping. This study expects to gain us an in-depth understanding of features that contribute to tidal area classification and greatly improve the automation and prediction accuracy. We emphasise model interpretability and knowledge mining. By comparing and combining object-based and deep learning-based models, this study contributes to the development and integration of both methodology domains for semantic segmentation.

1. INTRODUCTION

1.1 Background

Geomorphological mapping of intertidal areas in space and time is essential for understanding system dynamics and provides information for ecological conservation and management (Bouma et al., 2005). Ecological quality of intertidal areas is important because of the European Water Framework Directive and as they are designated as Natura 2000 areas, which is the implementation of the European bird directive and the European habitat directive. Mapping the geomorphology of intertidal areas is a considerable challenge mainly because spectral differences are oftentimes relatively small while transitions between geomorphological units are oftentimes gradual. Also, the intertidal areas are highly dynamic (Kleinhans et al., 2019). Surface water, saltmarsh, and tidal flats are relatively simple to distinguish but considerable challenges remain for distinguishing between different types of tidal flats, specifically, low and high dynamic shoal flats, sandy and silty low dynamic flats, and mega-ripple areas (Douma et al., 2019). In this study, we harness machine learning methods in automated geomorphological mapping and extend object-based methods with deep neural network-based semantic segmentation methods. We focus on distinguishing between 5 classes: sandy low dynamic flats, silty low dynamic flats, mega-ripples, high dynamic shoal flats, and hard substrates. We evaluate the methods used with an extensive data set of visually interpreted photos.

1.2 Objective

The objectives of this study is to compare between the methods that are based on ensemble tree-based modeling on features of

segments from classical OBIA vs. deep neural networks-based methods.

1.3 Impact

This study is both application- and method-driven. It provides us with an in-depth understanding of features that contribute to tidal area classification and is expected to greatly improve the automation of the procedure and the prediction accuracy. We emphasise model interpretability and knowledge mining. By comparing and combining object-based and deep learning-based models, this study contributes to the integration of thus far largely separate approaches for semantic segmentation.

2. METHODOLOGY

The study is mainly developed in Western Scheldt, the Netherlands (fig. 1), where a detailed manually delineated expert classification map is available. Figure 1 (b) shows our classification scheme. Importantly, we distinguish between the sandy and silty areas within the shoal flat low dynamics. We used aerial imagery with red, green, and NIR bands (0.25 m resolution), DEM from laser altimeters (2 m resolution), and derived indices. Imagery and DEM are acquired within the same year and season. Two methods are developed and compared in terms of the priors integrated, model interpretability, and the prediction accuracy and patterns. The first method firstly applies OBIA (Object-based Image Analysis) for segmentation and then uses an ensemble tree-based method, XGBoost (Chen and Guestrin, 2016), for classification, this method is referred to (OBIA-XGB) in this study. The challenges of this method lie in the identification of the optimal spatial unit of objects. The second method applies an encoder-decoder deep neural network archi-

* Corresponding author

lectures, U-net (Ronneberger et al., 2015), for end-to-end classification.

2.1 Preprocessing

We focus on distinguishing between 5 classes and for convenience, we assign a code to each of them:

- P1a1: sandy low dynamic flats
- P1a2: silty low dynamic flats
- P2b: mega-ripples
- P2c: high dynamic shoal flats
- H1: hard substrates

For the definition of the classes please refer to Douma et al. (2019). We also refer to the P1a1, P1a2, P2b, and P2c classes as the "P" class and the H1 class "H" class. The other classes (i.e. classes that do not belong to the P and H1 classes) are removed.

We filtered out water by firstly using eCognition (Developer 9.4.0) to do segmentation, and then using the NDWI ($NDWI = \frac{green - NIR}{green + NIR}$) with a threshold of 0.44. Mega-ripples naturally consist of water. In this study, we also filtered out water in the P2b (mega-ripple) class to gain an initial understanding of the behaviour of machine learning models.

2.2 Sampling scheme

As the unit of method OBIA-XGB is OBIA segments and method U-net image pixels, we develop different sampling schemes for each of them. For method OBIA-XGB, it is necessary to develop a sampling scheme that accounts for the object size for the tree classifier. We evaluated two sampling schemes: 1) random sampling and 2) stratified sampling according to the size of OBIA segments. For 1), we used random under-sampling to sample the majority class(es), without replacement, to ensure the classes have balanced samples. For 2), we firstly divided the objects into 5 categories, cut at each 20th, 40th, 60th, 80th, 100th percentiles of the object size of the minority class (i.e. class with the least number of objects). These percentiles are used so that the number of objects in these categories are similar. Within each category, 70% of the objects form the training set and the rest forms the test set. Then within the training set of each category, random under-sampling is applied.

