Hard-Constrained Hopfield Neural Network for Subpixel Mapping

Chengyuan Zhang, Qunming Wang, Peter M. Atkinson

***Abstract*—Subpixel mapping (SPM) can address the mixed pixel problem by producing land cover maps at a finer spatial resolution than the input images. The Hopfield neural network (HNN) method has shown great advantages in SPM and various extended versions have been developed recently. However, a longstanding issue in the HNN, especially in the multi-class scenario, is its tendency to fall into local optima with vanished gradients, failing to push subpixels to the hard class label of 0 or 1. This can lead to great uncertainties in determining hard class labels, and moreover, the disappearance of many small-sized land cover features and spatial details. In this paper, we proposed a hard-constrained HNN (H-HNN) model that introduces hard label-based constraints at both the subpixel and coarse pixel scales. These constraints aim to increase the accuracy of SPM by guiding the optimization process fully towards obtaining hard classification maps at the subpixel level. Experimental evaluations against benchmark methods demonstrated the effectiveness of the H-HNN. The findings reveal that the H-HNN method is a general and robust alternative to the HNN, which can increase the overall accuracy of the SPM results by about 1%. Additionally, the H-HNN can effectively reduce the uncertainties by predicting more accurate hard class labels and coarse proportions (with root mean square error (RMSE) of the coarse proportions decreased by about 0.015).**

***Index Terms*—Mixed pixel, subpixel mapping (SPM), Hopfield neural network (HNN).**

I. Introduction

Land use and land cover mapping is required for understanding and managing Earth surface dynamics, making it crucial for environmental research and decision-making processes [1], [2]. Remote sensing satellite images enable the characterization of diverse land cover types synoptically, with a large field of view. However, interpretation of remote sensing data encounters a significant challenge known as the mixed pixel problem [3], [4], where pixels may contain a mixture of different land cover types due to the inherent limitations of sensor resolution and the intricate heterogeneity of land surfaces. Conventional hard classification, which assigns each pixel to a single class label, can lead to inaccurate interpretation and the disappearance of small-sized targets within the mixed pixels [5].

Spectral unmixing, or soft classification [6], has emerged to address the mixed pixel problem by decomposing the spectrum of every pixel into class proportions (also referred to as fractions or abundance), but these provide no information on how the land cover is distributed within each pixel [7]. Subpixel mapping (SPM), also termed super-resolution mapping, has been proposed to address this problem by predicting the spatial distribution of land cover classes within the mixed pixels. Specifically, by SPM, each mixed pixel is subdivided into $s \times s$ subpixels and class labels are allocated for each subpixel. Thus, the native spatial resolution of remote sensing images is broken down into a finer spatial resolution [8]. Due to the trade-off between spatial and temporal resolution, it is challenging to acquire temporally frequent time-series images with fine spatial resolution. In this scenario, SPM is an effective solution to obtain fine spatial resolution land cover maps from temporally frequent but spatially coarse time-series. Up to now, SPM has proven effective in various fields, such as surface water body mapping [9], [10], forest fire detection [11] and urban mapping [12].

Taking a set of coarse proportion images as input, SPM produces fine spatial resolution land cover maps, posing an inherently ill-posed problem. Over the decades, various SPM algorithms have been developed [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. In the absence of auxiliary data, the SPM methods typically aim to maximize the spatial attraction between subpixels of the same land cover class. With a growing availability of multi-source data, researchers have exploited various sources of auxiliary information to enhance SPM recently, such as temporally adjacent images [22], [24], edges of objects [25], [26], digital evaluation models (DEM) [27], [28], open-sourced land cover products [29], training data [12], [23] and statistical land cover indices [30]. Considering the mechanisms of these various methods, two basic types of SPM methods were summarized by Wang et al. [31]. The first type conducts initialization by assigning each subpixel to a certain class (binary indicators with 1 or 0 to represent belonging to the class or not) based on the coherence constraint imposed by the coarse proportions, followed by optimization through exchanging the labels of subpixels (i.e., the number of subpixels for each class is fixed), such as the pixel-swapping (PSA) method [15], [32]. The second type, known as the soft-then-hard SPM (STHSPM) approach, estimates soft class values (between 0 and 1 to represent the possibility of belonging to the class) for all classes at the subpixel scale and then allocates class labels to the subpixels. Typical methods include the subpixel/pixel spatial attraction model [16] and radial basis function (RBF). The advanced deep learning-based methods [12], [23] also adopt the strategy of predicting soft class possibilities. Generally, the SPM process is obedient to the coherence constraint imposed by the

This work was supported by the National Natural Science Foundation of China under Grants 42222108 and 42171345. (*Corresponding author: Q. Wang*.)

C. Zhang is with the Key Laboratory of Ethnic Language Intelligent Analysis and Security Governance of MOE, Minzu University of China, Beijing 100081, China, and also with the College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China

Q. Wang is with the College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China (e-mail: wqm11111@126.com).

P.M. Atkinson is with the Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK, and also with Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK.

coarse proportions. However, spectral unmixing remains a challenging issue, leading to prediction of the coarse proportions with inevitable uncertainties [33], [34], which subsequently affect the results of both types of SPM methods greatly.

As another typical SPM method, the Hopfield neural network (HNN) [35] method is performed with an iterative optimization process by pushing the initialized subpixel class labels (soft values between 0 and 1) towards the hard label of 0 or 1 for each class, which does not perfectly honor the coarse proportions. The HNN has gained great attention and various extended versions have been developed in recent years. Specifically, the HNN was revised by taking account of anisotropic gradient [36] and self-similarity [37] in the coarse proportions. In addition, the HNN was combined with a genetic algorithm [38] and contouring method [39]. Numerical methods were also developed to speed up the iteration of the HNN [40]. Moreover, ancillary data were integrated into the HNN, including a temporally adjacent fine spatial resolution land cover map [41], [42], panchromatic images [43], fused images [44], LIDAR data [45] and shifted images [46]. Overall, the HNN method demonstrates unique advantages over other SPM methods appreciating its flexibility in maintaining the class proportions, but not honoring them perfectly [47], allowing for the alleviation of errors in the original coarse proportions to some extent. However, there has been a long standing, fundamental key issue in the HNN. Specifically, the final subpixel class labels after thousands of iterations may not be sufficiently close to the hard class label of 0 or 1, especially for the multi-class scenario.

The HNN employs a multi-class constraint of 'sum to one' to accommodate the layers of all land cover classes in a multi-class scenario [13]. Specifically, for each subpixel, the sum of soft attributes (i.e., class labels) across all layers (denoting different land cover classes) sums to one. While this condition is reasonable and needed, some concerns may arise. Specifically, the HNN result may fall into a steady state, but still lack specific signs (desired to be 1 or 0 to indicate whether belonging to the class or not) for determining the hard class labels of subpixels. Moreover, the spatial attraction term in the HNN, activated by the average signals of surrounding subpixels for each class, may encounter gradient disappearance, resulting in great under-estimation of small-sized patches (e.g., smaller than 0.5 pixel area) [48]. Consequently, significant uncertainties persist in the optimization of the class labels, leading to the disappearance of small-sized objects or spatial details along the class boundaries. However, these small-sized objects can be crucial in land cover detection. For example, small-sized forest fires and urban street trees, may be presented by weak signals in coarse data, resulting in their elimination in the final land cover map produced by HNN.

To address the above challenge, in this paper, we proposed the hard-constrained HNN (H-HNN) model, which imposes hard label-based constraints on the class label estimation process. Specifically, a 'one and only one' constraint and reinforced proportion constraint were proposed. The former means that each subpixel belongs to only one class, suggesting that the class label should be 1 for a certain class, but 0 for all the other classes. The latter means that based on the new 'one and only one' constraint, the area constraint needs to be reinforced correspondingly to minimize the difference between the estimated area coverage and input class proportion. It should be noted that the proposed H-HNN aims to reduce uncertainties in the optimizing process of the original HNN while retaining the inherent advantages in alleviating the impact of errors in the coarse proportions. The contributions are summarized as two aspects.

1) Theoretically, we presented a new insight to mitigate the ill-posed nature of HNN-based SPM by further steering the prediction towards achieving a hard-classified map at the subpixel scale, which is aligned with the essential task of SPM.
2) Technically, we introduced the H-HNN model by considering the 'one and only one' constraint at the subpixel scale and reinforced area constraint at the coarse pixel scale.

The remainder of this paper is as follows. Section II introduces briefly the original HNN model and highlights the uncertainties in predicting hard class labels. Following this, the proposed H-HNN is presented, along with the two new hard label-based constraints. Section III presents the experimental results for validation of H-HNN. Section IV engages in open issues of the proposed H-HNN. Finally, Section V summarizes the key findings of this study.

II. Methods

A. HNN-based SPM

The HNN [49] is a type of artificial neural network distinct from learning-based networks, which is operated without the need for training data. The basic architecture of HNN-based SPM is given below.

