


Please cite the Published Version

Zhang, X , Du, L, Tan, S, McCarty, GW and Zou, Z (2024) Wetland classification based on depth-adaptive convolutional neural networks using leaf-off SAR imagery. *Science of the Total Environment*, 957. 177768 ISSN 0048-9697

DOI: <https://doi.org/10.1016/j.scitotenv.2024.177768>

Publisher: Elsevier BV

Version: Accepted Version

Downloaded from: <https://e-space.mmu.ac.uk/637604/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: This is an author-produced version of the published paper. Uploaded in accordance with the University's Research Publications Policy

Data Access Statement: The data and codes that support the findings of this study are available from the corresponding author, LD, upon reasonable request.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

1 **Types of article:** Research paper

2

3 **Wetland classification based on depth-adaptive convolutional neural networks**
4 **using leaf-off SAR imagery**

5 Xin Zhang ^{a, †}, Ling Du ^{b, c, †, *}, Shen Tan ^{d, e}, Gregory W. McCarty ^c, Zhenhua Zou
6 ^f,

7 *^a Department of Computing and Mathematics, Manchester Metropolitan University, Manchester*
8 *M1 5GD UK; x.zhang@mmu.ac.uk*

9 *^b Department of Environmental Science & Technology, University of Maryland, College Park,*
10 *MD 20742, USA; lingdu@umd.edu*

11 *^c Hydrology and Remote Sensing Laboratory, USDA-ARS, Beltsville, MD 20705, USA;*
12 *greg.mccarty@usda.gov*

13 *^d Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences,*
14 *Peking University, Beijing 100871, China; tanshen@pku.edu.cn*

15 *^e Department of Earth System Science, Tsinghua University, Beijing 100084, China;*

16 *^f Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA;*
17 *zhzhou@umd.edu*

18

19 **† Xin Zhang and Ling Du co-led this manuscript.**

20 *** Correspondence: Ling Du, email: lingdu@umd.edu.**

21

22

23 **Abstract**

24 The recent development of deep learning (DL) techniques has created opportunities for
25 classifying wetlands from remote sensing data (mainly optical data). However, the methods
26 for accurately and efficiently classifying large-scale wetlands using DL and radar data that
27 can be more effective than optical data still needs evaluation. In this study, we developed
28 an end-to-end depth-adaptive convolutional neural network (CNN) for mapping wetlands
29 using leaf-off time-series Sentinel-1 Synthetic Aperture Radar (SAR) imagery along with
30 ancillary data. We examined the inclusion of multi-land cover proximity information and a
31 CNN-based self-supervised SAR denoising procedure for enhancing wetland classification
32 accuracy. The depth-adaptive CNN based on U-Net architecture was designed to classify
33 wetland classes (emergent wetland, scrub-shrub wetland, forested wetland, and open water)
34 in Delaware, U.S. while achieving optimization between model complexity (network
35 depths) and accuracy. Results show that our proposed DL method (OA=0.93, MIoU=0.60)
36 not only produced a higher classification accuracy than the traditional RF method (OA =
37 0.89, MIoU=0.18) but also had a significantly reduced computational cost compared to
38 established state-of-the-art CNNs (e.g., DeepLabV3+ and DANet) without loss of
39 accuracy. The inclusion of multi-land cover proximity information (especially distances to
40 forest and water) and the CNN-based self-supervised SAR denoising procedure can both
41 enhance wetland classification accuracy, especially for forested wetland using traditional
42 RF methods. These results demonstrated the novelty and efficiency of our proposed DL
43 method for classifying wetlands by combining denoised SAR imagery and ancillary
44 information, which provides insights on integration of DL approach and radar data for
45 supporting operational wetland mapping at large spatial scales.

46 **Keywords:** wetland classification, deep learning, leaf-off SAR, proximity information,
47 SAR denoising

48 **1. Introduction**

49 Wetlands cover only 5-8% of the terrestrial land surface but provide essential ecosystem services
50 to human society, such as water storage, flood regulation, and mitigation of climate change (Junk
51 et al. 2012; Mitsch et al. 2012). They also provide habitats for various plants and animals (Cohen
52 et al. 2016). However, wetlands can be extremely difficult to map compared to other permanent
53 or open-surface water wetlands, due to the interplay among water, soils, and vegetation (Gallant
54 2015). This may be especially true for extensive wetlands in the eastern U.S. where they are
55 heavily vegetated (i.e., from sparse emergent herbaceous species to dense woody plants) with
56 significant seasonality (Tiner 2003). These wetlands have been subject to loss in recent decades
57 due to drainage and conversion for large scale agriculture and development, which strongly
58 influences hydrological and biochemical cycles at watershed scale (Lang et al. 2024). Accurate
59 and efficient wetland mapping approaches are critical for quantifying wetland changes due to
60 climate change or human activity and assessing their impacts on regional hydrological and
61 biochemical cycles in earth system modelling.

62 Synthetic Aperture Radar (SAR) can provide observation under most weather conditions
63 and can penetrate through vegetation canopy to some extent, making it a promising data source
64 for large-scale wetland mapping (Adeli et al. 2020; Lang et al. 2008; Li et al. 2014;
65 Mohammadimanesh et al. 2019; Scepanovic et al. 2021). As reported by many studies, with the
66 presence of water underneath vegetation, like-polarized SAR backscatter (i.e., HH and VV) can
67 significantly increase due to the double-bounce interaction between the water surface and
68 vertical structures of the vegetation, providing a useful tool for distinguishing inundated
69 vegetation, especially during the leaf-off season (Henderson and Lewis 2008; Hess et al. 1990;
70 Lang and Kasischke 2008). Cross-polarized channels (i.e., HV and VH) are suitable to describe

71 variations in volume scattering from vegetation, allowing for discriminating different vegetation
72 structures (Baghdadi et al. 2001; Henderson and Lewis 2008). However, SAR data reprocessing
73 is notably more computation intensive than optical image processing due to speckle noise
74 originated from coherent imaging systems. Neglecting speckle noise degrades the radiometric
75 quality of the image and thus hinders image segmentation and classification. Many traditional
76 denoising methods such as filter-based methods (both in spatial and transform domains) have
77 been performed on SAR images as one of preprocessing steps (Argenti et al. 2013; Jamali et al.
78 2021a; Mahdavi et al. 2017). However, these methods usually require a noise-free image for
79 training and usually introduce a “wash out” effect that substantially decreases spatial detail
80 (Frost et al. 1982). Recently, self-supervised denoising methods that do not require clean images
81 have been demonstrated effective for SAR image denoising in terms of noise reduction and fine
82 feature preservation (Lin et al. 2023; Tan et al. 2022). The effectiveness of these advanced SAR
83 denoising procedures in wetland classification could be further examined.

84 Ancillary spatial data layers that provide descriptive information such as topographic and
85 proximity/adjacency characteristics can enhance wetland classification. The topographic
86 information such as light detection and ranging (LiDAR) derived topographic metrics has been
87 demonstrated to slightly or significantly improve wetland classification (Du et al. 2020; Hogg
88 and Holland 2008; Lang et al. 2012; O'Neil et al. 2018). Additionally, wetlands have both
89 hydrological and biotic characteristics that connect with their surroundings. It has been reported
90 that with the distance to the nearest stream or water body decreasing, the proportion of wetlands
91 increased significantly, and the geographic proximity to water represented an essential data layer
92 in wetland and land cover classification (Clewley et al. 2015; Hermosilla et al. 2022; Whitcomb
93 et al. 2014). Moreover, wetland vegetation composition gradients can be affected by adjacent

94 land covers that influence the sources/dispersal of plant propagules and physicochemical
95 conditions of wetlands (Houlahan et al. 2006; Kraft et al. 2019). For instance, the more abundant
96 a species is in nearby upland forest types, the more likely it is to occur in a swamp (Pitman et al.
97 2014). We hypothesize that the inclusion of distance-to information with regards to other land
98 cover features besides water can also enrich the features of wetland classes and constrain the
99 classification result.

100 In last decade, deep learning (DL) techniques, notably convolutional neural network
101 (CNN)-based methods, have led to great success in image segmentation and outperformed
102 traditional pixel- and object-oriented classification methods, due to their ability to capture
103 contextual information from images (Du et al. 2020; Gonzalez-Perez et al. 2022; Zhang et al.
104 2020). For wetland mapping, applications of DL methods have been mostly limited to use or
105 incorporation of optical data (Dang et al. 2020; DeLancey et al. 2019; Du et al. 2020; Dutt et al.
106 2024; Gonzalez-Perez et al. 2022; Gunen 2022; Hosseiny et al. 2021; Hu et al. 2021; Jamali et al.
107 2021b; Li et al. 2021; Lv et al. 2023; Mainali et al. 2023) and have been investigated to a lesser
108 extent using radar data exclusively, due to complex scattering mechanisms for landcover classes
109 and speckle noise of radar imaging (Guo et al. 2023; Lam et al. 2023; Mohammadimanesh et al.
110 2019; Scepanovic et al. 2021). CNN-based architecture like U-Net remains popular for its
111 simplicity and effectiveness and has been widely introduced in landscape monitoring and
112 wetland mapping (Du et al. 2020; Dutt et al. 2024; Gonzalez-Perez et al. 2022; Li et al. 2021).
113 Recently, various innovative DL techniques such as attention mechanisms for enhancing model
114 focus on relevant features and transformers allowing models trained on large datasets have also
115 emerged for improving wetland classification performance (Jamali and Mahdianpari 2022;
116 Jamali et al. 2023; Marjani et al. 2024; Radman et al. 2024). However, these advanced DL

117 techniques tend to become computationally intense by introducing massive number of
118 parameters. Thus, when developing DL models, attention also needs to focus on the trade-off
119 between model complexity (e.g., network depths) and accuracy.