For OBIA-XGB, we selected three tiles that are abundantly covered by the focused classes (i.e. P and H1 classes) for identifying the optimum spatial unit and model training. For the method based on U-net, we selected 80 500 pixels \times 500 pixels tiles for training (42 tiles) and validation (18 tiles). Three 4000 \times 4000 tiles are used for testing both of the methods and calculating the accuracy metrics. The train-test splitting is shown in fig. 2.

2.3 Accuracy assessment

We focus on the precision (True Positives)/(True Positives + False Positives) and recall (True Positives)/(True Positives + True Negatives) as indicators of the prediction accuracy. For OBIA-XGB, the precision and recall are weighted by the area size of each object for the final recall and precision. In addition, we present the precision and recall based on objects (i.e. not weighted by the area size). This serves to inspect how the objects with various sizes are identified.

3. EXPERIMENTS

3.1 Method OBIA-XGB

The inputs of the OBIA-XGB method are aerial imagery (with NIR, red, green bands) and DEM, as well as slope (Zevenbergen and Thorne, 1987), NDVI, and aspect ((Horn, 1981)) derived from eCognition.

We firstly use eCognition for object segmentation at various spatial units. Then, we combine segmentation levels to get sub-objects and related features. We export the objects with in total 48 features describing the spectral and texture (using Harralick's grey level co-occurrence matrices) features of objects and sub-objects.

Then, to determine the optimal spatial unit for each class, we each time regroup the five classes into the "target class" and the "other class" and iterate over all the five classes. The XGBoost is applied to classify the OBIA calculated object features and select the best spatial unit for each class based on the prediction accuracy.

We found no considerable differences in the optimal spatial unit for each class (will be described in section 4.1, fig. 3), therefore, we used a single optimal spatial unit for all classes. The identified optimal spatial unit is used for multi-class classification. The loss functions for the binary and multi-class classifications are respectively the *logistic loss* and *softmax* loss (i.e. a softmax activation followed by a cross-entropy loss).

The hyperparameters we tuned for XGBoost are learning rate, maximum tree-depth, number of estimators, and the Lasso regularisation term α , using 5-fold cross-validation. The modeling process using XGBoost and deep neural networks is implemented in the Python environment (version 3.6).

3.2 Method based on U-net

The input for the U-net is the aerial imagery, DEM, and derived slope (Horn, 1981), NDVI, and Brightness layers. The slope was calculated in QGIS3.18 using the same Horn's formula (Horn, 1981) as in eCognition, and is expressed in degrees. The NDVI and brightness are calculated in Python. The Brightness is defined as $Brightness = (NIR + red + green)/3$.

The loss functions we tested are categorical cross-entropy loss, IoU loss (van Beers et al., 2019), FocalTverskyLoss (Abraham and Khan, 2019), and Lovasz Hinge Loss (Berman et al., 2018). In the end, we chose to use the IoU loss as it provides the best cross-validation results with the training set. The hyperparameters we tuned are learning rate, batch size, and image size. The image size tested are from 2^5 (64) to 2^9 (512) and is optimised at 128. The batch size tested are from 8 to 32 and is optimised at 16. The learning rate are tested with various scheduling, ranging from 0.0001 to 0.01 at various epochs.

4. RESULTS

4.1 Spatial unit optimisation for OBIA-XGB

Figure 3 shows the precision and recall of different sampling schemes and cross-validation methods for the binary classification in the Western Scheldt study area. The stratified sampling greatly improves the prediction accuracy, comparing

