

# ハイパースペクトルデータ分類における深層特徴再構成の評価

岡田 純京†, 小篠 裕子†

概要画素単位のハイパースペクトル画像(HSI)分類のためのほとんどの最先端手法は、畳み込みニューラルネットワーク(CNN)に基づいている。本論文では、分類精度を向上させるために、ピクセル単位のHSI分類のCNNベースのネットワークに特徴再構成モジュール(FRM)を導入する。FRMは、行列分解を用いた低ランク近似により、CNN特徴量の元の行列に内在する特徴を抽出することができる。FRMをCNNベースのピクセル単位のHSI分類ネットワークに導入する前と後の分類精度を比較し、その有効性を検証する。実験結果は、この方法が分類精度を向上させたことを示している。また、分類精度の向上に寄与した特徴量を評価するために、オリジナルのCNN特徴量と再構成されたCNN特徴量を可視化し、比較した。

キーワード：ハイパースペクトル画像, 画素毎分類, 機械学習

## 1. Introduction

ピクセル単位の分類は、ハイパースペクトル画像(HSI)のような地球観測衛星ベースの画像に意味クラスを割り当てることを目的としたリモートセンシングの基本的なタスクである。HSIはシーンのスペクトル特性に関する豊富な詳細情報を提供する<sup>4)</sup>。ピクセル単位のHSI分類は、農業<sup>17)</sup>、土地被覆分類<sup>5)</sup>、その他の関連タスクなど、様々なアプリケーションで研究の注目を集めている。従来、機械学習では、サポートベクターマシン<sup>12)</sup>やk-nearest neighbor<sup>18)</sup>などの画素単位のHSI分類が利用され、現在、深層学習が注目されている。特に、畳み込みニューラルネットワーク(CNN)は、ピクセル単位のHSI分類に広く使われているディープラーニングアプローチである。

RGB 画像のセマンティックセグメンテーションにおいて、CNN 特徴の行列を低ランク行列に再構成することで、セグメンテーション性能が向上することが先行研究で報告されている<sup>7)</sup>。本研究では、行列分解を用いて、行列をデータの本質的な特徴を表す2つ以上の低ランク行列に分解し<sup>15)</sup>、これらの低ランク行列を用いて行列を再構成する。

再構成後のCNN特徴量は、本来のCNN特徴量よりも低いランクでありながら、その本質的な特徴量は維持されている。CNNとその特徴再構成により、セマンティックセグメンテーションの性能が向上した。

本論文では、分類精度を向上させるために、CNNに基づくピクセル単位のHSI分類のためのCNN特徴再構成を紹介する。このタスクのためにCNN特徴再構成を導入したのは我々が初めてである。この手法の有効性を検証するために、土地被覆分類の3つのHSIベンチマークにおいて、分類精度におけるCNN特徴再構成の導入前後を比較する。

さらに、特徴マップの可視化、計算時間、パラメータ数を比較することで、CNN特徴再構成の導入の効果を評価する。我々は、元のCNN特徴と再構成されたCNN特徴を可視化し、比較することで、どの特徴が分類精度の向上に寄与したかを議論した。リモートセンシングの分野では、衛星と地上との通信を最小化するために、オンボード処理が注目されている<sup>1)</sup>。この分野での効率的な処理の重要性を考慮し、計算時間とパラメータ数の観点からもそのコストを評価する。

実験結果は、CNNに基づくピクセル単位のHSI分類にCNN特徴再構成を導入することの分類精度の有効性を実証し、リモートセンシングタスクにおける将来の応用の可能性を示している。CNN特徴再構成は、1画素のスペクトルデータよりも高次元のRGB画像を用いたタスクにおいて有効であることが示されている。

受理；改訂；受理 2024年2月9日；改訂 2024年5月2日；受理 2024年5月29日

†Graduate School of Advanced Science and Technology, Tokyo Denki University  
(Tokyo, Japan)

# Evaluation of Deep Feature Reconstruction for Pixel-Wise Hyperspectral Image Classification

Junkei Okada<sup>†</sup>, Yuko Ozasa<sup>†</sup>

**Abstract** Most state-of-the-art methods for pixel-wise hyperspectral image (HSI) classification are based on the Convolutional Neural Network (CNN). In this paper, we introduce a feature reconstruction module (FRM) into the CNN-based network of pixel-wise HSI classification to improve classification accuracy. FRM can extract essential characteristics in the original matrix of CNN features by low-rank approximation using matrix factorization. We compare the classification accuracy before and after the introduction of FRM into the CNN-based network of pixel-wise HSI classification to validate its effectiveness. Experimental results demonstrate this method improved classification accuracy. We also visualized and compared the original CNN features and the reconstructed CNN features to evaluate which features contributed to the improvement in classification accuracy.