120 This study aims to propose a DL-based wetland classification method by integrating
121 Sentinel-1 Synthetic Aperture Radar (SAR) imagery and ancillary data. The innovations and
122 contributions of the methodology include:

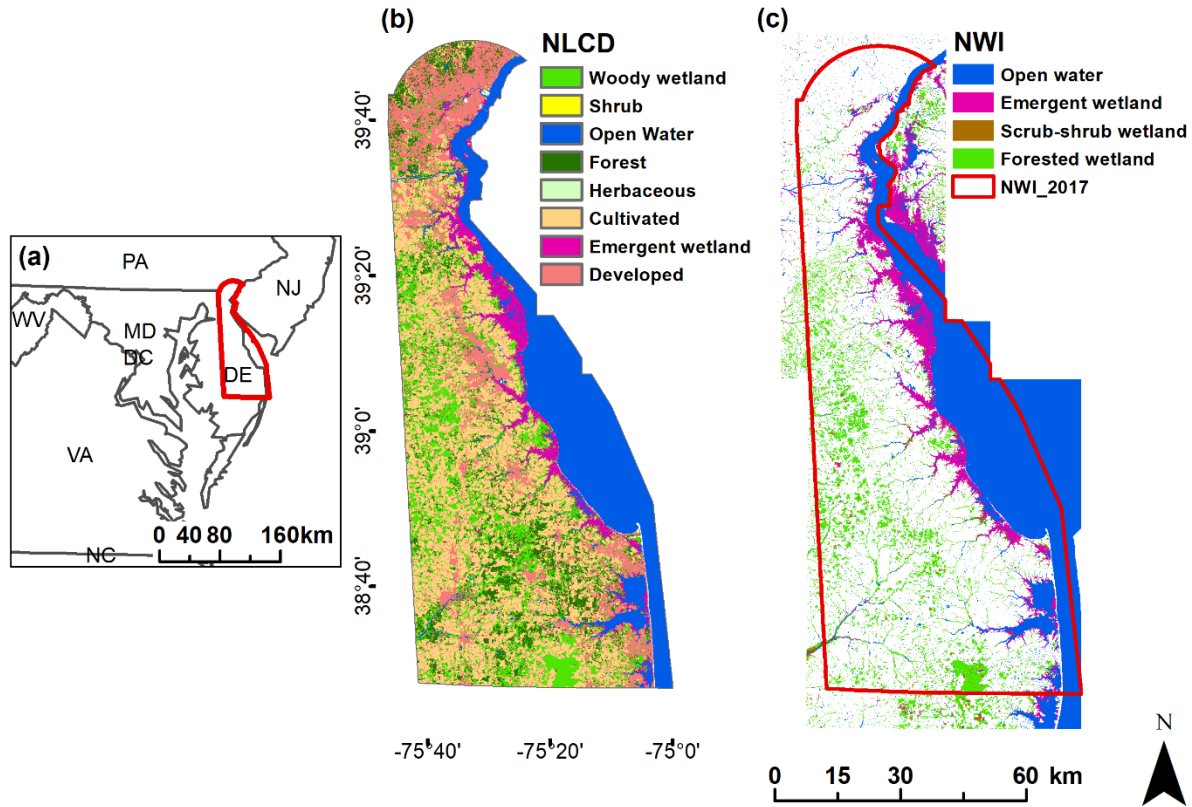
- 123 I. A depth-adaptive CNN model based on U-Net architecture with various depths (U-
124 Net A, B, C, D) was proposed, taking into account the trade-off between model
125 computational cost and accuracy.
- 126 II. Time-series SAR data during leaf-off season and a CNN-based self-supervised SAR
127 denoising procedure [Enhanced Noise2Noise (EN2N) model] (Tan et al. 2022) were
128 employed, allowing for large-scale wetland mapping without considering weather
129 conditions.
- 130 III. Multi-land cover proximity information that quantifies the nearest neighbour
131 distances was introduced for the first time to enhance classification accuracy.

132 The updated U.S. Fish and Wildlife Service National Wetlands Inventory (NWI) product in
133 Delaware was employed to verify model for distinguishing typical wetland classes, i.e., emergent
134 (EM) wetlands, scrub-shrub (SS) wetlands, forested (FO) wetlands, and open waters. We also
135 evaluated the method generalizability at a larger spatial extent (i.e., the entire Delmarva
136 Peninsula including portions of Maryland and Virginia) through comparisons with existing NWI
137 and land cover products.

138 **2. Materials and methods**

139 **2.1 Study area**

140 The study area is within the Delmarva Peninsula, adjacent to the Chesapeake Bay, U.S. (Figure
141 1). It is characterized by a low relief landscape with an average elevation of 26 m above sea
142 level. The temperature ranges from an average of approximately 2 °C in January and February to
143 25 °C in July and August (Shedlock et al. 1999). Annual precipitation is ~1200 mm with an even
144 distribution throughout the year, and the annual evapotranspiration is ~600mm, with a peak in
145 the summer and a trough in winter. Abundant water supply and poorly drained soil on lowlands
146 together contribute to the widespread vegetated wetlands in this region (Lowrance et al. 1997). In
147 this region, many wetlands are inundated or saturated for a short period with a peak normally
148 occurring in early spring before leaf-out (March/April) with low evapotranspiration conditions.
149 Land cover of this area is dominated by croplands (~32%), forests (~25%), and grasslands
150 (~5%), according to the 2019 National Land Cover Database (NLCD), (Figure 1b). A
151 considerable portion (~60%) of forested areas are forested wetlands which is the predominant
152 wetland class in the study area. This region also has other nontidal wetlands distributed over the
153 inland portion of the Delmarva, including EM wetlands (i.e., marshes and wet meadows
154 dominated by emergent plants), SS wetlands (i.e., swamps with shrubs or trees), and open
155 shallow water bodies (Figure 1c).



156

157 **Figure 1.** Location of study area (a). Panels (b) and (c) respectively display the land cover map
 158 derived from 2019 National Land Cover Database (NLCD) and wetland classes extracted from
 159 Delaware NWI updated in 2017.

160 **2.2 Data and processing**

161 **2.2.1 SAR imagery and denoising**

162 We used C-band SAR imagery from Sentinel-1 satellite with ground range detected (GRD)
 163 projection as the primary data input to classify wetlands (Table 1). The Sentinel-1 SAR mission
 164 has a regular revisit interval (12 days) and high spatial resolution (typically 10-m grid). Images
 165 in the winter-spring (from November 1st, 2017 to March 1st, 2018, 11 dates in total) with VH and
 166 VV polarizations and ascending orbit were downloaded. The selection of winter-spring is a time
 167 of year when the expression of wetland inundation is maximized in the study area. The original

168 records downloaded from the ESA were calibrated and ortho-corrected using the Sentinel-1
169 Toolbox (S1TBX) and the Graph Processing Framework from ESA's Sentinel Application
170 Platform. SAR reprocessing included the application of precise orbit files, border and thermal
171 noise removal, radiometric calibration, and orthorectification to project the images from slant
172 range to ground range. Finally, backscatter coefficient (σ°) was converted to a decibel (dB) scale
173 by $10 \times \log_{10} \sigma^{\circ}$. The mean values of VV and VH (VV_mean, and HH_mean) were also
174 calculated as data input to of wetland classification.

175 The leaf-off time-series SAR imagery was denoised using the EN2N model, a CNN-
176 based self-supervised SAR denoising procedure (Tan et al. 2022) (Figure S1-S2). This denoising
177 method introduced a self-supervised training strategy that time-series SAR data were denoised
178 without clean reference images, and a feature loss function was used to repair the spatial details
179 (Figure S2). This denoising method can also save a significant amount of time in image
180 processing while achieving good quality denoising performance (Tan et al. 2022).

181 *2.2.2 Topographic information*

182 Topographic information including elevation and slope derived from the elevation were
183 employed as ancillary datasets (Table 1). The elevation information was provided from 3D
184 Elevation Program (3DEP) bare earth Digital Elevation Model (DEM) with 1/3 arcsecond grid
185 (~10m) (Thatcher and Lukas 2021). The 3DEP data holdings provide seamless multi resolution
186 elevation data for earth science studies and mapping applications in the United States.

187 *2.2.3 Proximity information*

188 Multi-land cover proximity information that quantifies the distance of a location to the nearest

189 land covers were also introduced as additional data layers (Table 1). Maps of proximity to four
190 relevant land covers (i.e., forest, shrubland, herbaceous/grassland, and permanent water surface)
191 were derived based on the existing 30-m land cover product (2019 NLCD). The 2019 NLCD is
192 available via the Multi-Resolution Land Characteristics Consortium.

193 Specifically, we defined the Euclidean distance of a location to a particular proximal land
194 cover category as its proximity metric. First, we calculated the distance of each pixel to all
195 polygon objects belonging to a particular land cover type. Then, for each pixel, the proximity to
196 a particular land cover type was calculated as the sum of the distance to nearest and second
197 nearest targeted land cover polygons by use of a pixelwise sort (Figure S3). Once the pixel was
198 assigned to a given land cover category, the distance was set as 0. Finally, the proximity to each
199 land cover was normalized by the image size during model training (Equation 1).

$$200 \quad \text{Distance}(p^c, W, H) = \frac{\text{Distance}_{p_1}^c(x,y) + \text{Distance}_{p_2}^c(x,y)}{\sqrt{W^2 + H^2}} \quad (1)$$

201 where W and H are the width and height of the image, respectively. p^c refers to the polygons
202 belonging to a specific land cover type c . $\text{Distance}_{p_1}^c$ and $\text{Distance}_{p_2}^c$ refer to the distance of
203 each pixel (x, y) to the nearest and second-nearest land cover polygons.

204 *2.2.4 Reference data*

205 The reference data used for model verification were wetland polygons derived from the updated
206 2017 NWI dataset in Delaware (Table 1). There are five major categories in NWI classification
207 system, i.e., Marine, Estuarine, Riverine, Lacustrine, and Palustrine (Cowardin, 1979). We
208 extracted four wetland classes (i.e., EM, SS, FO wetlands, and open waters) from these
209 categories based on the “attribute” field of wetland polygons. Extracted wetland polygons were

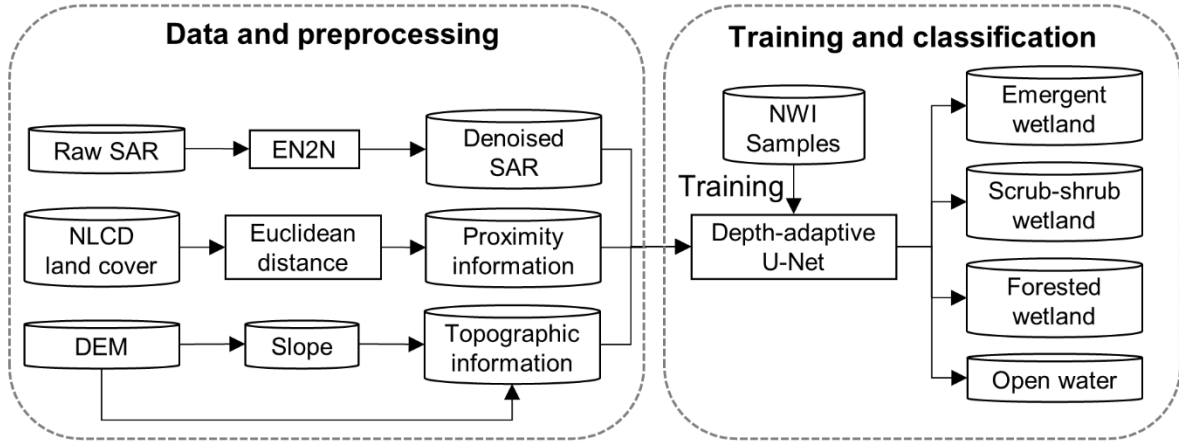
210 converted into binary rasters (hereinafter referred to wetland labels) to align with SAR imagery.