1) *Neurons in HNN:* With a coarse proportion image sized $M \times N \times K$ (M, N and K denote the number of rows, columns and land cover classes, respectively) as input, the HNN network is structured into K layers, with each representing a land cover class. Each layer comprises $M \times s \times N \times s$ neurons (i.e., denoting the subpixel labels), where s denotes the zoom factor.

The neuron located at row i and column j has an input signal u_{ijk} and an output signal v_{ijk}. The value of v_{ijk} represents the probability that the subpixel at (i, j) belongs to class k, which ranges from 0 to 1. The input u_{ijk} and output v_{ijk} are transformed by the activation function

$$v_{ijk} = \frac{1}{2}[1 + \tanh(\lambda u_{ijk})] \tag{1}$$

where λ determines the steepness of the tanh function.

2) *Feedback of HNN:* As a recurrent network, the output of each neuron v_{kij} influences the activation of neurons in subsequent iterations, as shown in Fig. 1. For example, the input u_{ijk} for the next iteration becomes

$$u_{ijk}(t+dt) = u_{ijk}(t) + \frac{du_{ijk}(t)}{dt}dt = u_{ijk}(t) - \frac{dE_{ijk}(t)}{dv} \tag{2}$$

where dt is the iteration step, and $\frac{du_{ijk}(t)}{dt}$ represents the change in energy, which can be calculated as the gradient descent of the energy function $\frac{dE_{ijk}(t)}{dv}$. The network reaches a stable solution

when the total energy is minimized. The energy function is defined as

$$E_{ijk} = \sum_i \sum_j (w_1 G_{ijk} + w_2 P_{ijk}) + \sum_k w_3 M_{ijk} \quad (3)$$

in which G_{ijk} is a spatial clustering function aimed at maximizing spatial correlation, P_{ijk} is the coarse proportion constraint and M_{ijk} represents the multi-class constraint (i.e., 'sum to one'), with w_1, w_2 and w_3 being the weights of each term. The neurons are interconnected and influence each other through the energy function within the spatial support extents of each term. Further elaborations can be found in [13], [49]. In the following, we focus mainly on the uncertainty associated with the energy function.

3) *Connections between neurons:* The connections between the target neuron v_{ijk} and each term of the energy function in Eq. (3) are depicted in Fig. 1 based on their spatial supports.

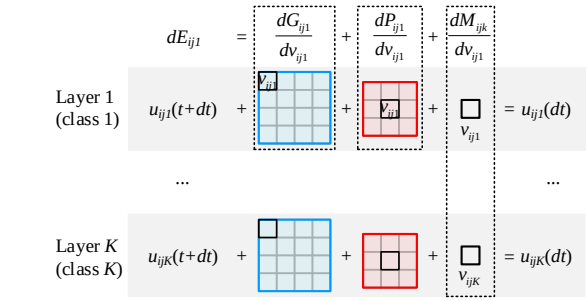

Fig. 1. Illustration of spatial connections between neurons (subpixel labels) of the energy function system in HNN for the spatial clustering function, proportion constraint and multi-class constraint.

More specifically, the spatial clustering function is activated by computing the average of the eight neighbors of v_{ijk}, which is expressed as

$$dG_{ijk} = dG1_{ijk} + dG2_{ijk}$$

$$\frac{dG1_{ijk}}{dv_{ijk}} = \frac{1}{2}\left(1 + \tanh\left(\frac{1}{8}\sum_{\substack{d=i-1 \\ d\neq i}}^{i+1}\sum_{\substack{e=j-1 \\ e\neq j}}^{j+1} v_{dek} - 0.5\right)\lambda\right)\left(v_{ijk} - 1\right) \quad (4)$$

$$\frac{dG2_{ijk}}{dv_{ijk}} = \frac{1}{2}\left(1 - \tanh\left(\frac{1}{8}\sum_{\substack{d=i-1 \\ d\neq i}}^{i+1}\sum_{\substack{e=j-1 \\ e\neq j}}^{j+1} v_{dek} - 0.5\right)\lambda\right) v_{ijk}$$

in which the first term $dG1_{ijk}$ pushes v_{ijk} towards 1 when the average of the eight neighbors exceeds 0.5, whereas the second term encourages v_{ijk} to approach 0 when the average is below 0.5.

To ensure class proportion consistency for each pixel, the coarse proportion constraint is activated by calculating the difference between the average of the $s \times s$ subpixel labels within the coarse pixel and the corresponding coarse proportion, which is formulated as

$$\frac{dP_{ijk}}{dv_{ijk}} = \frac{1}{2 \times s \times s}\sum_{d=x\times s+1}^{x\times s+s}\sum_{e=y\times s+1}^{y\times s+s}[1 + \tanh(v_{dek} - 0.5)\lambda] - F_{xyk} \quad (5)$$

where (x, y) is the spatial location of the coarse pixel that the subpixel at (i, j) falls within and F_{xyk} is the coarse proportion of class k for the coarse pixel.

To ensure that the sum of the outputs for each neuron equals to 1, often referred to as 'sum to one', the multi-class constraint for each subpixel is given by

$$\frac{dM_{ijk}}{dv_{ijk}} = \left(\sum_{k=1}^{K} v_{ijk}\right) - 1 \quad (6)$$

B. Uncertainties in predicting hard class labels in original HNN

In the original HNN model, as shown in Fig. 1, the energy delivery between different layers relies primarily on the multi-class constraint M_{ijk}. In a scenario with multiple land cover classes, when all neurons for a subpixel have values that are very similar (e.g., equal to one divided by the number of classes), the 'sum to one' condition can be met for that subpixel. However, in the final output of each neuron in the HNN, the values may then fail to be decisively pushed towards either 0 or 1 so that only one of the subpixel neurons is equal to one and the others are zeros, resulting in an inability to allocate a single class to the subpixel.

Furthermore, the spatial clustering function G_{ijk} and the coarse proportion constraint P_{ijk} are operated within the target layer corresponding to the observed land cover class. That is, the neuron in each layer is pushed by the spatial clustering function. Specifically, a neuron v_{ijk} in the layer of class k requires a positive gradient of the first term $dG1_{ijk}$ in the spatial energy function Eq. (4) to push it towards 1. However, if the neuron is located at the center of a coarse pixel with class proportion less than 0.5, leading to the average of neighbors being less than 0.5, the term $dG1_{ijk}$ would receive a small gradient, and the term $dG2_{ijk}$ would be activated to push v_{ijk} towards 0, resulting in a disappearance of the class.

To illustrate the inadequacy of relying solely on the multi-class constraint in the HNN, Fig. 2 presents a scenario with four land cover classes, where two possible cases are listed. It should be noted that both cases satisfy the multi-class constraint of 'sum to one' and the coarse proportion constraint. However, the second case also satisfies the hard label-based constraints. The resulting land cover maps of the two cases exhibit significant differences. Specifically, in the first case, there is considerable uncertainty in determining the appropriate class for neurons with soft values of 0.25 or 0.5. In contrast, the second case provides a more satisfactory prediction and is less ambiguous because each subpixel is hard-classified. This comparison highlights the importance of accounting for hard class label-based constraints in HNN-based SPM, as they can impact significantly the reliability of the final results.

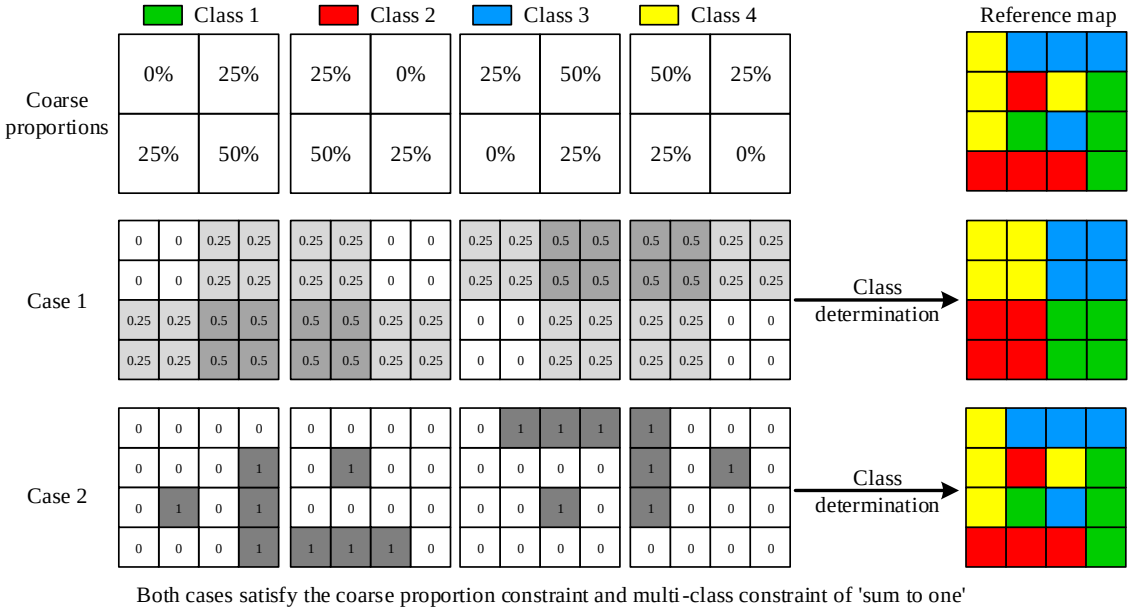

Fig. 2. Two possible cases depicting class labels optimized by the HNN where each fulfills both the coarse proportion constraint and the multi-class constraint of 'sum to one'. However, Case 2 also satisfies the hard-label based constraint of 'one and only one' concurrently, which involves less uncertainty in determining the final hard class labels in SPM.