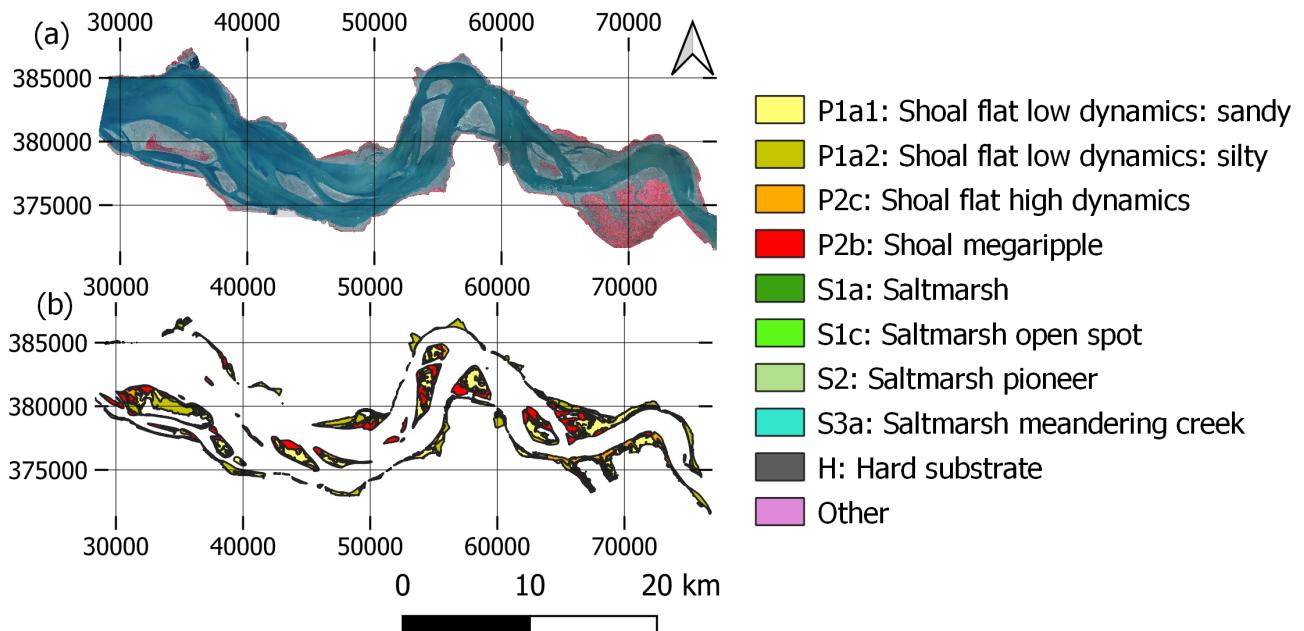


Figure 1. The Western Scheldt study area. Figure (a): false color aerial image with the red, blue, and Near-Infrared bands, 0.25 m resolution. Figure (b): the expert map and classification schemes.

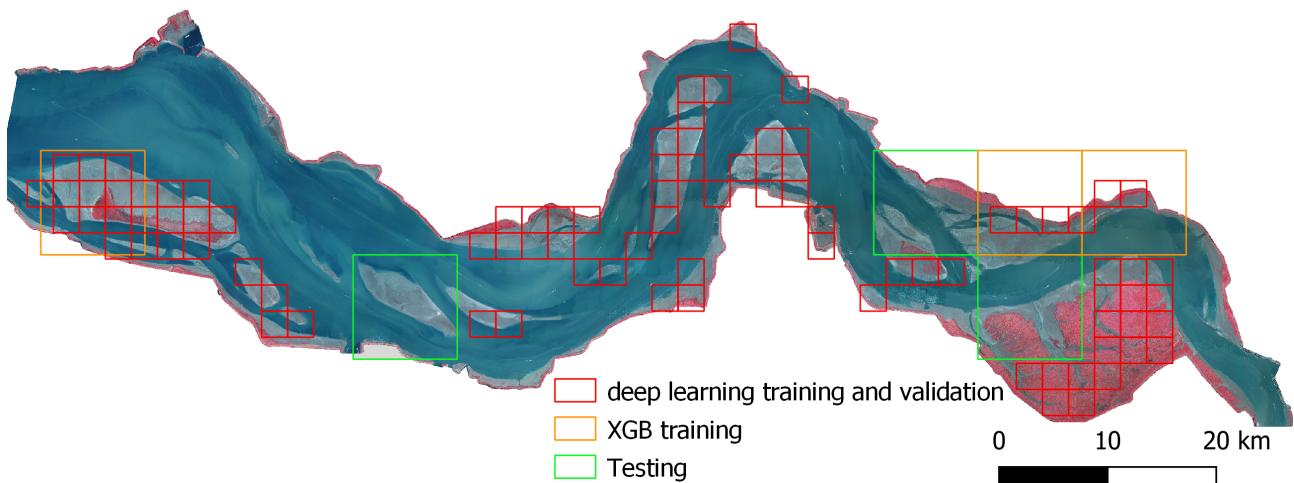


Figure 2. Training, validation, and test areas used by the machine learning methods, Western Scheldt.

the cross validation between applying our Stratified Sampling (SS) scheme to random sampling of the objects (object).