211 To evaluate the generalizability of our DL method at a large spatial extent (e.g., entire
 212 Delmarva Peninsula), we also downloaded the newly released NWI dataset in Maryland and
 213 Virginia (updated in 2019) as well as the 10-m 2020 ESA WorldCover dataset ([https://esa-
 214 worldcover.org/en](https://esa-worldcover.org/en)) for comparisons. It should be noted that the 2020 ESA WorldCover only
 215 includes one similar wetland class for comparison (i.e., herbaceous wetland which we treated as
 216 being similar to EM wetlands). This ESA WorldCover product was generated based on 131
 217 spatial localizing features including Sentinel-1 and 2 data, topographic features, and positional
 218 features using machine learning algorithms and demonstrated a higher accuracy than other land
 219 cover products, e.g., ESRI land cover product (Wang et al. 2022).

220 Table 1. Datasets used in this study.

Data	Description	Source	Acquisition Data
SAR imagery	Time-series C band imagery (VV, VH, VV_mean, and HH_mean)	Sentinel-1 from European Space Agency	Nov. 1 st , 2017- Mar. 1 st , 2018
Topographic information	DEM Slope	USGS 3DEP products Derived from DEM	
Proximity information	Distance to water (Distance_W) Distance to forest (Distance_F) Distance to shrub (Distance_S) Distance to herbaceous (Distance_H)	Calculated from 30-m NLCD	2019
Reference data	NWI	National Wetlands Inventory dataset for Delaware (2017), Maryland and Virginia portions of Delmarva (2019)	2017&2019
	ESA WorldCover	2020 ESA WorldCover	2020

221



222

223 **Figure 2.** Deep learning framework for wetland classification.

224

225 **2.3 Methods**

226 **2.3.1 Proposed depth-adaptive U-Net**

227 U-Net is a popular semantic segmentation model characterized by a symmetric U-shaped
 228 architecture, which includes an encoder-decoder structure with long skip connections (Figure 3).
 229 The encoder part is a feature extraction process implemented using multiple convolution
 230 operations, in which the spatial dimension is reduced while the channel information is enhanced.
 231 The decoder part is an expanding process that combines the feature and spatial information
 232 through a sequence of transposed convolution operations and concatenations with high-
 233 resolution features from the encoder path. By utilizing the concatenations that bypass layers in
 234 the encoder part, high-resolution features from earlier stages can be directly integrated. This
 235 enhances localization and prevents loss of spatial information.

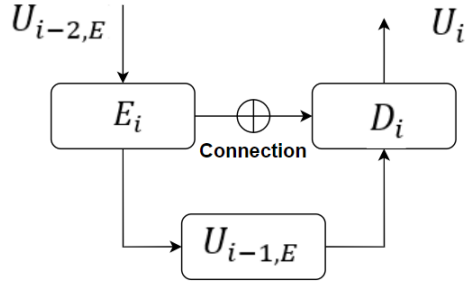
236 **2.3.1.1 Depth-adaptive U-Net architecture.** The full U-Net is designed with five stages of
 237 convolution operations (i.e., U-Net D), which requires massive computational resources,

238 especially when the input image is large. To optimize computational resources and accuracy, we
 239 developed a depth-adaptive U-Net network with three strengths: automatic depth optimization,
 240 multiscale fusion, and model compression. Four different depths of U-Net (U-Net A, B, C, and
 241 D) (Figure 4 and Figure S4) were integrated in one encoder part, and the output of each model
 242 becomes the hidden layer in the ensemble model. The depth that achieves best accuracy was
 243 automatically determined during training. A full U-Net can be represented as a recursive
 244 structure:

$$245 \quad U_i = D_i(U_{i-1,E} + U_{i-2,E}) \quad (2)$$

$$246 \quad U_{i-1,E} = E_i(U_{i-2,E}) \quad (3)$$

247 Where $i=1,2,3,4,5$. E and D refer to the encoder and decoder parts, respectively, $U_{i,E}$ refers to
 248 the output of the encoder part in U_i .



249

Figure 3. Recursive structure of a U-Net

250

251 Each encoder part contains a max-pooling layer and two repeat convolution operations
 252 (Conv). The max-pooling down-sampling the input representation to half size. The Conv
 253 operation has a 3×3 convolution kernel followed by a rectified linear unit (ReLU) and a
 254 BatchNorm layer. The operation can be formulated as:

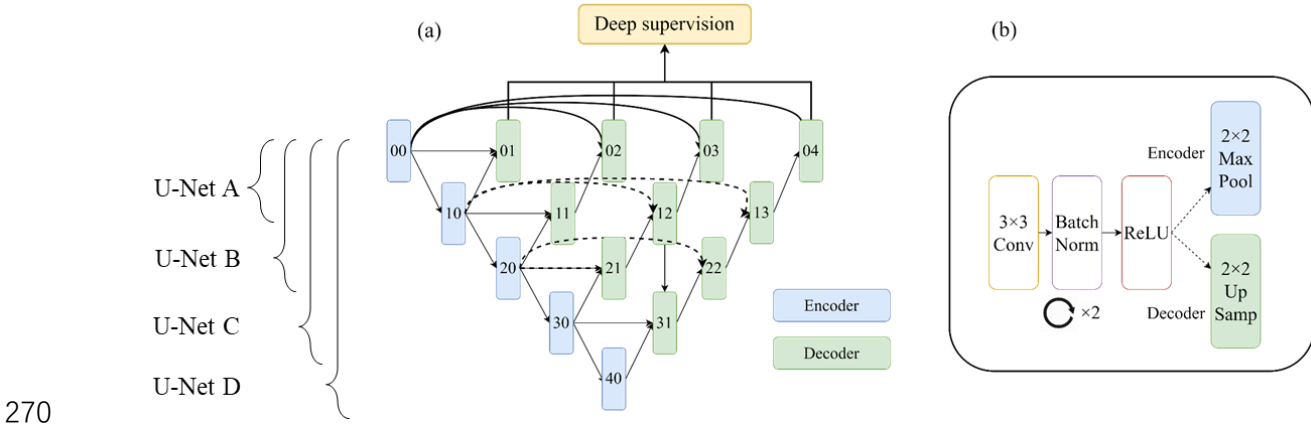
$$255 \quad E_{i,j} = f_{pool}(f_{conv2}(f_{conv1}(E_{i-1,j}))) \quad (4)$$

256 where the E_{ij} is the hidden feature in j th of depth i . The $E_{i-1,j}$ is the upper hidden feature of E_{ij} .
 257 The f_{pool} represents the downsampling operation by max-pooling and f_{Conv} represents the Conv
 258 operation.

259 This decoder adopts a structure similar to the encoder by replacing the max-pooling layer
 260 with an up-sampling layer to bilinearly extend the in-depth feature to the original size. For each
 261 decoder block, the upscale was set to 2 to ensure that the output size is the same as the forward
 262 encoder output. A skip connection was used to concatenate the encoder and decoder at each
 263 layer. The original skip connection was simply as the residual learning. The hidden feature from
 264 the encoder was directly concatenated to the decoder part. The decoder operation can be
 265 formulated as:

$$266 \quad D_{i,j} = (f_{TConv2}(f_{TConv1}(D_{i+1,j-1}))) + E_{i,0} \quad (5)$$

267 where the $D_{i,j}$ is the hidden feature of decoder part in j th of depth i . The $D_{i+1,j-1}$ is the
 268 lower hidden feature of $D_{i,j}$. The f_{TConv} presents the transposed convolution that up-sampling the
 269 feature. $E_{i,0}$ presents the output of encoder in the same depth.



271 **Figure 4.** Diagram of depth-adaptive U-Net with four different depths (U-Net A, B, C, and D) (a)
 272 and structure of each encoder/decoder node (b). Each encoder/decoder node includes a 3×3
 273 convolution block, two batch norm blocks, a ReLU block, and a Max Pooling block (for encoder)
 274 or Up-Sampling block (for decoder).

275 *2.3.1.2 Fusion loss function with deep supervision.* A fusion loss function was introduced by
 276 supervising all outputs of models with different depths:

$$277 \quad L_{fusion} = L_2 + L_3 + L_4 + L_5 \quad (6)$$

278 For each model, a hybrid segmentation loss, including a pixelwise loss-based Cross-Entropy loss
 279 with class weight, an intersection over union (IoU) based Dice Loss, and an edge loss based on
 280 Binary Cross-Entropy loss, is introduced:

$$281 \quad L = \alpha L_{cross\ entropy} + (1 - \alpha) L_{Dice} + L_{Edge} \quad (7-a)$$

$$282 \quad L_{cross\ entropy}(y, \hat{y}) = - \sum_{i=0}^{h \times w} \sum_{c=1}^C w_c \log(\hat{y}) y \quad (7-b)$$

$$283 \quad L_{Dice}(y, \hat{y}) = 1 - \frac{2y\hat{y}+1}{y+\hat{y}+1} \quad (7-c)$$

$$284 \quad L_{Edge}(e, \hat{e}) = \sum_{i=0}^{h \times w} e \cdot \log \hat{e} + (1 - e) \cdot \log (1 - \hat{e}) \quad (7-d)$$

285 Here, \hat{y} is the predicted confidence value for one class by the model. \hat{e} is the predicted
 286 confidence value for the edge. w_c It is the class weight that can be calculated as:

$$287 \quad w_c = 1 + 1 / Size_c \quad (8)$$

288 where $Size_c$ is the covering area of different wetland classes. By training L_{fusion} , we can
289 simultaneously monitor models' performance with different depths and select the best model
290 structure.