C. H-HNN

To prevent the occurrence of stable, but sub-optimal conditions in the optimization of class labels (i.e., the predicted class label is far from the ideal value of 0 and 1), this paper introduces a H-HNN with two new constraints. The energy function of H-HNN is expressed as

$$E_{ijk} = \sum_i \sum_j (w_1 G_{ijk} + w_2 P_{ijk} + w_3 P'_{ijk}) + \sum_k (w_4 M_{ijk} + w_5 M'_{ijk}) \quad (7)$$

in which w_i represents the weight for each term, M'_{ijk} and P'_{ijk} denote the proposed 'one and only one' constraint at the subpixel scale and the reinforced proportion constraint at the coarse pixel scale, respectively.

1) The 'one and only one' constraint at the subpixel scale. We introduced the 'one and only one' constraint to strengthen the multi-class constraint between different layers in the HNN and, more importantly, to push the predicted class labels of subpixels fully towards 0 or 1. Furthermore, as each subpixel belongs to only one class, the class label should be 1 for only one class, but 0 for the other classes. Accordingly, we proposed a square-based constraint expressed as

$$dM'_{ijk} = 1 - \left(\sum_{k=1}^{K} v_{ijk}^2\right) \quad (8)$$

Since each v_{ijk} ranges between 0 and 1, coupled with the multi-class constraint in Eq. (6), the new constraint described in Eq. (8) will yield a zero energy only when the neuron has solely one output signal equal to 1 for one class, while the rest are equal to 0. That is,

$$\min \begin{pmatrix} \sum_{k=1}^{K} v_{ijk} - 1 \\ \sum_{k=1}^{K} v_{ijk}^2 - 1 \end{pmatrix} \Rightarrow \sum_{\substack{k=1 \\ k \neq k_0}}^{K} v_{ijk} = 0 \text{ and } v_{ijk_0} = 1 \quad (9)$$

Since the maximum value of dM'_{ijk} is $1 - \frac{1}{K}$ (occurs when the values of v_{ijk} for all K classes are equal), this term is normalized by dividing the maximum value.

Algorithm 1 H-HNN

Inputs: $\{F_k(V) | \ k=1,2, \ldots,K\}$, s, λ, w_1, w_2, w_3, w_4, w_5, *num_iteration*

```
Calculate initial u_ijk and v_ijk of all neurons.
for itr=1:num_iteration
  for k=1: K
    for i=1: s
      for j=1: s
        Calculation of dP_k and dM_ijk using (5), (6);
        Calculation of the proposed two constraints dP'_k and
        dM'_ijk using (8) and (10);
        Calculation of dG_ijk from spatial neighbors using (4);
        For the neuron v_ijk, calculate E_ijk of all terms with
        weights using (7) and update u_ijk using (2);
        Update v_ijk using (1) as the input of next iteration.
      end
    end
  end
end
```

Outputs: $\{v_{ijk} | \ i=1,\ldots,s; \ j=1,\ldots,s; \ k=1,2,\ldots K\}$

2) The Reinforced proportion constraint at the coarse pixel scale. With a relatively low area proportion (e.g., lower than 0.5 [48]), the spatial attraction term *dG1* in Eq. (4) would not be activated, but *dG2* instead. Consequently, these land cover types with small coarse proportions would eventually disappear.

Hence, a reinforced proportion constraint is necessary. Accordingly, we proposed a square-based proportion constraint:

$$dP'_k = s^2 \times F_k(V) - (\sum_{j=1}^{s\times s} v_{ijk}^2) \quad (10)$$

The reinforced proportion constraint is also expected to be minimized to 0. The maximum value of dP'_k is $F_k(V) - F_k^2(V)$. Therefore, this term is normalized by dividing the corresponding maximum value $F_k(V) - F_k^2(V)$.

The pseudocode in Algorithm 1 shows the optimization process of H-HNN. It can be seen that the proposed two terms are activated for all $s \times s$ subpixels in the mixed pixels.

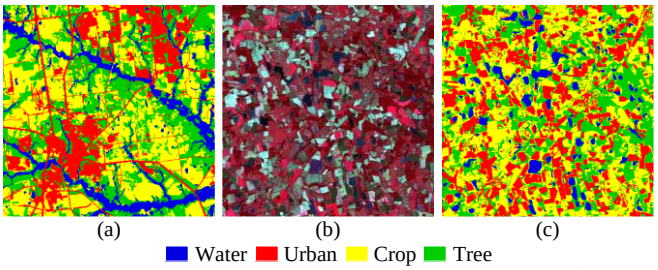

Fig. 3. The two datasets used in the experiments. (a) NLCD map of the South Carolina dataset (30 m). (b) False color image of the multispectral dataset (10 m). (c) Non-linear kernal SVM-derived land cover map of the France dataset (10 m).

III. Experiments

The proposed H-HNN method was evaluated using two datasets. The first dataset was obtained from the National Land Cover Database 2001 (NLCD 2001) with a spatial resolution of 30 m. It is located in South Carolina and has a size of 480×480 pixels (Fig. 3(a)). The second dataset was acquired from Sentinel-2 MSI images on Octorber 5th, 2023, with a spatial resolution of 10 m, covering a spatial size of 960×960 pixels in France (Fig. 3(b)). The multispectral image was classified to a land cover map by a support vector machine (SVM) with a non-linear-based kernel (i.e., radial basis function) (Fig. 3(c)). Both land cover maps contain four land cover classes: water, urban, crop and tree.

In SPM, the input coarse proportions were simulated by degrading the fine spatial resolution land cover maps in Fig.3(a) and Fig. 3(c) with zoom factors of 3, 4, 6 and 8. The original fine spatial resolution land cover maps cover the same spatial extent as the coarse images, and serve as references for both qualitative and quantitative evaluation of the SPM methods [8]. The performance of the proposed H-HNN was compared to that of the RBF, PSA and original HNN methods. It should be noted that both RBF and PSA adhere strictly to the coarse proportions, while HNN and H-HNN do not compeletely preserve them. Addtionally, the impact of proportion errors was also investigated.