It could be observed that when using both stratified sampling and area-weighted cross-validation (SSAC), all the classes obtained a very high precision with each of the tested spatial unit, indicating that larger objects are better identified than the smaller objects. The recall remains high for the H class but is lower for the P classes, especially the P1a1. Compared to SS, which has a relatively high recall for all the classes but much lower and diverse precision for different spatial units and classes, it is found that objects with smaller sizes are more likely to mix with other classes, especially for the P1a1, P2c, and H class. As the SSAC shows homogeneous recall and precision between spatial units, to optimise the classification for small objects and ensure the highest model generality, we use SS to identify the optimal spatial unit. The optimal spatial unit identified for all classes are 50_15 (object level 50, sub-object level 15).

4.2 Prediction accuracy

4.2.1 Method OBIA-XGB The accuracy metrics of the OBIA-XGB are shown in tables 1 and 2. SSAC is used to indicate the performance and SS for indicating how well the smaller objects are classified. According to SSAC, the megaripple class obtained the highest precision, indicating the method OBIA-XGB is already promising in distinguishing it from other P classes and the H1 class. The sandy and silty low dynamic flats are not well separated.

Comparing tables 1 and 2, it is observed that small objects better in the P2b and H1 classes but are less satisfying in other classes. From the spatial prediction maps, We can observe that mis-classification mostly occur as classifying P1a1, P1a2 and P2b as H1, especially along some channel outlets or close to water; and mis-classify P1a1 into P1a2 or p2b. The probability map also shows a high uncertainty (indicated as low probability) in separating between P1a1 and P1a2, as well as H1 and



Figure 3. Cross-Validation results for all the classes, with the XGBoost classifier on OBIA objects. Objects: random sampling and cross-validation based on objects. SSAC: stratified-area sampling with area-weighted cross-validation results. SS: stratified-area sampling.

	P1a1	P1a2	P2b	P2c	H1
P1a1	7.0	4.2	1.2	9.1	0.1
P1a2	8.0	3.3	4.9	2.6	1.1
P2b	5.2	2.0	24.2	8.0	0.8
P2c	1.4	3.9	1.0	11.5	0.1
H1	0.0	0.0	0.1	0.1	0.0
precision	32.4	24.4	77.2	36.8	1.3
recall	32.6	16.5	60.2	64.1	8.9
overall accuracy	46.1				

	P1a1	P1a2	P2b	P2c	H1
P1a1	4.2	4.4	1.4	5.4	0.1
P1a2	5.4	2.7	9.2	2.3	1.6
P2b	3.1	2.6	36.6	6.4	1.6
P2c	0.8	2.9	1.3	7.3	0.2
H1	0.0	0.0	0.2	0.2	0.1
precision	30.7	21.6	75.1	33.8	1.9
recall	26.9	12.8	72.7	58.5	14.0
overall accuracy	50.8				

Table 1. Accuracy assessed with the sampling and validation scheme SSAC, on test tiles. Confusion matrix is normalised by the entire area. Precision, recall, and overall accuracy (%) are calculated using area-weighted accuracy assessment.

other classes.

4.2.2 Method based on U-net The U-net segmentation obtained a considerably lower accuracy for P2b and P2c compared to the method OBIA-XGB (tables 1 and 3) but obtained a much better result for P1a1. P2b again obtained the best result compared to the other classes but is often confused with P1a1. The model failed completely to predict the H1 class.

5. CONCLUSION

In this study, we compared machine learning methods that are based on OBIA-derived features vs. end-to-end representative learning, for inter-tidal area classification, with respectively XGBoost and U-net. The method that applies XGBoost

Table 2. Accuracy assessed with the sampling and validation scheme SS. Confusion matrix normalised by the number of total objects. Precision, recall, and overall accuracy (%) are calculated using area-weighted accuracy assessment.

	P1a1	P1a2	P2b	P2c	H1
P1a1	17.0	4.4	3.6	5.8	0.0
P1a2	5.8	1.8	0.5	1.2	0.0
P2b	6.3	2.2	19.5	8.9	0.0
P2c	7.5	3.6	5.3	6.3	0.0
H1	0.0	0.0	0.1	0.1	0.0
precision	46.3	14.7	67.0	28.2	0.0
recall	55.0	19.1	52.8	27.8	0.0
overall accuracy	45.0				

Table 3. Confusion matrix normalised by the number of pixels, precision, recall, and overall accuracy (%) of the U-net model.