291 *2.3.2 Comparison with other classifiers*

292 A comprehensive comparison experiment was designed in this study. We started by a
293 comparison with traditional classification methods. Machine learning algorithms such as the
294 random forest (RF), support vector machine, and boosted regress trees have been commonly
295 used in land cover classification tasks from remote sensing data (Zhang et al. 2020). In
296 particular, the RF was considered to outperform other machine learning classifiers due to its
297 ability to handle high-dimensional datasets and mitigate overfitting and has been widely applied
298 in wetland mapping (Adugna et al. 2022; Amani et al. 2019; Jamali et al. 2021a; Rodriguez-
299 Galiano et al. 2012). In our study, we examined the inclusion of multi-land cover proximity
300 information and a CNN-based self-supervised SAR denoising procedure in our proposed DL
301 method and the RF method for classifying different wetland classes. Moreover, because DL
302 models require a significantly higher number of parameters, it is inadequate to only compare
303 with machine learning methods. Therefore, we additionally compared the efficiency of our
304 proposed DL model against two established state-of-the-art CNN-based models: DeepLabv3+
305 (Chen et al. 2018) and DANet (Fu et al. 2019) in terms of accuracy, number of parameters, and
306 processing time. We have not chosen any transformer-based model for comparison, because this
307 type of model often requires massive training samples to achieve optimum performance, which
308 does not apply to our datasets.

309 *2.3.3 Model training and classification schemes*

310 For DL model training, wetland labels combined with all data inputs were split into small image
311 patches using a moving window (512*512 pixels) to allow for model training and classification.
312 Based on all image patches that contain wetland categories, we randomly generated training,
313 validation, and testing image sets according to the 6:2:2 ratio (Figure S5 shows the distribution
314 of validation and test sets). The proposed DL model was written by PyTorch and trained with the
315 AdamW optimizer (Loshchilov and Hutter 2017). A batch size of 64, distributed over 3 GPUs
316 (GeForce RTX 2080 Ti) was used. The learning rate was linearly ramped up during the first ten
317 epochs as $1e-3$. After this warmup, we decayed the learning rate with a cosine schedule. The
318 weight decay also followed a cosine schedule from 0.04 to 0.4. These parameters were applied to
319 all DL models used in this study.

320 For RF model training, we generated a total of 8,000 random sample points (2,000 for
321 each category: open water, EM, SS, and FO) from the DL training image patches. To enhance
322 RF training data quality, the random selection of training sample points also followed the
323 criterion that the neighbouring 3×3 domain (i.e., $30m \times 30m$) around each sample point has
324 uniform land cover. The validation and test data used for assessing accuracy of RF were the
325 same as those used for DL to make model accuracy metrics comparable. In RF classification, we
326 used constant ntree (the number of trees) of 500, and mtry (the number of variables at each split)
327 equal to the square root of the number of total inputs.

328 *2.3.4 Accuracy evaluation metrics*

329 To evaluate the classification performance of each method, five accuracy metrics: precision,
330 recall, overall accuracy (OA), F1-score, and mean IoU (MIoU), were calculated on test set to

331 assess model accuracy.

332 OA is the most intuitive performance measure, and it is simply a ratio of correctly
333 predicted observations to the total observations, which can be written as:

$$334 \quad OA = S_d/n \times 100 \quad (9)$$

335 where S_d is the total number of correctly classified targets, n is the total number of validation
336 targets.

337 The Precision measures the fraction of true positive detections ($X_{ij}/X_j \times 100$), and the
338 Recall measures the fraction of correctly identified positives ($X_{ij}/X_i \times 100$), where X_{ij} is the
339 observation in row i column j in the confusion matrix, X_i is the marginal total of row i and X_j is
340 the marginal total of column j in the confusion matrix. F1-score is the harmonic average of
341 Precision and Recall:

$$342 \quad F1 - score = \frac{Precision \times Recall}{Precision + Recall} \times 2 \quad (10)$$

343 The segmentation performance was evaluated by IoU, which is the ratio of overlapped
344 area to the area of union between predicted and ground truth categories, and is written as:

$$345 \quad IoU(A, B) = \frac{Area(A \cap B)}{Area(A \cup B)} \quad (11)$$

346 where A and B correspond to ground truth and predicted wetland objects, respectively. IoU
347 ranges from 0 to 1, where 0 represents no overlap and 1 represents perfect segmentation. The
348 MIoU is calculated by averaging the IoU of all wetland classes.

349 **3. Results**

350 ***3.1 Wetland classification performance***

351 The overall classification performance of our DL method was satisfactory (OA = 0.93), which
352 was higher than that of RF (OA = 0.89) (Table 2 and Figure 5) using all denoised SAR and
353 topographic and proximity information. The F1-score of each wetland class using DL was
354 significantly higher than that using RF, especially for SS and FO wetlands (e.g., F1-score was
355 0.54 and 0.78 for SS and FO, respectively, using DL, and was 0.00 and 0.41 for SS and FO,
356 respectively, using RF). RF generated high Precision estimates but very low Recall estimates for
357 SS and FO. Based on the confusion matrix shown in Figure 5, there was a considerable amount
358 of SS and FO wetlands classified into other categories in RF.

359 At the object level, the wetland types predicted by the DL method was more comparable
360 with the NWI wetland labels than RF predicted wetland types. The RF presented extensive “salt-
361 and-pepper” appearance in its results (Figure 6). This was also demonstrated by the higher MIoU
362 of our DL method (0.60) than RF (0.18) for all wetland classes (Table 2). Additionally, by visual
363 check, the pattern of DL predicted EM wetland was also comparable with the herbaceous
364 wetland from 2020 ESA WorldCover (Figure 6).

365

366

367

368

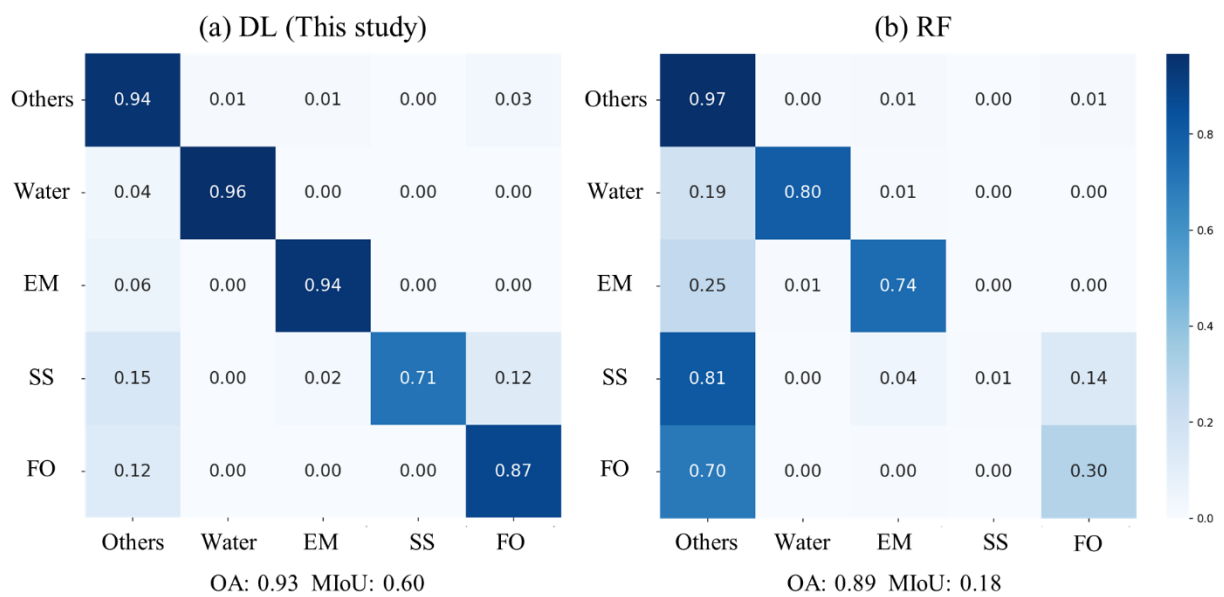
369

370 **Table 2.** Classification accuracy of the proposed deep learning (DL) method and random forest
 371 (RF) method using denoised SAR, topographic data, and proximity information.

Category	DL (this study)			RF		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Water	0.76	0.96	0.85	0.89	0.80	0.84
EM	0.82	0.94	0.87	0.73	0.74	0.74
SS	0.44	0.71	0.54	0.25	0.00	0.00
FO	0.71	0.87	0.78	0.64	0.30	0.41
Others	0.98	0.94	0.96	0.91	0.97	0.94
OA		0.93			0.89	
MIoU		0.60			0.18	