**Key words:** hyperspectral image, pixel-wise classification, machine learning

## 1. Introduction

Pixel-wise classification is a fundamental task of remote sensing that aims at assigning a semantic class in Earth-observing satellite-based images such as hyperspectral images (HSIs). HSIs provide abundant detailed information about the spectral properties of a scene<sup>4</sup>. Pixel-wise HSI classification has been attracting research attention in various applications, such as agriculture<sup>17</sup>, land-cover classification<sup>5</sup>, and other related tasks.

Traditionally, pixel-wise HSI classification has been exploited in machine learning, including support vector machine<sup>12</sup> and k-nearest neighbor<sup>18</sup>, while deep learning is currently attracting attention. In particular, Convolutional Neural Network (CNN) is widely used deep learning approach for pixel-wise HSI classification.

In semantic segmentation of RGB images, a previous work reported that reconstructs a matrix of CNN features into a low-rank matrix, resulting in improved segmentation performance<sup>7</sup>. This work used matrix decomposition to decompose the matrix into two or more low-rank matrices that represent the data's essential characteristics<sup>15</sup> and reconstruct the matrix using these low-rank matrices. The CNN feature after reconstruction is of lower rank than the original CNN fea-

tures while preserving its essential characteristics. The CNN with its feature reconstruction improved the performance of semantic segmentation.

In this paper, we introduce the CNN feature reconstruction for the CNN-based pixel-wise HSI classification to improve classification accuracy. We are the first to introduce the CNN feature reconstruction for this task. To validate the effectiveness of this method, we compare before and after the introduction of CNN feature reconstruction in classification accuracy on three HSI benchmarks of land-cover classification.

Furthermore, we evaluate the effects of introducing the CNN feature reconstruction by comparing feature map visualization, computation time, and number of parameters. We visualized and compared original CNN features with reconstructed CNN features to discuss which features contributed to the improvement of classification accuracy. In the remote sensing field, on-board processing has been attracting attention in order to minimize satellite-ground communication<sup>1</sup>. Given the importance of efficient processing in this field, we also evaluate its cost in terms of computation time and the number of parameters.

Experimental results demonstrated the effectiveness of introducing the CNN feature reconstruction into CNN-based pixel-wise HSI classification in terms of classification accuracy and indicate its potential for future applications in remote sensing tasks. The CNN feature reconstruction has been shown to be effective in tasks with RGB images that have higher di-

Received February 9, 2024; Revised May 2, 2024; Accepted May 29, 2024

<sup>†</sup>Graduate School of Advanced Science and Technology, Tokyo Denki University  
(Tokyo, Japan)

この方法は、次世代HSカメラからの高次元スペクトルデータを入力としてでも、様々なアプリケーションのピクセル単位のHSI分類の性能を向上させることができる。本論文の主な貢献は、ピクセル単位のHSI分類にCNN特徴再構成を導入したことと、再構成されたCNN特徴による分類精度の向上に寄与した特徴かについての議論である。

## 2. 関連研究

### 2.1 ピクセル単位のHSI分類

現在、先行研究の多くは、CNNベースの手法を用いたピクセル単位のHSI分類に焦点を当てている。Huら<sup>9)</sup>は、畳み込み層、プーリング層、完全連結層を含む1次元CNNを提案し、2層ニューラルネットワークやサポートベクターマシン分類器よりも高い精度を示した。Sidikeら<sup>16)</sup>は、限られた学習データに対してネットワークユニット数を増やすために漸進的拡張を導入した深い漸進的拡張ネットワーク(dPEN)を提案した。dPENは1次元CNN<sup>9)</sup>や他の一般的な機械学習手法を凌駕した。しかし、CNN特徴量から特徴量を抽出して分類精度を向上させる試みはなされていない。