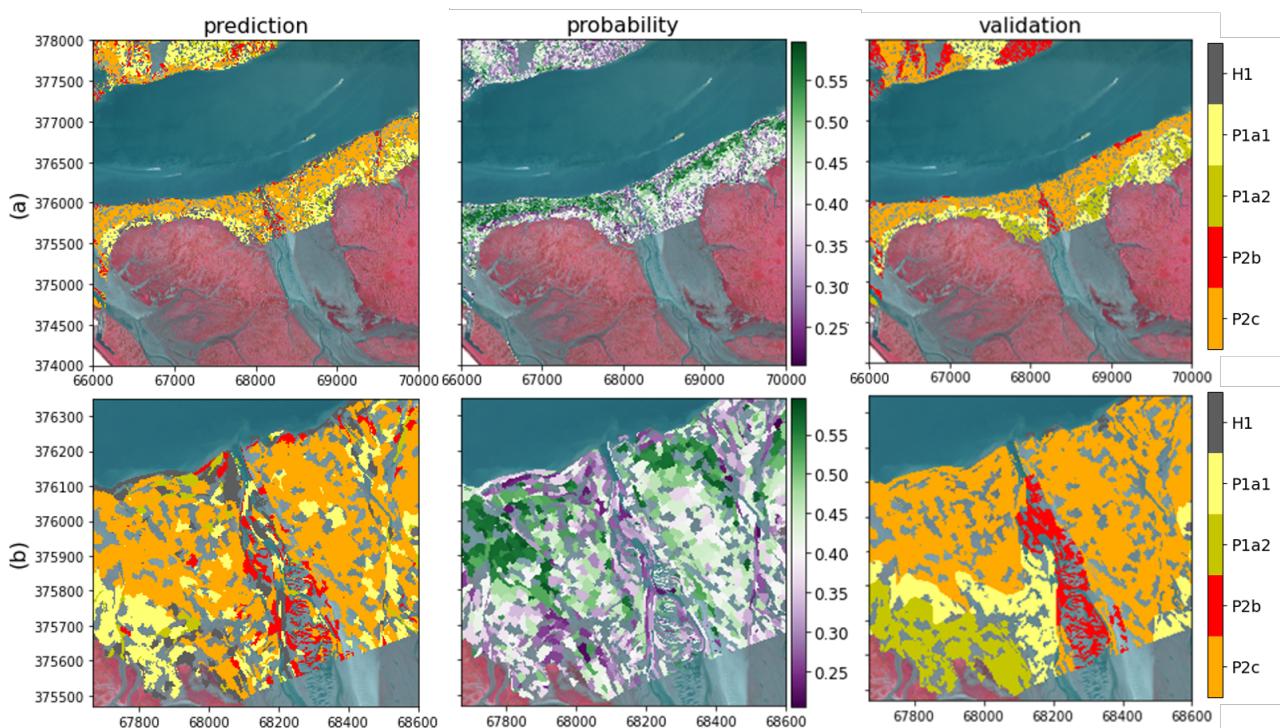


Figure 4. Maps of method OBIA-XGB. Figure (b) is a zoom-in of the figure (a). From left to right: predicted labels, predicted probabilities of the class with highest value, ground truth (labels).

to OBIA-derived features (OBIA-XGB) outperforms the U-net based methods in P2b and P2c, but the U-net method can better separate between P1a1 and P1a2. The OBIA-XGB obtained satisfying results in separating between the P2b and other classes, which is sometimes more difficult to distinct manually or basing on the rule-based OBIA methods alone. Therefore, our study indicates that machine learning methods can help to improve the rule-based OBIA method. Future studies aim at improve over the basic U-net method we applied in this study.

6. ACKNOWLEDGEMENT

This study was performed at Utrecht University with a contribution from the project 'Optimalisatie productie GMK met behulp van OBIA' financed by Rijkswaterstaat (Contract number: 31160225).