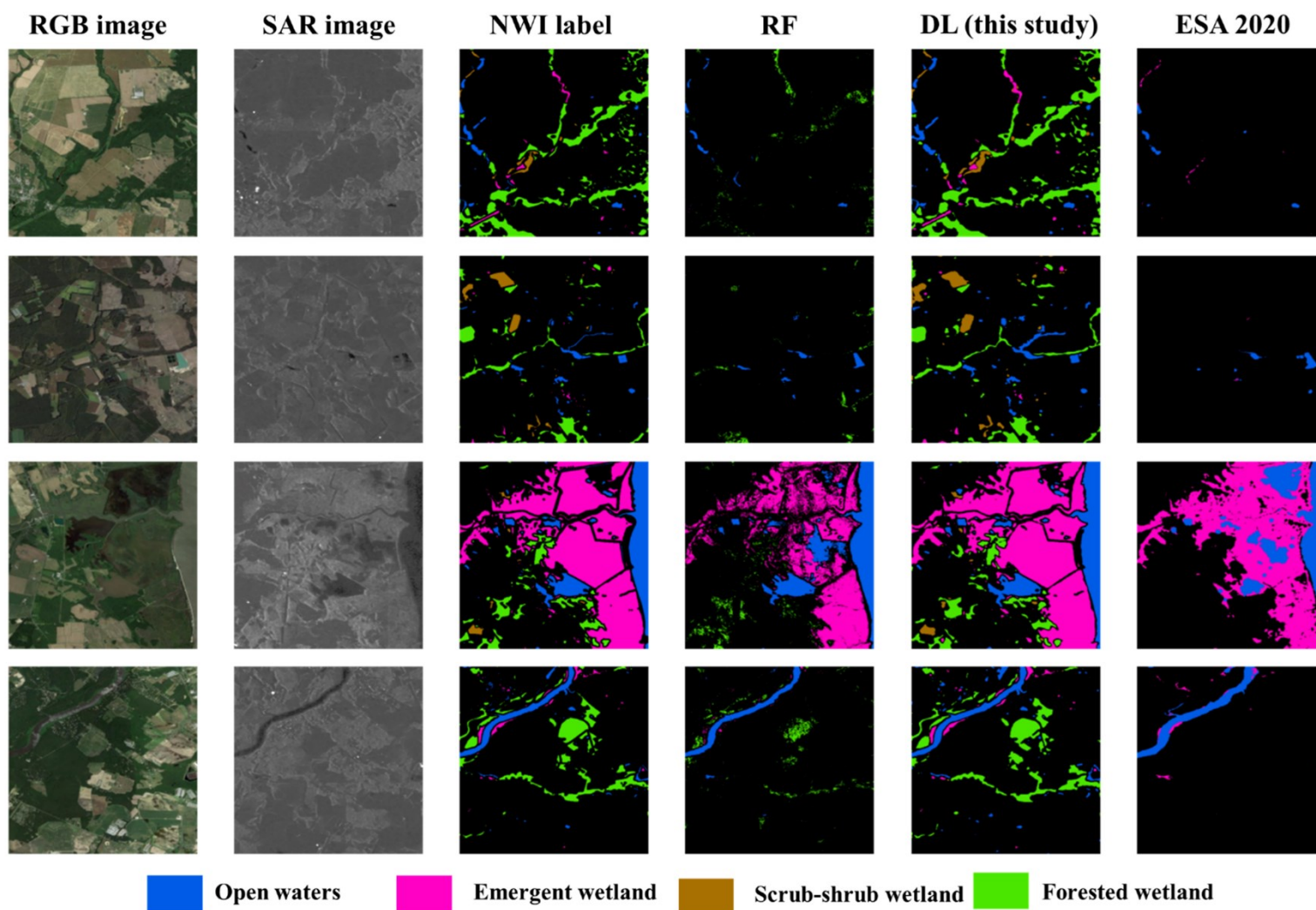
372

373



374

375 **Figure 5.** Confusion matrix of the proposed DL method (a) and random forest (RF) method (b).
 376 y axis represents actual, and x axis represents predicted categories.



377

378 **Figure 6.** Comparison of our deep learning (DL) predicted wetland types with random forest (RF) output, 2020 ESA WorldCover
 379 product (only herbaceous wetland), and NWI wetland labels. The first column shows the true colour combination of Red-Green-Blue
 380 (RGB) bands from Google Earth; The second column shows denoised SAR images with VH polarization.

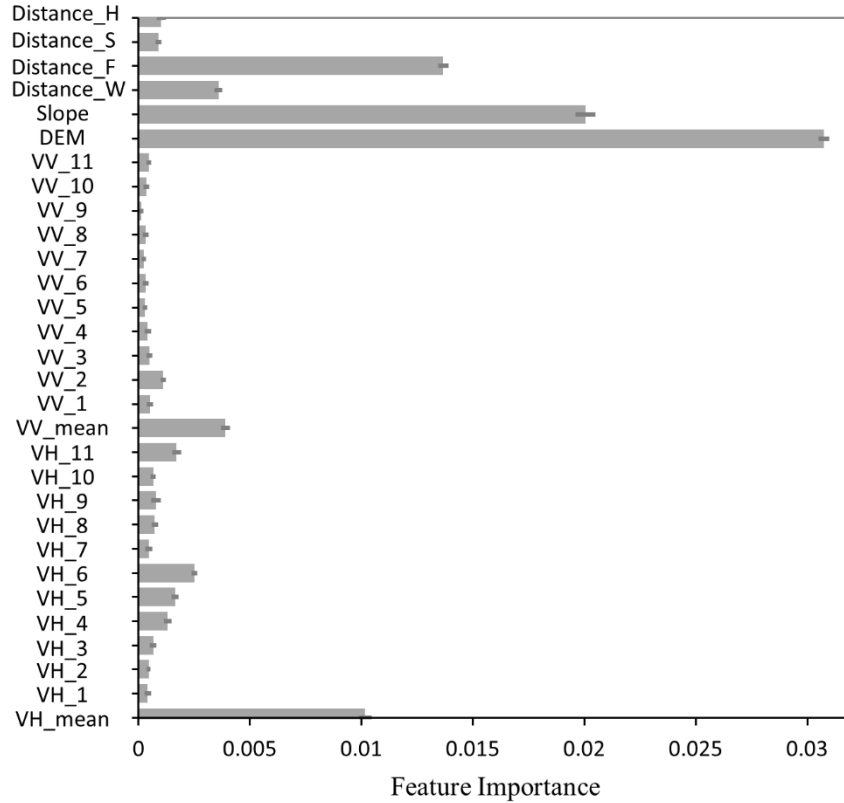
381 **3.2 Inclusion of multi-land cover proximity information**

382 The inclusion of multi-land cover proximity information was investigated in both DL and RF
 383 methods by comparing classification accuracy using (Table 2) and without using (Table 3)
 384 proximity information. As we expected, excluding proximity information resulted a decrease in
 385 classification accuracy for all wetland classes in both methods by comparing Table 3 to Table 2.
 386 For example, the MIoU of DL decreased from 0.60 to 0.54, and the MIoU of RF decreased from
 387 0.18 to 0.16. There was a significantly decreased accuracy for FO wetland in RF method without
 388 using proximity information, e.g., the F-1 score of FO decreased from 0.41 to 0.03. According to
 389 the relative importance score generated from RF, the proximity to forest (Distance_F) and
 390 proximity to water (Distance_W) contributed most in classification, following topographic
 391 information (DEM and slope) (Figure 7).

392 **Table 3.** Classification accuracy of the proposed deep learning (DL) method and random forest
 393 (RF) method without using proximity information.

Category	DL (this study)			RF		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Water	0.72	0.95	0.82	0.89	0.76	0.82
EM	0.82	0.93	0.87	0.76	0.73	0.73
SS	0.51	0.71	0.60	1.00	0.00	0.00
FO	0.70	0.78	0.74	0.46	0.01	0.03
Others	0.97	0.94	0.96	0.86	0.99	0.92
OA		0.93			0.86	
MIoU		0.54			0.16	

394



395

396 **Figure 7.** Relative importance of all data input from random forest model. Distance_H,
 397 Distance_S, Distance_F, and Distance_W represent the distance to herbaceous, shrubland, forest,
 398 and open water, respectively. VV_x (or VH_x) represents the 11 time series SAR images with
 399 VV (or VH) polarization acquired during the leaf-off period. VV_mean and VH_mean are mean
 400 values of VV bands and VH bands, respectively.

401 **3.3 Effectiveness of CNN-based self-supervised denoised SAR imagery**

402 The effectiveness of the SAR denoising procedure using EN2N was also investigated in both DL
 403 and RF methods by comparing classification accuracy using (Table 2) and without using (Table
 404 4) the SAR denoising procedure. There was also decreased classification accuracy when
 405 excluding the denoising procedure (i.e., using the raw SAR) in both methods. The MIoU of DL
 406 and RF decreased from 0.93 to 0.47, and from 0.18 to 0.15, respectively (Table 2 and Table 4). It
 407 is also notable that, without the SAR denoising procedure, there was a significantly decreased

408 accuracy for FO wetland in RF method, e.g., the F-1 score decreased from 0.41 to 0.07.

409 **Table 4.** Classification accuracy of the proposed DL method (this study) and RF method without
410 the SAR denoising procedure. “Others” represents all other land cover categories.

Category	DL (this study)			RF		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Water	0.73	0.92	0.82	0.87	0.76	0.81
EM	0.81	0.87	0.87	0.70	0.70	0.70
SS	0.55	0.71	0.61	0.00	0.00	0.00
FO	0.73	0.75	0.76	0.52	0.04	0.07
Others	0.95	0.91	0.94	0.89	0.98	0.93
OA		0.89			0.86	
MIoU		0.47			0.15	

411

412 *3.4 Computational cost and accuracy*

413 Table 5 provides a comparison of computational cost and accuracy obtained from the proposed
414 depth-adaptive U-Net and two CNN-based DL models (DeepLabV3+ and DANet) using all data
415 put. As seen, as the depth of U-Net increases, the resulting accuracy increased while requiring
416 more calculation resources. The U-Net A with relatively simple network structure (104,064
417 parameters) had the lowest accuracy (MIoU = 0.33, OA = 0.88), and the U-Net D with relatively
418 complex network structure (8,854,176 parameters) achieved the best resulting accuracy (MIoU =
419 0.60, OA = 0.93). By comparison, model application could be pruned to U-Net C to achieve a
420 satisfactory performance at large scale, because U-Net C achieved a MIoU of 0.58 close to that
421 of U-Net D (MIoU=0.60) but had less than one quarter of parameters and a significantly reduced
422 processing time. A pattern worth noticing is that, when compared to our U-Net C, the two CNN-
423 based DL models (DeepLabV3+ and DANet) that comprise of much higher number of
424 parameters (>20 times) demonstrated the same level accuracy (MioU of 0.58~0.60).

425

426 **Table 5.** Computational cost and model accuracy based on U-Net with different depth (A, B, C,
 427 and D), DeepLabV3+, and DANet. Unit of inference time is seconds per Sentinel-1 scene
 428 (25341×19433 px).