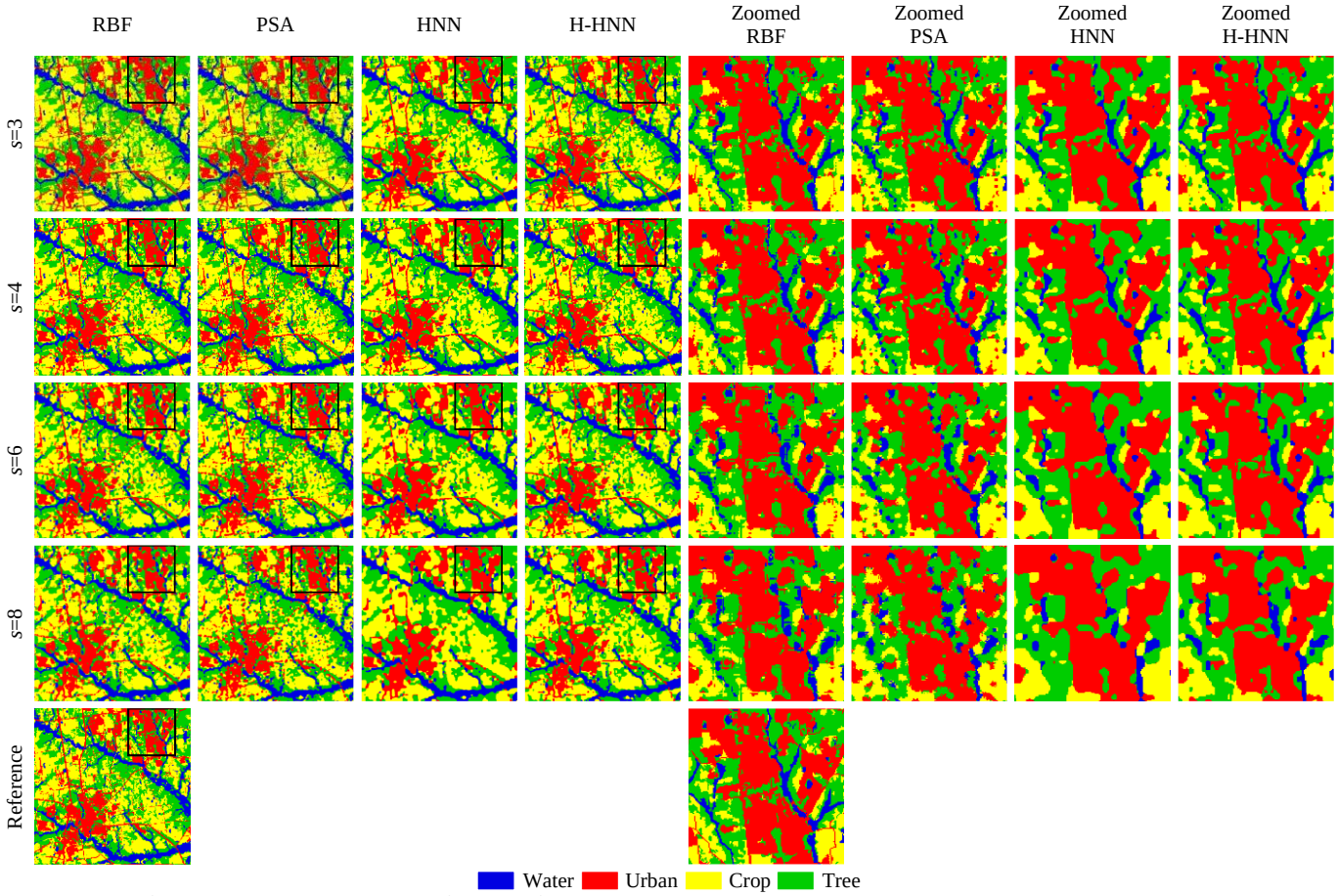

Fig. 4. SPM results for the South Carolina dataset with four zoom factors.

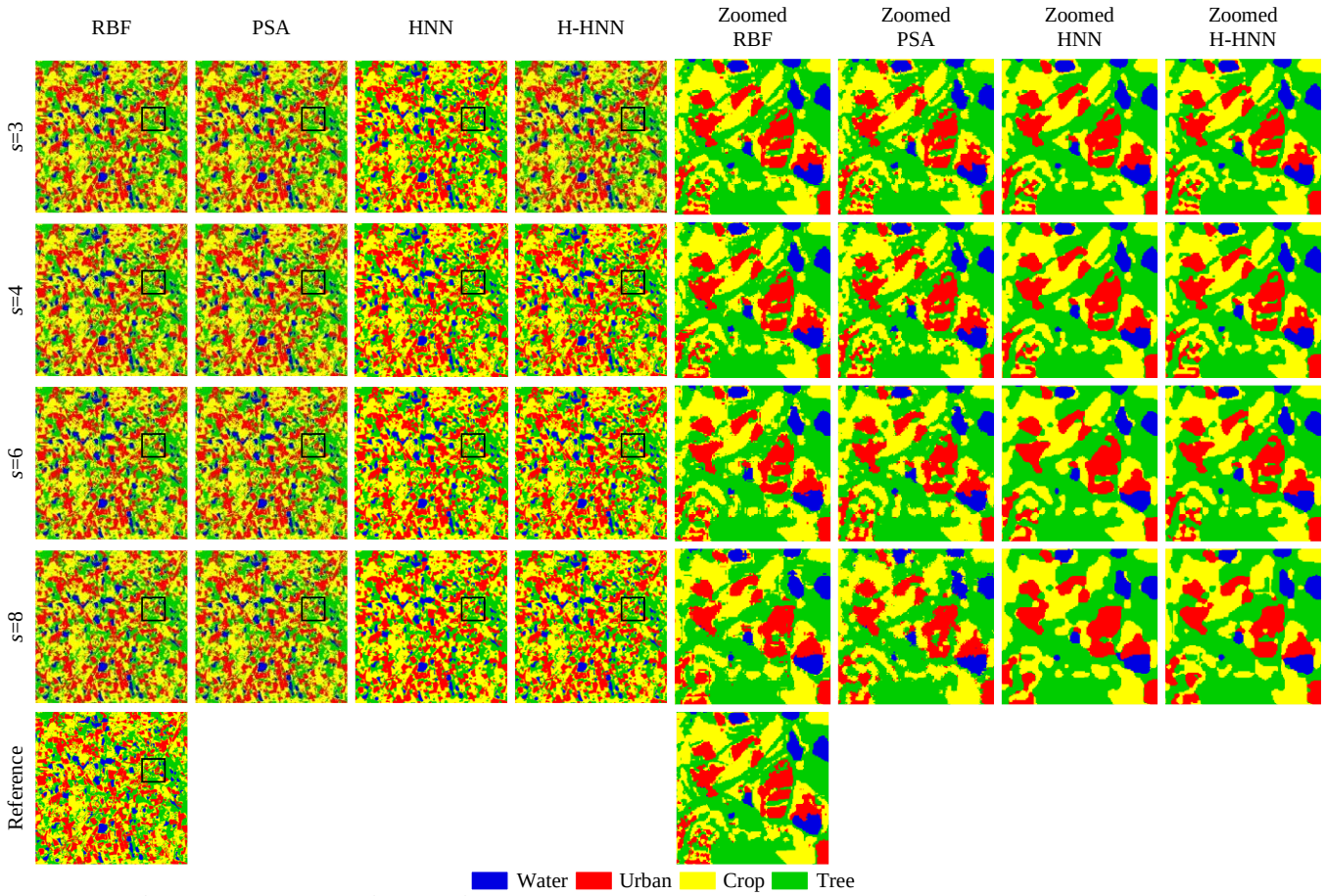

Fig. 5. SPM results for the France dataset with four zoom factors.

In the experiments, the parameters for all methods were set empirically. The window size for the spatial attraction term of the RBF was set to 3×3, 3×3, 5×5 and 5×5 coarse pixels for zoom factors of 3, 4, 6 and 8, respectively, and that for the PSA was 3×3, 3×3, 5×5 and 5×5 subpixels for the zoom factors of 3, 4, 6 and 8, respectively. The weights in the energy function for HNN and H-HNN were set equally to 1. The number of iteration was set to 1000, the steepness parameter λ was set to 10 and the iteration step *dt* was set to 0.001.

A. Qualitative and quantitative assessment

Figs. 4 and 5 depict the SPM results of the four methods (RBF, PSA, original HNN and H-HNN) for the two datasets with different zoom factors. The results show that the HNN and H-HNN demonstrate smoother boundaries compared to RBF and PSA, albeit at the expense of some loss in elongated structures. Notably, the H-HNN tends to produce visually more reliable results than the HNN, especially for small but continuous land cover features that disapear in the HNN results.

The overall accuracies (OA) of the four methods for the two datasets are listed in Table 1. It is that evident that the OA values of the four methods decrease with a larger zoom factor. However, the OA of the proposed H-HNN is generally the largest among the four methods under various circumstances, which is consistent with the visual inspection in Figs. 4 and 5. For example, for the South Carolina map, the OA gains of the H-HNN over HNN are 1.57%, 1.23%, 0.83%, and 0.97% for the zoom factors of 3, 4, 6 and 8, respectively.

The producers' accuracies (PA) of the four methods for the two datasets are provided in Fig. 6. The PA values of the H-HNN are generally the largest among the four methods. For example, for the South Carolina dataset, the gains in PA of the proposed H-HNN over HNN are 2.63%, 2.35%, 2.29%, 1.76%, 1.35% and 0.91% for the water, urban, crop and tree classes, respectively, with a zoom factor of 4.