References

- Abraham, N., Khan, N. M., 2019. A novel focal tversky loss function with improved attention u-net for lesion segmentation. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, IEEE, 683–687.
- Berman, M., Triki, A. R., Blaschko, M. B., 2018. The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4413–4421.
- Bouma, H., De Jong, D., Twisk, F., Wolfstein, K., 2005. Zoute wateren ecopenstelsel (zes. 1). *Voor het in kaart brengen van het potentiële voorkomen van levensgemeenschappen in zoute en brakke rijkswateren. Rapport RIKZ/2005.024. Rijkswaterstaat, Middelburg*.
- Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, ACM, New York, NY, USA, 785–794.
- Douma, H., Addink, E. A., Kleinhans, M. G., 2019. Nadere uitwerking productie Geomorfologische Kaart met behulp van Object-Based Image Analysis voor de Waddenze. *Universiteit Utrecht, Departement Fysische Geografie, rapport i.o. Rijkswaterstaat CIV*.
- Horn, B. K., 1981. Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1), 14–47.
- Kleinhans, M. G., Douma, H., Addink, E. A., Jentink, R., 2019. Tidal flats, megaripples and marsh: automated recognition on aerial images. *MARID VI*, 135.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.
- van Beers, F., Lindström, A., Okafor, E., Wiering, M. A., 2019. Deep neural networks with intersection over union loss for binary image segmentation. *ICPRAM*, 438–445.
- Zevenbergen, L. W., Thorne, C. R., 1987. Quantitative analysis of land surface topography. *Earth surface processes and landforms*, 12(1), 47–56.

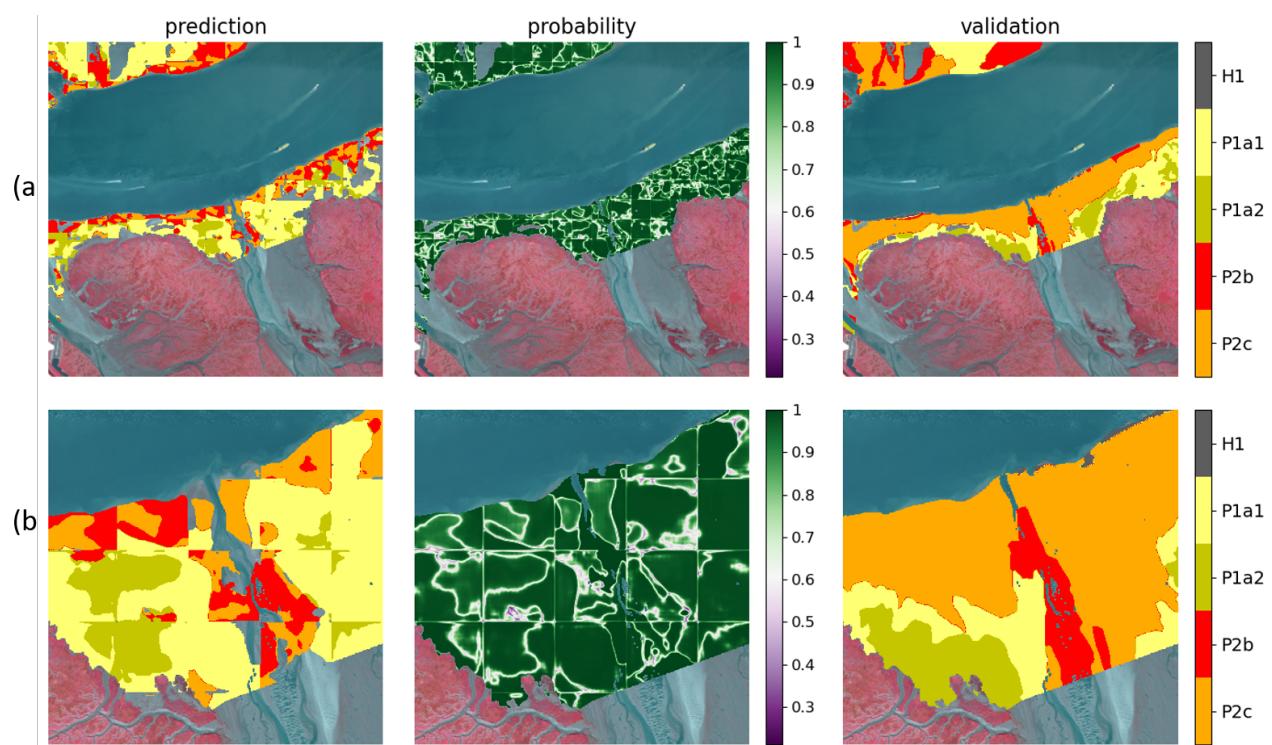


Figure 5. Map of the best U-net model. Figure (b) is a zoom-in of the figure (a). From left to right: predicted labels, predicted probabilities of the class with highest value, ground truth (labels).