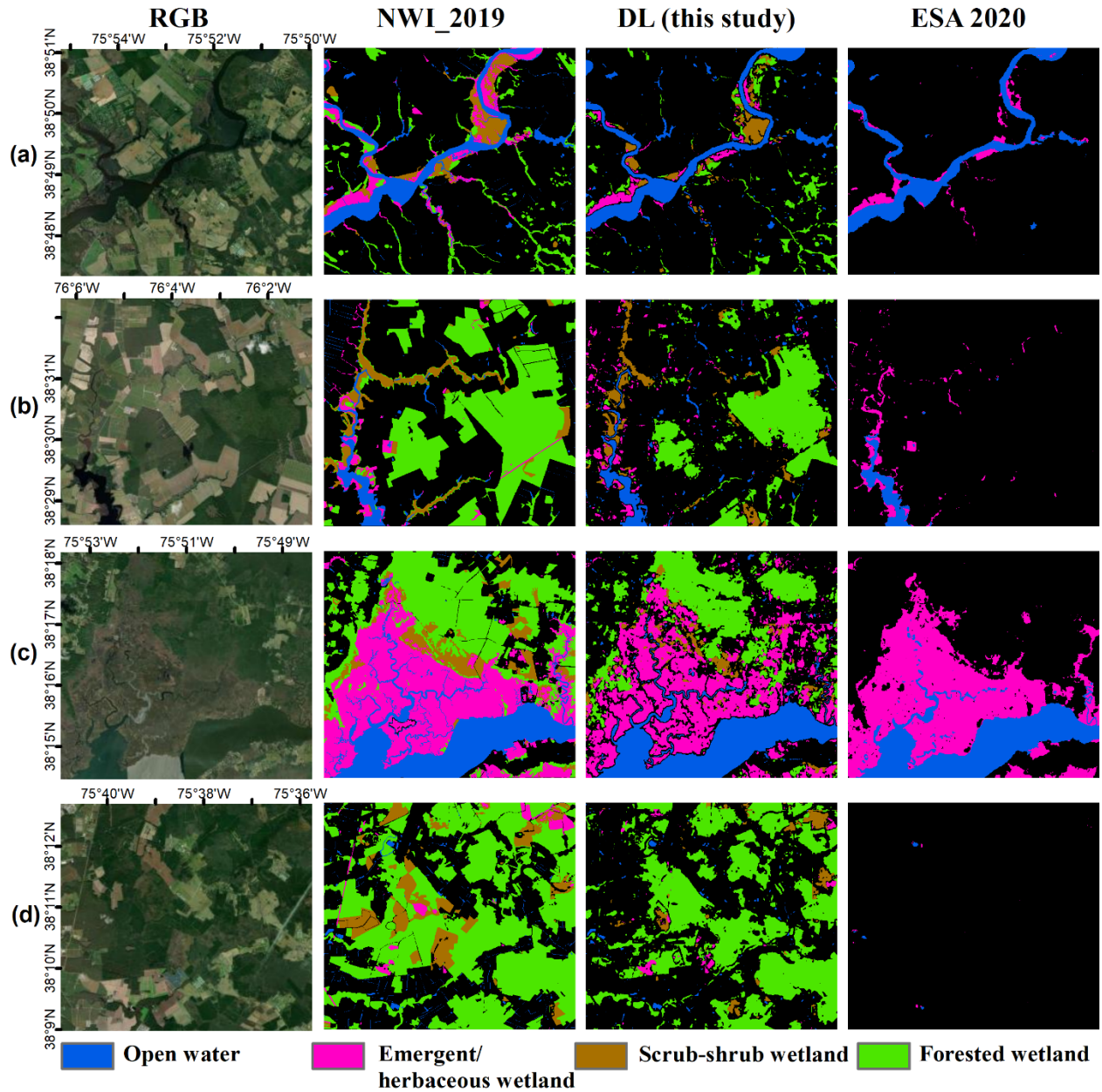
U-Net depth	MACs (GB)	Parameters	Inf. Time	MIoU	OA
U-Net A	17.83	104,064	73 Sec	0.33	0.88
U-Net B	31.19	473,856	79 Sec	0.51	0.92
U-Net C	44.50	1,950,720	88 Sec	0.58	0.93
U-Net D	57.81	8,854,176	137 Sec	0.60	0.93
DeepLabV3+	101.787	41,292,390	367 Sec	0.58	0.93
DANet	115.816	50,081,850	390 Sec	0.60	0.93

429

430 **3.5 Method generalizability**

431 To further test the generalizability of our DL method at a large spatial extent, we predicted
 432 wetland classes for the entire Delmarva Peninsula during the leaf-off season between 2017-2018
 433 using the trained model and compared it with the newly released 2019 NWI wetland and the
 434 2020 ESA WorldCover (only herbaceous wetland) products (Figure S6 and Figure 8). Generally,
 435 the spatial extent of the DL predicted wetland classes was comparable with that from NWI and
 436 2020 ESA WorldCover for the entire Delmarva. For example, EM wetlands mostly occurred
 437 along the coastal areas, and FO/SS wetlands were distributed within the inland portion of the
 438 Delmarva (Figure S6). The total area of open water surface in Delmarva was 459,360 ha from
 439 our DL prediction, which was close to that from NWI (464,660 ha) and 2020 ESA WorldCover
 440 products (455,540 ha) (Table 6). There was a conservative pattern of wetland extent predicted for
 441 the entire Delmarva in our DL method compared to NWI (Figure 8), with an underestimation by
 442 20%-30% for EM and FO wetlands and 60% for SS wetlands. We observed numerous cases
 443 where SS wetlands in NWI were omitted or classified as FO wetlands in our DL prediction (e.g.,

444 Figure 8d).



451 **Table 6.** Comparisons of wetland areas (Unit: ha) from the proposed deep learning (DL) method
 452 with NWI and 2020 ESA WorldCover products in the Delmarva Peninsula.

Category	Entire Delmarva Peninsula			Delmarva Peninsula excluding Delaware		
	NWI	DL(This study)	2020 ESA WorldCover	NWI	DL(This study)	2020 ESA WorldCover
Water	464,660	459,360	455,540	402,010	398,850	400,250
EM	146,020	112,310	121,230	114,380	86,410	93,990
SS	31,160	12,070		27,770	9,110	
FO	214,530	150,400		156,880	102,550	

453

454 **4. Discussion**

455 ***4.1 Significance of this study***

456 Compared to optical data, design of a robust method for mapping wetlands based on SAR data is
 457 challenging, due to complex scattering mechanisms and the speckle noise caused by the coherent
 458 nature of the SAR imaging. Recent development of DL techniques has shown great advantages
 459 in learning complex contextual information from images, but its application with remote sensing
 460 imagery like SAR usually require prediction over a massive area, making computational
 461 efficiency a critical factor to be considered. In this study, we were able to classify wetland
 462 classes based on depth-adaptive U-Net by combining leaf-off Sentinel-1 C-band SAR imagery
 463 and ancillary data in eastern U.S.. We found that our model not only outperformed the traditional
 464 RF methods in terms of accuracy but also had a significantly reduced computational cost
 465 compared to state-of-the-art CNN models (e.g., DeepLabv3+ and DANet) without loss of
 466 accuracy. The inclusion of multi-land cover proximity information and CNN-based self-
 467 supervised SAR denoising procedure (EN2N) can both be means to enhance classification
 468 accuracy. These indicate that our proposed DL method is efficient and could be integrated for

469 automatic recognition of wetland classes for supporting operational wetland mapping.

470 An important innovation of this study involves inclusion of multi-land cover proximity
471 metrics (i.e., distances to water, forest, shrub, and herbaceous) as additional data layers to
472 constrain classification models. These proximity metrics helped capture hydrological
473 connectivity (e.g., distance to water) and ecological dispersal dynamics (e.g., distance to forest),
474 which are known drivers of wetland type distribution and composition. To the best of our
475 knowledge, the incorporation of multi-land cover proximity to constrain wetland classification in
476 DL methods has not been explored to any significant extent. Our results showed that adding
477 these proximity layers improved the mean MIoU and F1-score in both DL and RF methods, with
478 a more pronounced improvement in FO detection using RF (Table 2-3). This finding underlines
479 the importance of contextual information regarding wetland adjacency effects for pixel-oriented
480 classifiers like RF, which benefited significantly from additional geographic data to constrain
481 and refine classification decisions. Moreover, the relative importance of topographic information
482 was the highest, followed by proximity to forest (Distance_F) and proximity to water
483 (Distance_W), which highlight the importance of understanding the local environmental
484 condition and mesoscale adjacency effects in effective classification.

485 The CNN-based self-supervised denoising method (EN2N) employed in our study
486 bypasses the need for clean SAR images and has an acceptable computational efficiency cost,
487 and thus can be easily transplanted in other geographic locations for SAR denoising. When
488 excluding use of the SAR denoising method in wetland classification, both the DL and RF
489 methods generated decreased accuracy compared to those using denoised SAR data (Table 2 and
490 Table 4). There was also significantly decreased accuracy for FO wetlands in RF without using
491 denoised SAR, which demonstrated the importance of SAR denoising for woody wetland

492 classification when using pixel-oriented classification methods. Use of the CNN-based U-Net in
493 the DL classifier can actually help capture the contextual feature information and thus reduce the
494 speckle noise. This result indicated that the inherent denoising characteristic of CNN did not
495 perform noise reduction as well as the EN2N. Thus, the individual CNN-based self-supervised
496 SAR denoising procedure is recommended for DL classification tasks with SAR imagery.

497 ***4.2 Limitations and future work***

498 This study employed leaf-off C-band SAR imagery from Sentinel-1, which can penetrate cloud
499 and sparse canopy with medium penetration depth. However, the penetration capacity of C-band
500 measurements is not sufficient to penetrate dense canopies such as evergreen broadleaf forests.
501 In our study area, inundated wetlands are principally covered by deciduous forests (Lang et al.
502 2020), and the winter season with leaf-off condition was focused on. Therefore, the limited
503 capacity of C-band SAR to penetrate high-density canopy had substantially less influence on the
504 mapping result. However, to improve the applicability of the method, SAR images from sensors
505 with broader wavebands (i.e., L-band and P-band SAR) could be preferred over C-band images
506 in areas with high-density canopies.

507 The accuracy for SS wetland class was not high, although it was improved in our DL
508 model compared to the RF model, indicating the difficulty of SS wetland detection. There was an
509 omission of SS to some degree (especially in RF), i.e., a number of SS wetlands classified into
510 FO and other categories (Figure 5). The lack of SS training labels relative to other wetland
511 classes as well as the backscattering similarity between SS and FO/other ecosystems could be the
512 causes for the misclassification especially when they mixed with each other. Other ancillary
513 information like canopy height could further contribute to recognition of different wetland

514 classes (Gonzalez-Perez et al. 2022). Concerns related to how to more adequately group wetland
515 classes together to improve accuracy could also be addressed in future studies. In addition, the
516 DL method also generated a conservative/underestimated pattern of wetland classes (especially
517 SS) for the entire Delmarva, compared to 2019 NWI and 2020 ESA WorldCover products. A
518 model retraining using local samples in different geographic areas may be further tested to
519 improve model generalizability (Mainali et al. 2023).