Table 1 OA (%) of SPM results for the two datasets

		RBF	PSA	HNN	H-HNN
South Carolina	*s*=3	91.12	89.45	90.03	91.59
	s=4	87.72	86.36	87.40	88.63
	s=6	82.09	80.83	82.54	83.37
	s=8	77.74	75.22	78.24	79.21
France	*s*=3	96.17	94.61	95.48	96.62
	s=4	93.79	92.71	93.47	94.53
	s=6	89.98	88.78	89.71	90.45
	s=8	86.35	84.00	86.26	86.77

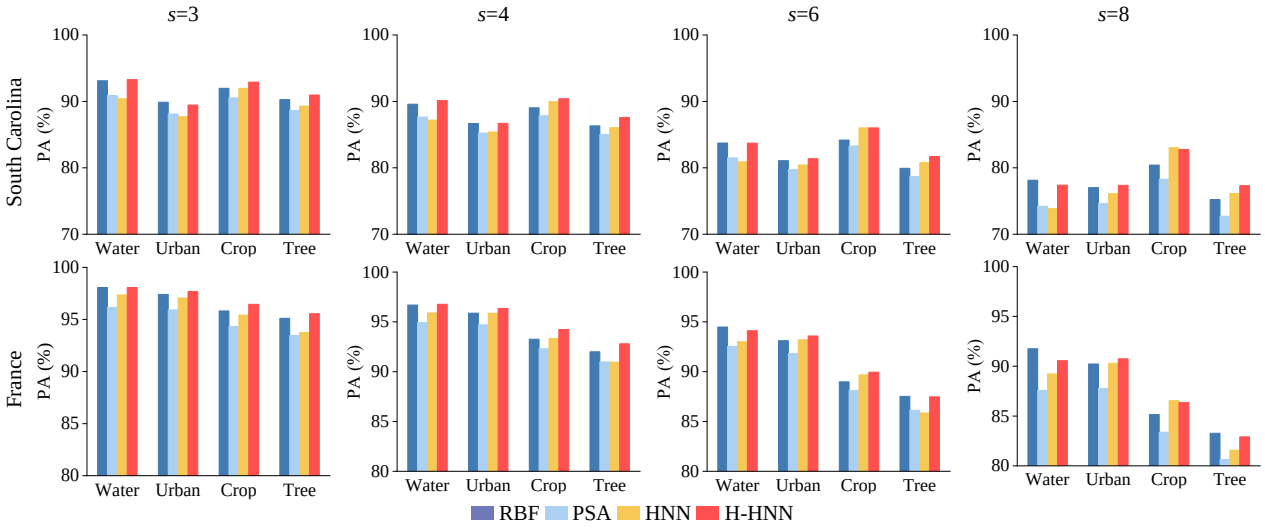

Fig. 6. Producer's accuracy (PA) of SPM results for the two datasets.

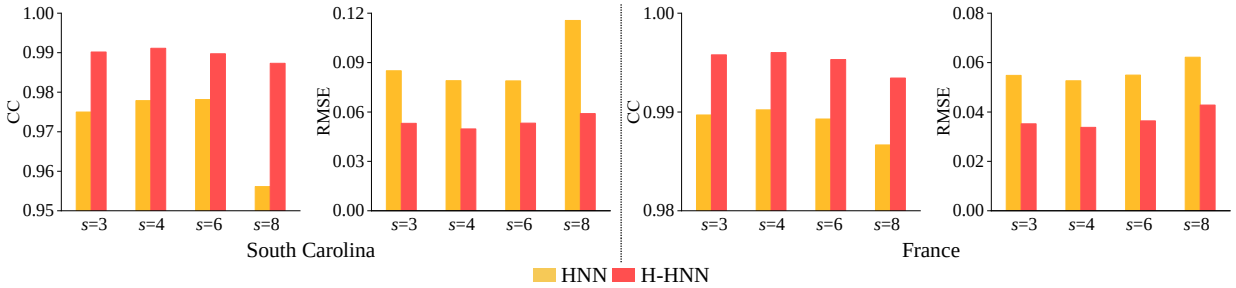

Fig. 7. Assessments of SPM results based on the CC and RMSE between the degraded SPM results and original error-free coarse proportions.

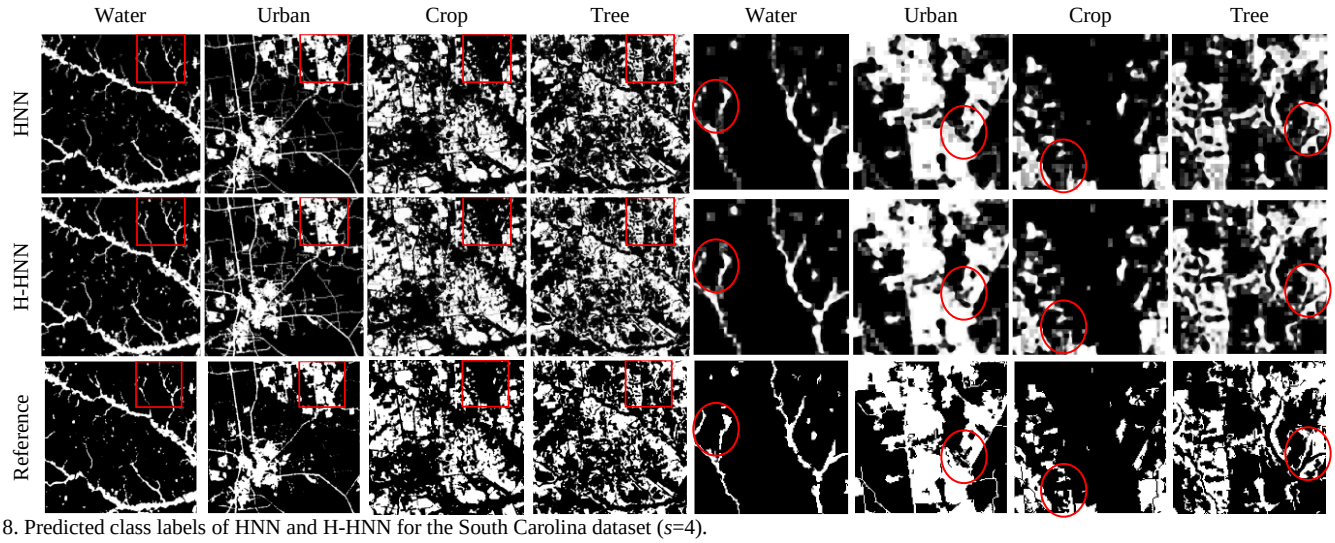

Fig. 8. Predicted class labels of HNN and H-HNN for the South Carolina dataset (s=4).

B. Accuracy assessment based on coarse proportion reproduction

The reliability of reproducing the coarse proportions by the SPM predictions of HNN and H-HNN is evaluated. Specifically, the degraded SPM results are compared with the ideal error-free coarse proportions using the correlation coefficient (CC) and RMSE, as shown in Fig. 7. Note that the results of PSA and RBF are not included because these two methods maintain strictly the coarse proportions, and any errors are retained in the results directly. It is observed that the H-HNN produces smaller RMSE values and larger CC values for the two datasets compared to the original HNN, suggesting the stronger ability of the H-HNN to preserve the original coarse proportions.

C. Comparison of predicted class labels

To illustrate the differences in predicting class labels between the H-HNN and original HNN, the results for the South Carolina dataset with a zoom factor of 4 are shown in Fig. 8. It can be seen that the HNN result exhibits gray blocks, indicating that subpixel labels within the coarse pixels failed to be pushed to hard class labels. In the H-HNN result, the small-sized patches are reproduced more satisfactorily and, more importantly, the

occurrence of blocky artifacts is reduced obviously compared to the HNN. The CC and RMSE between the predicted and ideal hard labels for each class are listed in Table 2. Generally, the proposed H-HNN exhibits a smaller RMSE and larger CC than the HNN, indicating that the H-HNN can generate more accurate hard class labels.

Table 2 The RMSE and CC between the predicted and ideal hard class labels for the South Carolina dataset (s=4)

	RMSE		CC	
	HNN	H-HNN	HNN	H-HNN
Water	0.1396	0.1355	0.9025	0.9105
Urban	0.1997	0.2017	0.8713	0.8725
Agriculture	0.2358	0.2340	0.8686	0.8741
Forest	0.2588	0.2554	0.8345	0.8445

D. Impact of errors in coarse proportions

The impact of proportion errors was analyzed using the France dataset. Specifically, a fine spatial resolution land cover map, as shown in Fig. 9(a), was obtained by applying a SVM classifier with linear kernel to the 10 m Sentinel-2 image in Fig. 3(b). The non-linear kernel SVM-derived land cover map is shown in Fig. 9(b). Generally, non-linear SVM classifiers tend to produce more accurate results than linear-based SVM classifiers [50], [51]. Assuming the map derived from the non-linear SVM as reference, the differences between Fig. 9(a) (i.e., assumed to contain spatially correlated errors) and Fig. 9(b) were considered to simulate misclassified pixels in the 10 m land cover map, as depicted in Fig. 9(c). Coarse proportion images with errors were then simulated by degrading the linear kernel SVM-derived land cover map with zoom factors of 3, 4, 6 and 8. The errors in the coarse proportions were evaluated in terms of the RMSE by referring to the non-linear SVM-derived coarse proportions (i.e., assumed as error-free coarse proportions) for each zoom factor, as listed in Table 3. Then, the coarse proportion images with errors were used as input for the four SPM methods, including RBF, PSA, HNN and H-HNN. The SPM results of the four methods for zoom factors of 4 and 6 are depicted in Fig. 10.

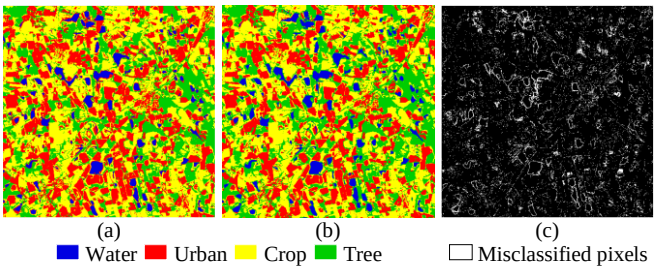

Fig. 9. SVM-derived land cover maps for the France dataset. (a) Linear kernel SVM-derived land cover map (i.e., considered as simulated map with classification errors). (b) Non-linear kernel SVM-derived land cover map (i.e., assumed as reference). (c) Locations of misclassified pixels in the linear SVM-derived land cover map.

Table 3 Evaluation (in terms of RMSE) of simulated errors in the coarse proportions (i.e., comparing the coarse proportions by degrading Fig. 9(a) and Fig. 9(b))

	Water	Urban	Crop	Tree	Mean
s=3	0.0724	0.1349	0.1150	0.1330	0.1138
s=4	0.0654	0.1215	0.0988	0.1189	0.1012
s=6	0.0563	0.1049	0.0811	0.1023	0.0862
s=8	0.0506	0.0944	0.0698	0.0918	0.0767

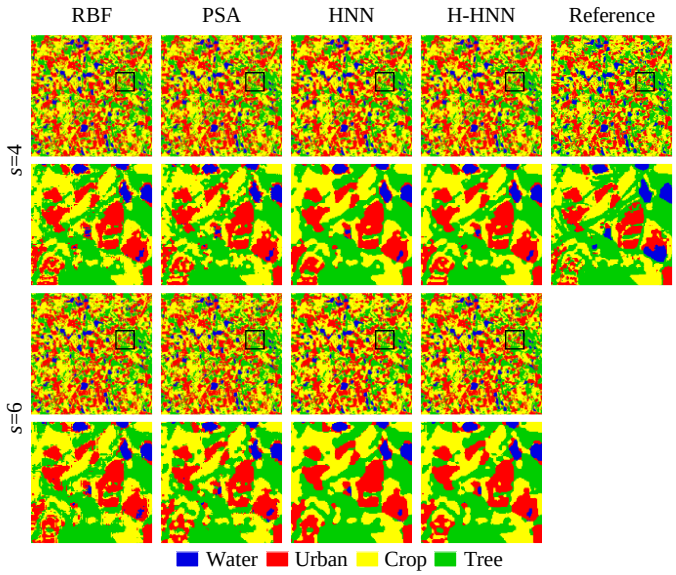

Fig. 10. SPM results with errors in the input coarse proportions for the France dataset (s=4 and s=6).

Table 4 OA (%) of SPM results with errors in coarse proportions for the France dataset

	RBF	PSA	HNN	H-HNN
s=3	90.70	89.75	90.68	91.30
s=4	89.24	88.29	89.43	90.06
s=6	86.11	85.23	86.62	87.10
s=8	83.42	81.25	83.79	84.21

When errors exist in the coarse proprotions, the PSA and RBF results exhibit plentiful sporadic artifacts along the boudaries. Conversely, the results of the HNN and H-HNN are less afffected by these errors since the HNN-based (HNN and H-HNN) methods do not preserve strictly the coarse proportions. It is worth noting that the proposed H-HNN retains the ability to eliminate small errors in the coarse proportions. Moreover, the H-HNN tends to produce visually more reliable results than the HNN, especially for small but continuous land cover features that disapear in the HNN results. The OA values of the four methods for the case with errors in coarse proportions are shown in Table 4. It can be seen that the proposed H-HNN produces the largest OA among the four methods for all cases, which is consistent with the visual inspection in Fig. 10.

IV. Discussion

A. Ability to reserve small-sized patches of H-HNN

To assess the accuracy of restoring small-sized patches, the accuracy for land cover classes with coarse proportions less than 0.5 was evaluated for both the H-HNN and HNN on the two datasets, as depicted in Fig. 11. The results indicate that the H-HNN achieves larger accuracy for small-sized patches of land cover classes. For example, for the South Carolina dataset, with a zoom factor of 3, the gains in OA of the proposed H-HNN over the HNN are 3.90%, 2.36%, 3.02% and 3.39% for the four land cover classes, respectively. This advantage can be attributed to the stricter nature of the proposed hard label-based proportion constraint compared to the original coarse proportion constraint in the HNN.

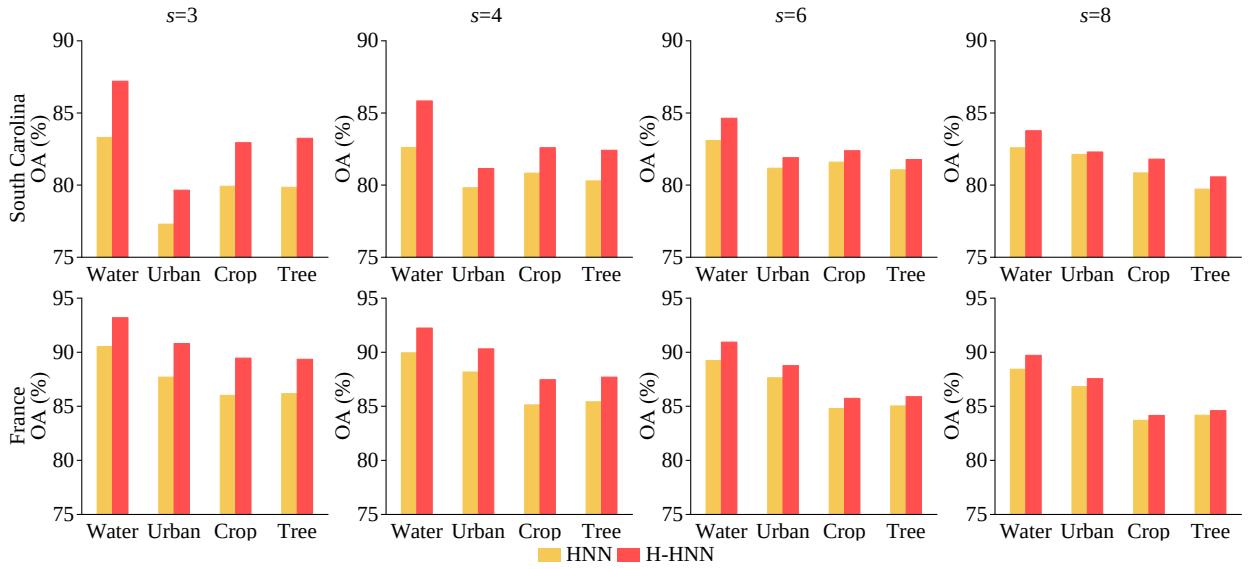

Fig. 11. OA (%) of small-sized land cover classes (i.e., with coarse proportions less than 0.5) for the two datasets.

B. *Comparison with the class allocation in units of class*

Generally, the final SPM results of HNN-based methods (i.e., H-HNN and original HNN in this paper) are generated by assigning to each subpixel the class with the highest neuron output value (normally between 0 and 1). Thus, this approach does not obey the coarse proportions exactly. Alternatively, the final determination of class labels can be realized using the class allocation in units of class (UOC) strategy, as proposed by Wang et al [31]. Specifically, the number of subpixels assigned to each class is determined exactly by the coarse proportions. By UOC, the subpixels are allocated a class label in turn. The UOC-based strategy obeys the coarse proportions entirely. Nevertheless, the spatial details of small-sized land cover patches and errors are preserved simultaneously in the UOC-based SPM results.

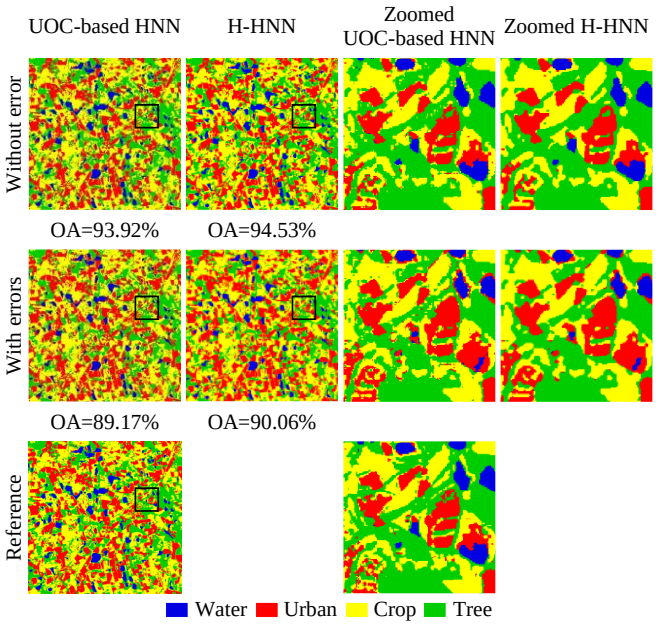

Fig. 12. SPM results of UOC-based HNN and H-HNN for the France datasets (*s*=4).