520 A lack of sufficient ground truth labels for training and validation is a common issue that
521 hampers the application and limits the performance of DL-based classification approaches.
522 Annotating labels manually, especially polygon objects to feed CNN, usually costs considerable
523 human labour and requires prior knowledge. In this study, we benefited from the availability of
524 the updated 2017 NWI product for Delaware as the source of training and validation labels.
525 However, such high accuracy datasets may not be available for national or global scale
526 applications. Also, there often exist insufficient training samples for certain categories, e.g. the
527 training labels of SS wetland were less relative to other classes in the NWI reference. Self-
528 Supervised Learning (SSL) technology is an innovative unsupervised approach poised to solve
529 the challenges posed by the over-dependence of labelled data in DL and is now considered to be
530 the future of machine learning (Tao et al. 2023; Zhang and Han 2023). SSL can learn
531 intermediate representation of data, which is useful in understanding the underlying semantic or
532 structural meanings that benefit a variety of practical downstream tasks. Recently, the emergence
533 of diffusion models also provides opportunities for generating controllable samples consistent
534 with real scenes (Yuan et al. 2023). A future direction of our work will be to find a more
535 conventional way to improve the efficiency in data organizing tasks based on these technologies.

536 **5. Conclusions**

537 To accurately and efficiently map different wetland classes with readily available datasets and
538 ensure robust methodology under clouds, a novel CNN-based DL classification method
539 incorporating denoised Sentinel-1 SAR and topographic and multi-land cover proximity
540 information was developed in this study. This method was verified in a typical wetland
541 landscape in Delaware using the updated NWI product and was further evaluated at a larger
542 spatial extent (i.e., the Delmarva Peninsula). The DL method significantly outperformed the
543 traditional RF methods for mapping different wetlands in terms of accuracy and was more
544 efficient than two well-known, state-of-the-art CNN-based models. Moreover, both the CNN-
545 based self-supervised SAR denoising procedure and inclusion of multi-land cover proximity
546 information further enhanced the classification accuracy of wetland classes, with a significant
547 improvement in forested wetland detection using RF methods. The depth-adaptive CNN
548 developed in this study helped to address trade-off between model performance and
549 computational cost and showed a reasonable generalizability when extended to the Delmarva
550 Peninsula. Our study demonstrates that this method holds promise for operational wetland
551 mapping using SAR at large scales.

552 **Acknowledgments**

553 This work was supported by the U.S. Department of Agriculture (USDA) Natural Resources
554 Conservation Service under their Conservation Effects Assessment Project (CEAP). This
555 research was a contribution from the Long-Term Agroecosystem Research (LTAR) network.
556 LTAR is supported by the USDA. Any use of trade, firm, or product names is for descriptive
557 purposes only, and does not imply endorsement by the U.S. Government.

558 **Disclosure statement**

559 No potential conflict of interest was reported by the author(s).

560 **Data availability statement**

561 The data and codes that support the findings of this study are available from the corresponding
562 author, LD, upon reasonable request.

563

564 **Reference**

- 565 Adeli, S., Salehi, B., Mahdianpari, M., Quackenbush, L.J., Brisco, B., Tamiminia, H., & Shaw, S. (2020).
566 Wetland Monitoring Using SAR Data: A Meta-Analysis and Comprehensive Review. *Remote*
567 *Sensing, 12*
- 568 Adugna, T., Xu, W., & Fan, J. (2022). Comparison of Random Forest and Support Vector Machine
569 Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote*
570 *Sensing, 14*
- 571 Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mohammad Javad Mirzadeh, S., White, L.,
572 Banks, S., Montgomery, J., & Hopkinson, C. (2019). Canadian Wetland Inventory using Google Earth
573 Engine: The First Map and Preliminary Results. *Remote Sensing, 11*
- 574 Argenti, F., Lapini, A., Bianchi, T., & Alparone, L. (2013). A Tutorial on Speckle Reduction in Synthetic
575 Aperture Radar Images. *IEEE Geoscience and Remote Sensing Magazine, 1*, 6-35
- 576 Baghdadi, N., Bernier, M., Gauthier, R., & Neeson, I. (2001). Evaluation of C-band SAR data for
577 wetlands mapping. *International Journal of Remote Sensing, 22*, 71-88
- 578 Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous
579 separable convolution for semantic image segmentation. In, *Proceedings of the European*
580 *conference on computer vision (ECCV)* (pp. 801-818)
- 581 Clewley, D., Whitcomb, J., Moghaddam, M., McDonald, K., Chapman, B., & Bunting, P. (2015).
582 Evaluation of ALOS PALSAR Data for High-Resolution Mapping of Vegetated Wetlands in Alaska.
583 *Remote Sensing, 7*, 7272-7297
- 584 Cohen, M.J., Creed, I.F., Alexander, L., Basu, N.B., Calhoun, A.J., Craft, C., D'Amico, E., DeKeyser, E.,
585 Fowler, L., Golden, H.E., Jawitz, J.W., Kalla, P., Kirkman, L.K., Lane, C.R., Lang, M., Leibowitz, S.G.,
586 Lewis, D.B., Marton, J., McLaughlin, D.L., Mushet, D.M., Raanan-Kiperwas, H., Rains, M.C., Smith, L.,
587 & Walls, S.C. (2016). Do geographically isolated wetlands influence landscape functions? *Proc Natl*
588 *Acad Sci U S A, 113*, 1978-1986
- 589 Dang, K.B., Nguyen, M.H., Nguyen, D.A., Phan, T.T.H., Giang, T.L., Pham, H.H., Nguyen, T.N., Tran,
590 T.T.V., & Bui, D.T. (2020). Coastal Wetland Classification with Deep U-Net Convolutional Networks
591 and Sentinel-2 Imagery: A Case Study at the Tien Yen Estuary of Vietnam. *Remote Sensing, 12*
- 592 DeLancey, E.R., Simms, J.F., Mahdianpari, M., Brisco, B., Mahoney, C., & Kariyeva, J. (2019).
593 Comparing Deep Learning and Shallow Learning for Large-Scale Wetland Classification in Alberta,
594 Canada. *Remote Sensing, 12*

595 Du, L., McCarty, G.W., Zhang, X., Lang, M.W., Vanderhoof, M.K., Li, X., Huang, C., Lee, S., & Zou, Z.
596 (2020). Mapping Forested Wetland Inundation in the Delmarva Peninsula, USA Using Deep
597 Convolutional Neural Networks. *Remote Sensing*, *12*

598 Dutt, R., Ortals, C., He, W., Curran, Z.C., Angelini, C., Canestrelli, A., & Jiang, Z. (2024). A Deep
599 Learning Approach to Segment Coastal Marsh Tidal Creek Networks from High-Resolution Aerial
600 Imagery. *Remote Sensing*, *16*

601 Frost, V.S., Stiles, J.A., Shanmugan, K.S., & Holtzman, J.C. (1982). A Model for Radar Images and Its
602 Application to Adaptive Digital Filtering of Multiplicative Noise. *IEEE Transactions on Pattern
603 Analysis and Machine Intelligence, PAMI-4*, 157-166

604 Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., & Lu, H. (2019). Dual attention network for scene
605 segmentation. In, *Proceedings of the IEEE/CVF conference on computer vision and pattern
606 recognition* (pp. 3146-3154)

607 Gallant, A. (2015). The Challenges of Remote Monitoring of Wetlands. *Remote Sensing*, *7*, 10938-
608 10950

609 Gonzalez-Perez, A., Abd-Elrahman, A., Wilkinson, B., Johnson, D.J., & Carthy, R.R. (2022). Deep and
610 Machine Learning Image Classification of Coastal Wetlands Using Unpiloted Aircraft System
611 Multispectral Images and Lidar Datasets. *Remote Sensing*, *14*

612 Gunen, M.A. (2022). Performance comparison of deep learning and machine learning methods in
613 determining wetland water areas using EuroSAT dataset. *Environ Sci Pollut Res Int*, *29*, 21092-
614 21106

615 Guo, R., Zhao, X., Zuo, G., Wang, Y., & Liang, Y. (2023). Polarimetric Synthetic Aperture Radar Image
616 Semantic Segmentation Network with Lovász-Softmax Loss Optimization. *Remote Sensing*, *15*

617 Henderson, F.M., & Lewis, A.J. (2008). Radar detection of wetland ecosystems: a review.
618 *International Journal of Remote Sensing*, *29*, 5809-5835

619 Hermosilla, T., Wulder, M.A., White, J.C., & Coops, N.C. (2022). Land cover classification in an era of
620 big and open data: Optimizing localized implementation and training data selection to improve
621 mapping outcomes. *Remote Sensing of Environment*, *268*

622 Hess, L.L., Melack, J.M., & Simonett, D.S. (1990). Radar detection of flooding beneath the forest
623 canopy: a review. *International Journal of Remote Sensing*, *11*, 1313-1325

624 Hogg, A.R., & Holland, J. (2008). An evaluation of DEMs derived from LiDAR and photogrammetry
625 for wetland mapping. *Forestry Chronicle*, *84*, 840-849

626 Hosseiny, B., Mahdianpari, M., Brisco, B., Mohammadimanesh, F., & Salehi, B. (2021). WetNet: A
627 Spatial-Temporal Ensemble Deep Learning Model for Wetland Classification Using Sentinel-1 and
628 Sentinel-2. *IEEE Transactions on Geoscience and Remote Sensing*, 1-14

629 Houlahan, J.E., Keddy, P.A., Makkay, K., & Findlay, C.S. (2006). The effects of adjacent land use on
630 wetland species richness and community composition. *Wetlands*, *26*, 79-96

631 Hu, Q., Woldt, W., Neale, C., Zhou, Y., Drahota, J., Varner, D., Bishop, A., LaGrange, T., Zhang, L., &
632 Tang, Z. (2021). Utilizing unsupervised learning, multi-view imaging, and CNN-based attention
633 facilitates cost-effective wetland mapping. *Remote Sensing of Environment*, *267*