The SPM results of the proposed H-HNN are compared with those of the UOC-based HNN for the France dataset in Fig. 12 with a zoom factor of 4. The cases involving ideal coarse proportions and coarse proportions with errors were considered, which were created by degrading the non-linear SVM and linear SVM-derived land cover maps shown in Fig. 3(c) and Fig. 9(a), respectively. The OA values are also provided below the results. It can be seen that the H-HNN produces more satisfactory results than the UOC-based HNN. Specifically, the errors in the coarse proportions have a noticeable negative effect on the UOC-based HNN results, exhibiting smaller OAs than the H-HNN results for both datasets. Overall, the H-HNN method demonstrates more accurate results both visually and through quantitative evaluation compared to the UOC-based HNN.

C. *Potential extensions of H-HNN*

The two introduced hard label-based constraints in the H-HNN align with the intrinsic hard classification task of SPM. In this paper, SPM was achieved by exploring only the spatial information (i.e., spatial dependence) based on the H-HNN. In cases where auxiliary data are available, the H-HNN can be extended effortlessly by incorporating such data while accommodating the two proposed hard label-based constraints. For example, within spatio-temporal-based SPM [52], the energy function of H-HNN can be extended to encompass the objectives of maximizing temporal dependency among time-series images, alongside the original goal of maximal spatial attraction, formulated as $E = G_{spatial} + G_{temporal} + C_{constraints}$.

In this paper, the common case of H-resolution (i.e., the spatial size of object is larger than one pixel) [8] was considered for SPM. For the L-resolution case (i.e., scene dominated by objects smaller than one pixel) where prior information on spatial patterns is required, the HNN has shown its advantages in reproducing small-sized patches [48]. Similarly, in the L-resolution case, it is difficult for the HNN to fully realize the coarse proportion constraint, resulting in the inevitable loss of some small-sized patches. By incorporating hard class label-based constraints, the H-HNN proposed in this paper has

demonstrated its advantages in reproducing more spatial details, and it should, therefore, be explored in future as a potential solution to enhance SPM in the L-resolution case.

In the H-HNN, the forms of the hard-based constraints are designed for the aim of minimizing the energy function. That is, the two constraints are supposed to be zero ideally. Theoretically, higher order forms (e.g., v^3 and v^4) also satisfy the zero condition. The accuracy assessments of H-HNN using higher order forms of constraints are shown in Table 5, using the South Carolina dataset with a zoom factor of 4 as an example. The results demonstrated similar accuracies. Therefore, the squared form is suggested for the proposed hard-based constraints in H-HNN. The proposed hard-based constraints can also be utilized for other SPM methods that consider the coarse proportion as constraints such as the multi-objective optimization-based algorithm [22].

Table 5 The SPM accuracy (%) of H-HNN using different forms of hard-label based constraints for the South Carolina dataset (s=4)

Constraint forms	Water	Urban	Crop	Tree	OA
v^2	90.69	87.05	90.32	87.52	88.63
v^3	91.13	87.24	90.02	87.44	88.59
v^4	91.15	87.27	89.99	87.44	88.57

D. Computational cost of H-HNN

The proposed H-HNN is constructed on the basis of HNN-based SPM by adding the hard label-based constraints, which are also involved in the iterative optimization process. Thus, it can be more time-consuming than HNN. The computation times of the four methods (PSA, RBF, HNN and H-HNN) for the South Carolina dataset are listed in Table 6. All experiments were conducted on an Intel Core i7 processor using the MATLAB 2023b software. For the HNN and H-HNN, the iteration step was set to 0.001, and the number of iterations was 1000. It was found that the RBF and PSA methods exhibit faster speed compared to the HNN and H-HNN, but show compromised accuracies in the experiments in Section III. The H-HNN method requires slightly more time (about 10%) than the HNN due to the incorporation of the proposed hard label-based constraints. With the rapid advancement of high-performance computing systems, the slight sacrifice in computational cost is worthwhile given the increased accuracy. For example, for the South Carolina dataset with a zoom factor of 4, H-HNN achieves approximately 1% increase in accuracy compared to HNN. Considering the difficulty in increasing SPM accuracy without any additional auxiliary data (i.e., under the same assumptions of spatial dependence), the 1% increase in accuracy with H-HNN is undoubtedly valuable.

Table 6 Consuming time of the four methods for the South Carolina dataset (in units of seconds)

Zoom factor	PSA	RBF	HNN	H-HNN
s=4	119	11	1632	1788
s=8	39	10	3267	3587

V. Conclusion

The HNN model has gained wide adoption in SPM due to its unique capability of reproducing the coarse proportion constraints, but not perfectly, thus exhibiting resilience to errors in the coarse proportions. However, during the optimization of class labels, the HNN model always exhibits vanished gradients, which make it unable to push the subpixels to the hard class label of 0 or 1, even after thousands of iterations. This challenging issue results in significant uncertainties in assigning a hard class label to each subpixel, and can lead to the loss of small-sized land cover features and spatial details. To address this issue, this paper introduces the H-HNN method, which integrates two hard label-based constraints. By imposing stricter constraints, the proposed H-HNN aims to reduce the inherent uncertainties in predicting hard class labels in the HNN model, and reproduce more spatial details at the target fine spatial resolution. Experiments were conducted on two datasets with varying levels of proportion errors to evaluate the performance of the H-HNN. Comparative analyses were carried out against three typical SPM methods including the PSA, RBF and HNN. The main findings are outlined as follows.

1) Both visual and quantitative assessments reveal that the H-HNN yields more accurate SPM predictions than the three benchmark methods.
2) The incorporation of the hard label-based constraints in the H-HNN results in more accurate class labels than the original HNN, and the H-HNN is advantageous for reconstruction of spatial details and small-sized land cover patches.
3) With errors exist in the coarse proportions, the SPM accuracies of all methods decrease. However, the H-HNN remains the most accurate and robust across all scenarios, demonstrating significant advantages in mitigating the impact of errors in the coarse proportions compared to the PSA and RBF methods.
4) The proposed H-HNN exhibits the advantages over the original HNN model in reconstructing more accurate proportions at the coarse pixel scale. For example, by H-HNN, the RMSE of the reproduced coarse proportions is decreased by 0.015.

References

[1] N. T. Hoang, and K. Kanemoto, "Mapping the deforestation footprint of nations reveals growing threat to tropical forests," *Nat Ecol Evol*, vol. 5, no. 6, pp. 845-853, 2021.

[2] W. Shi, M. Goodchild, M. Batty *et al.*, "Prospective for urban informatics," *Urban Informatics*, vol. 1, no. 1, pp. 2, 2022.

[3] P. Fisher, "The pixel: A snare and a delusion," *International Journal of Remote Sensing*, vol. 18, no. 3, pp. 679-685, 1997.

[4] F. Xu, S. Heremans, and B. Somers, "Urban land cover mapping with sentinel-2: A spectro-spatio-temporal analysis," *Urban Informatics*, vol. 1, no. 1, pp. 8, 2022.

[5] P. M. Atkinson, "Super-resolution land cover classification using the two-point histogram," in geoENV IV—Geostatistics for Environmental Applications: Proceedings of the Fourth European Conference on Geostatistics for Environmental Applications held in Barcelona, Spain, November 27–29, 2002. Dordrecht: Springer Netherlands, 2004.

[6] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing," vol. 32, pp. 167-203, 2020.

[7] X. Wang, P. M. Atkinson, Y. Zhang *et al.*, "Automatic mapping of 500 m daily open water body fraction in the american continent using goes-16 abi imagery," *Remote Sensing of Environment*, vol. 304, pp. 114040, 2024.

[8] P. M. Atkinson, "Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study," *International Journal of Remote Sensing*, vol. 30, no. 20, pp. 5293-5308, 2009.

[9] Y. Qin, C. Zhang, and P. Lu, "A fully automatic framework for sub-pixel mapping of thermokarst lakes using sentinel-2 images," *Science of Remote Sensing,* vol. 8, pp. 100111, 2023.

[10] I. Olthof, and R. H. Fraser, "Mapping surface water dynamics (1985–2021) in the hudson bay lowlands, canada using sub-pixel landsat analysis," *Remote Sensing of Environment,* vol. 300, pp. 113895, 2024.

[11] X. Li, G. Zhang, S. Tan *et al.*, "Forest fire smoke detection research based on the random forest algorithm and sub-pixel mapping method," *Forests,* vol. 14, no. 3, pp. 485, 2023.

[12] D. He, Q. Shi, J. Xue *et al.*, "Very fine spatial resolution urban land cover mapping using an explicable sub-pixel mapping network based on learnable spatial correlation," *Remote Sensing of Environment,* vol. 299, pp. 113884, 2023.

[13] A. J. Tatem, H. G. Lewis, P. M. Atkinson *et al.*, "Multiple-class land-cover mapping at the sub-pixel scale using a hopfield neural network," *International Journal of Applied Earth Observation and Geoinformation,* vol. 3, no. 2, pp. 184-190, 2001.

[14] K. C. Mertens, L. P. C. Verbeke, E. I. Ducheyne *et al.*, "Using genetic algorithms in sub-pixel mapping," *International Journal of Remote Sensing,* vol. 24, no. 21, pp. 4241-4247, 2003.

[15] P. M. Atkinson, "Sub-pixel target mapping from soft-classified, remotely sensed imagery," *Photogrammetric Engineering & Remote Sensing,* vol. 71, no. 7, pp. 839-846, 2005.

[16] K. C. Mertens, B. de Baets, L. P. C. Verbeke *et al.*, "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *International Journal of Remote Sensing,* vol. 27, no. 15, pp. 3293-3310, 2006.

[17] V. A. Tolpekin, and A. Stein, "Quantification of the effects of land-cover-class spectral separability on the accuracy of markov-random-field-based superresolution mapping," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, no. 9, pp. 3283-3297, 2009.

[18] Q. Wang, W. Shi, and P. M. Atkinson, "Sub-pixel mapping of remote sensing images based on radial basis function interpolation," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 92, pp. 1-15, 2014.

[19] Y. Zhong, Y. Wu, X. Xu *et al.*, "An adaptive sub-pixel mapping method based on map model and class determination strategy for hyperspectral remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 3, pp. 1411-1426, 2015.

[20] L. Li, Y. Chen, X. Yu *et al.*, "Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 101, pp. 10-21, 2015.

[21] P. Wang, Y. Zhang, L. Wang *et al.*, "Sub-pixel mapping of spectral imagery based on deviation information measurement," *IEEE Transactions on Instrumentation and Measurement,* vol. 72, pp. 1-16, 2023.

[22] M. Song, Y. Zhong, A. Ma *et al.*, "Multiobjective spatiotemporal subpixel mapping for remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 62, pp. 1-17, 2024.

[23] X. Zhang, Y. Ge, J. Chen *et al.*, "High-quality super-resolution mapping using spatial deep learning," *iScience,* vol. 26, no. 6, pp. 106875, 2023.

[24] Y. Xu, and B. Huang, "A spatio-temporal pixel-swapping algorithm for subpixel land cover mapping," *IEEE Geoscience and Remote Sensing Letters,* vol. 11, no. 2, pp. 474-478, 2014.

[25] F. Ling, X. Li, F. Xiao *et al.*, "Object-based sub-pixel mapping of buildings incorporating the prior shape information from remotely sensed imagery," *International Journal of Applied Earth Observation and Geoinformation,* vol. 18, pp. 283-292, 2012.

[26] D. He, Y. Zhong, Q. Shi *et al.*, "Generating continuous fine-scale land cover mapping by edge-guided maximum a posteriori based spatiotemporal sub-pixel mapping," *Science of Remote Sensing,* vol. 5, pp. 100041, 2022.

[27] X. Li, F. Ling, G. M. Foody *et al.*, "Monitoring high spatiotemporal water dynamics by fusing modis, landsat, water occurrence data and dem," *Remote Sensing of Environment,* vol. 265, pp. 112680, 2021.

[28] C. Huang, Y. Chen, and J. Wu, "Dem-based modification of pixel-swapping algorithm for enhancing floodplain inundation mapping," *International Journal of Remote Sensing,* vol. 35, no. 1, pp. 365-381, 2013.

[29] Y. Chen, G. Zhang, H. Cui *et al.*, "A novel weakly supervised semantic segmentation framework to improve the resolution of land cover product," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 196, pp. 73-92, 2023.

[30] Q. Hu, H. Yin, M. A. Friedl *et al.*, "Integrating coarse-resolution images and agricultural statistics to generate sub-pixel crop type maps and reconciled area estimates," *Remote Sensing of Environment,* vol. 258, pp. 112365, 2021.

[31] Q. Wang, W. Shi, and L. Wang, "Allocating classes for soft-then-hard subpixel mapping algorithms in units of class," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 5, pp. 2940-2959, 2014.

[32] Y. F. Su, G. M. Foody, A. M. Muad *et al.*, "Combining pixel swapping and contouring methods to enhance super-resolution mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 5, no. 5, pp. 1428-1437, 2012.

[33] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez *et al.*, "Spectral variability in hyperspectral data unmixing: A comprehensive review," *IEEE Geoscience and Remote Sensing Magazine,* vol. 9, no. 4, pp. 223-270, 2021.

[34] R. M. Cavalli, "Spatial validation of spectral unmixing results: A systematic review," *Remote Sensing,* vol. 15, no. 11, pp. 2822, 2023.

[35] A. J. Tatem, H. G. Lewis, P. M. Atkinson *et al.*, "Super-resolution target identification from remotely sensed images using a hopfield neural network," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 4, pp. 781-796, 2001.

[36] L. Xiaodong, D. Yun, L. Feng *et al.*, "Superresolution mapping of remotely sensed image based on hopfield neural network with anisotropic spatial dependence model," *IEEE Geoscience and Remote Sensing Letters,* vol. 11, no. 7, pp. 1265-1269, 2014.

[37] Y.-F. Su, "Integrating a scale-invariant feature of fractal geometry into the hopfield neural network for super-resolution mapping," *International Journal of Remote Sensing,* pp. 1-22, 2019.

[38] H. G. Cyril Amala Dhason, I. Muthaia, S. P. Sakthivel *et al.*, "Super‐resolution mapping of hyperspectral satellite images using hybrid genetic algorithm," *IET Image Processing,* vol. 14, no. 7, pp. 1281-1290, 2020.

[39] Y.-F. Su, G. M. Foody, A. M. Muad *et al.*, "Combining hopfield neural network and contouring methods to enhance super-resolution mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 5, no. 5, pp. 1403-1417, 2012.

[40] A. M. Muad, S. K. M. Zaki, and S. A. Jasim, "Optimizing hopfield neural network for super-resolution mapping," *J. Kejuruter,* vol. 32, pp. 91-97, 2020.

[41] C. Zhang, Q. Wang, P. Lu *et al.*, "Fast and slow changes constrained spatio-temporal subpixel mapping," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, pp. 1-16, 2022.

[42] X. Li, F. Ling, Y. Du *et al.*, "A spatial-temporal hopfield neural network approach for super-resolution land cover mapping with multi-temporal different resolution remotely sensed images," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 93, pp. 76-87, 2014.

[43] Q. M. Nguyen, P. M. Atkinson, and H. G. Lewis, "Super-resolution mapping using hopfield neural network with panchromatic imagery," *International Journal of Remote Sensing,* vol. 32, no. 21, pp. 6149-6176, 2011.

[44] M. Q. Nguyen, P. M. Atkinson, and H. G. Lewis, "Superresolution mapping using a hopfield neural network with fused images," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 3, pp. 736-749, 2006.

[45] M. Q. Nguyen, P. M. Atkinson, and H. G. Lewis, "Superresolution mapping using a hopfield neural network with lidar data," *IEEE Geoscience and Remote Sensing Letters,* vol. 2, no. 3, pp. 366-370, 2005.

[46] F. Ling, Y. Du, F. Xiao *et al.*, "Super-resolution land-cover mapping using multiple sub-pixel shifted remotely sensed images," *International Journal of Remote Sensing,* vol. 31, no. 19, pp. 5023-5040, 2010.

[47] Q. Wang, W. Shi, P. M. Atkinson *et al.*, "Land cover change detection at subpixel resolution with a hopfield neural network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 8, no. 3, pp. 1339-1352, 2015.

[48] A. M. Muad, and G. M. Foody, "Impact of land cover patch size on the accuracy of patch area representation in hnn-based super resolution mapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 5, no. 5, pp. 1418-1427, 2012.

[49] A. J. Tatem, H. G. Lewis, P. M. Atkinson *et al.*, "Super-resolution land cover pattern prediction using a hopfield neural network," *Remote Sensing of Environment,* vol. 79, no. 1, pp. 1-14, 2002.

[50] N. E. Ayat, M. Cheriet, and C. Y. Suen, "Automatic model selection for the optimization of svm kernels," *Pattern Recognition,* vol. 38, no. 10, pp. 1733-1745, 2005.

[51] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," *In 2019 International Conference on Intelligent Sustainable Systems (ICISS)*, IEEE, pp. 24-28, 2019.

[52] X. Li, F. Ling, G. M. Foody *et al.*, "Spatial-temporal super-resolution land cover mapping with a local spatial-temporal dependence model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4951-4966, 2019.