634 Jamali, A., & Mahdianpari, M. (2022). Swin Transformer and Deep Convolutional Neural Networks
635 for Coastal Wetland Classification Using Sentinel-1, Sentinel-2, and LiDAR Data. *Remote Sensing*,
636 *14*

637 Jamali, A., Mahdianpari, M., Brisco, B., Granger, J., Mohammadimanesh, F., & Salehi, B. (2021a).

638 Deep Forest classifier for wetland mapping using the combination of Sentinel-1 and Sentinel-2
639 data. *GIScience & Remote Sensing*, *58*, 1072-1089

640 Jamali, A., Mahdianpari, M., Brisco, B., Granger, J., Mohammadimanesh, F., & Salehi, B. (2021b).
641 Wetland Mapping Using Multi-Spectral Satellite Imagery and Deep Convolutional Neural Networks:
642 A Case Study in Newfoundland and Labrador, Canada. *Canadian Journal of Remote Sensing*, *47*,
643 243-260

644 Jamali, A., Roy, S.K., & Ghamisi, P. (2023). WetMapFormer: A unified deep CNN and vision
645 transformer for complex wetland mapping. *International Journal of Applied Earth Observation and*
646 *Geoinformation*, *120*

647 Junk, W.J., An, S., Finlayson, C.M., Gopal, B., Květ, J., Mitchell, S.A., Mitsch, W.J., & Robarts, R.D.
648 (2012). Current state of knowledge regarding the world's wetlands and their future under global
649 climate change: a synthesis. *Aquatic Sciences*, *75*, 151-167

650 Kraft, A.J., Robinson, D.T., Evans, I.S., & Rooney, R.C. (2019). Concordance in wetland
651 physicochemical conditions, vegetation, and surrounding land cover is robust to data extraction
652 approach. *PLoS One*, *14*, e0216343

653 Lam, C.-N., Niculescu, S., & Bengoufa, S. (2023). Monitoring and Mapping Floods and Floodable
654 Areas in the Mekong Delta (Vietnam) Using Time-Series Sentinel-1 Images, Convolutional Neural
655 Network, Multi-Layer Perceptron, and Random Forest. *Remote Sensing*, *15*

656 Lang, M., McCarty, G., Oesterling, R., & Yeo, I.-Y. (2012). Topographic Metrics for Improved
657 Mapping of Forested Wetlands. *Wetlands*, *33*, 141-155

658 Lang, M.W., Ingebritsen, J.C., & Griffin, R.K. (2024). Status and Trends of Wetlands in the
659 Conterminous United States 2009 to 2019. In. Washington, D.C. : Department of the Interior; Fish
660 and Wildlife Service, Washington, D.C. 43 pp.

661 Lang, M.W., & Kasischke, E.S. (2008). Using C-Band Synthetic Aperture Radar Data to Monitor
662 Forested Wetland Hydrology in Maryland's Coastal Plain, USA. *IEEE Transactions on Geoscience*
663 *and Remote Sensing*, *46*, 535-546

664 Lang, M.W., Kasischke, E.S., Prince, S.D., & Pittman, K.W. (2008). Assessment of C-band synthetic
665 aperture radar data for mapping and monitoring Coastal Plain forested wetlands in the Mid-
666 Atlantic Region, U.S.A. *Remote Sensing of Environment*, *112*, 4120-4130

667 Lang, M.W., Kim, V., McCarty, G.W., Li, X., Yeo, I.-Y., Huang, C., & Du, L. (2020). Improved Detection
668 of Inundation below the Forest Canopy using Normalized LiDAR Intensity Data. *Remote Sensing*,
669 *12*

670 Li, H., Wang, C., Cui, Y., & Hodgson, M. (2021). Mapping salt marsh along coastal South Carolina
671 using U-Net. *ISPRS Journal of Photogrammetry and Remote Sensing*, *179*, 121-132

672 Li, J., Chen, W., & Touzi, R. (2014). Optimum RADARSAT-1 configurations for wetlands
673 discrimination: a case study of the Mer Bleue peat bog. *Canadian Journal of Remote Sensing*, *33*,
674 S46-S55

675 Lin, H., Zhuang, Y., Huang, Y., & Ding, X. (2023). Unpaired Speckle Extraction for SAR Despeckling.
676 *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1-14

677 Loshchilov, I., & Hutter, F. (2017). Fixing Weight Decay Regularization in Adam. *CoRR*,
678 *abs/1711.05101*

679 Lowrance, R., Altier, L.S., Newbold, J.D., Schnabel, R.R., Groffman, P.M., Denver, J.M., Correll, D.L.,
680 Gilliam, J.W., Robinson, J.L., Brinsfield, R.B., Staver, K.W., Lucas, W., & Todd, A.H. (1997). Water

681 Quality Functions of Riparian Forest Buffers in Chesapeake Bay Watersheds. *Environ Manage*, 21,
682 687-712

683 Lv, Z., Nunez, K., Brewer, E., & Runfola, D. (2023). Mapping the tidal marshes of coastal Virginia: a
684 hierarchical transfer learning approach. *GIScience & Remote Sensing*, 61

685 Mahdavi, S., Salehi, B., Moloney, C., Huang, W., & Brisco, B. (2017). Speckle filtering of Synthetic
686 Aperture Radar images using filters with object-size-adapted windows. *International Journal of*
687 *Digital Earth*, 11, 703-729

688 Mainali, K., Evans, M., Saavedra, D., Mills, E., Madsen, B., & Minnemeyer, S. (2023). Convolutional
689 neural network for high-resolution wetland mapping with open data: Variable selection and the
690 challenges of a generalizable model. *Sci Total Environ*, 861, 160622

691 Marjani, M., Mahdianpari, M., Mohammadimanesh, F., & Gill, E.W. (2024). CVTNet: A Fusion of
692 Convolutional Neural Networks and Vision Transformer for Wetland Mapping Using Sentinel-1 and
693 Sentinel-2 Satellite Data. *Remote Sensing*, 16

694 Mitsch, W.J., Bernal, B., Nahlik, A.M., Mander, Ü., Zhang, L., Anderson, C.J., Jørgensen, S.E., & Brix, H.
695 (2012). Wetlands, carbon, and climate change. *Landscape Ecology*, 28, 583-597

696 Mohammadimanesh, F., Salehi, B., Mahdianpari, M., Gill, E., & Molinier, M. (2019). A new fully
697 convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex
698 land cover ecosystem. *ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 223-236

699 O'Neil, G.L., Goodall, J.L., & Watson, L.T. (2018). Evaluating the potential for site-specific
700 modification of LiDAR DEM derivatives to improve environmental planning-scale wetland
701 identification using Random Forest classification. *Journal of Hydrology*, 559, 192-208

702 Pitman, N.C.A., Andino, J.E.G., Aulestia, M., Cerón, C.E., Neill, D.A., Palacios, W., Rivas-Torres, G.,
703 Silman, M.R., & Terborgh, J.W. (2014). Distribution and abundance of tree species in swamp forests
704 of Amazonian Ecuador. *Ecography*, 37, 902-915

705 Radman, A., Mohammadimanesh, F., & Mahdianpari, M. (2024). Wet-ConViT: A Hybrid
706 Convolutional-Transformer Model for Efficient Wetland Classification Using Satellite Data. *Remote*
707 *Sensing*, 16

708 Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J.P. (2012). An
709 assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS*
710 *Journal of Photogrammetry and Remote Sensing*, 67, 93-104

711 Scepanovic, S., Antropov, O., Laurila, P., Rauste, Y., Ignatenko, V., & Praks, J. (2021). Wide-Area
712 Land Cover Mapping With Sentinel-1 Imagery Using Deep Learning Semantic Segmentation
713 Models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14,
714 10357-10374

715 Shedlock, R.J., Denver, J.M., Hayes, M.A., Hamilton, P.A., Koterba, M.T., Bachman, L.J., Phillips, P.J., &
716 Banks, W.S. (1999). Water-quality assessment of the Delmarva Peninsula, Delaware, Maryland, and
717 Virginia; results of investigations, 1987-91. In, *Water Supply Paper*. Reston, Virginia

718 Tan, S., Zhang, X., Wang, H., Yu, L., Du, Y., Yin, J., & Wu, B. (2022). A CNN-Based Self-Supervised
719 Synthetic Aperture Radar Image Denoising Approach. *IEEE Transactions on Geoscience and*
720 *Remote Sensing*, 60, 1-15

721 Tao, C., Qi, J., Guo, M., Zhu, Q., & Li, H. (2023). Self-Supervised Remote Sensing Feature Learning:
722 Learning Paradigms, Challenges, and Future Works. *IEEE Transactions on Geoscience and Remote*
723 *Sensing*, 61, 1-26

724 Thatcher, C.A., & Lukas, V. (2021). 3D Elevation Program supports broadband internet access. In,
725 *Fact Sheet* (p. 2). Reston, VA

726 Tiner, R.W. (2003). Geographically isolated wetlands of the United States. *Wetlands*, 23, 494-516

727 Wang, J., Yang, X., Wang, Z., Cheng, H., Kang, J., Tang, H., Li, Y., Bian, Z., & Bai, Z. (2022).
728 Consistency Analysis and Accuracy Assessment of Three Global Ten-Meter Land Cover Products in
729 Rocky Desertification Region—A Case Study of Southwest China. *ISPRS International Journal of*
730 *Geo-Information*, 11

731 Whitcomb, J., Moghaddam, M., McDonald, K., Kellndorfer, J., & Podest, E. (2014). Mapping
732 vegetated wetlands of Alaska using L-band radar satellite imagery. *Canadian Journal of Remote*
733 *Sensing*, 35, 54-72

734 Yuan, Z., Hao, C., Zhou, R., Chen, J., Yu, M., Zhang, W., Wang, H., & Sun, X. (2023). Efficient and
735 Controllable Remote Sensing Fake Sample Generation Based on Diffusion Model. *IEEE Transactions*
736 *on Geoscience and Remote Sensing*, 61, 1-12

737 Zhang, X., & Han, L. (2023). A generic self-supervised learning (SSL) framework for representation
738 learning from spectra-spatial feature of unlabeled remote sensing imagery. *arXiv preprint*
739 *arXiv:2306.15836*

740 Zhang, X., Han, L., Han, L., & Zhu, L. (2020). How Well Do Deep Learning-Based Methods for Land
741 Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery.
742 *Remote Sensing*, 12

743