### 2.2 行列の分解

行列分解は、観測された行列をいくつかの部分行列の積に因数分解する。部分行列の積からデータを生成するプロセスを仮定すると、行列分解も逆演算として働く。これは、観測データの潜在構造が、行列分解<sup>7)</sup>によって分解された部分行列の積から復元できることを示している。

与えられたデータが行列  $X = [x_1, \dots, x_n] \in \mathbb{R}^{d \times n}$  であるとする。一般的な仮定は、 $X$ に隠された低次元の部分空間または複数の部分空間の和が存在することである。すなわち、辞書行列  $D = [d_1, \dots, d_r] \in \mathbb{R}^{d \times r}$  と対応する  $C \in \mathbb{R}^{r \times n}$  が存在し、 $X$ は次式で表せる。

$$X = \bar{X} + E = DC + E, \quad (1)$$

ここで、 $\bar{X} \in \mathbb{R}^{d \times n}$  は出力低ランク再構成、 $E \in \mathbb{R}^{d \times n}$  は破棄するノイズ行列である。ここで、復元された行列 $X$ は、次のような低ランクの性質を持つと仮定する。

$$\text{rank}(\bar{X}) \leq \min(\text{rank}(D), \text{rank}(C)) \leq r \ll \min(d, n). \quad (2)$$

因数分解のランク $r$ は一般に  $(d+n)r < dn^{10)}$  となるように選ばれる。

### 2.3 深い特徴の再構成

Gengら<sup>7)</sup>はHamburgerと呼ばれる特徴再構成モジュール(FRM)を提案した。このモジュールは、深い特徴に対する低ランク行列補完問題を解き、CNNベースのネットワークにおける大域的特徴の抽出を導くことで、大域的文脈をモデル化することができる。ハンバーガーは1つの行列分解と2つの線形変換<sup>7)</sup>からなる。ハンバーガーはまず入力  $Z \in \mathbb{R}^d \times \mathbb{Z}^{n \times n}$  を線形変換  $W_l \in \mathbb{R}^{d \times d}$  で特徴空間に写像し、次に行列分解 $M$ を用いて低ランク信号部分空間を解き、最後に抽出した信号を別の線形変換  $W_u \in \mathbb{R}^d \times \mathbb{Z}^{d \times d}$  で出力に変換する、

$$H(Z) = W_u M(W_l Z), \quad (3)$$

ここで、 $M$ は明確な潜在構造を復元するための行列分解である。最後に、このネットワークはスキップ接続で $bmY$ を出力する、

$$Y = Z + H(Z). \quad (4)$$

$M$ をネットワークに融合させるためには、反復アルゴリズムがどのように勾配を逆伝播するかが重要な問題である。反復プロセスを区別するための標準的な選択として、リカレントニューラルネットワークのような動作の最適化がある。しかし、ランダムに初期化された因子行列は複数回の反復を必要とし、実際には勾配が消失するため性能が低下する。そこで、 $M$ の最適化ステップで消失する勾配を克服するために、1ステップの勾配<sup>7)</sup>を適用する。

## 3. Methods

3.1 ピクセル単位のHSI分類のプロセス 1ピクセル単位のHSI分類ネットワーク $f$ は、各ピクセルにラベルが与えられたHSIを用いて学習される。 $N$ 個の画素  $(x_1, x_2, \dots, x_N)$  は、各画素  $x_i$  が  $B$ 個のバンド  $(x_i \in \mathbb{R}^B)$  を持ち、HSI  $X \in \mathbb{R}^{B \times H \times W}$  から選択され、入力として使用される。ここで、 $H$ は高さ、 $W$ は幅、 $B$ はバンド数である。各 $N$ 個の画素に対応する $N$ 個のラベル  $\{y_1, y_2, \dots, y_N\}$  が与えられ、 $\forall y_i \in \{1, 2, \dots, C\}$ 、ここで $C$ はクラス数である。ピクセル単位のHSI分類ネットワーク $f$ は $x_i$ を入力とし、特徴量 $z_i$ を抽出する。抽出された特徴量 $z_i$ は、線形変換とソフトマックス関数により、クラス番号 $C$ の予測ラベル $\hat{y}_i$ を出力する。本論文では、ピクセル単位のHSI分類に1次元畳み込みニューラルネットワーク(1D-CNN)を使用する。

mensionality than the spectral data of a single pixel. This method could improve performance in pixel-wise HSI classification for various applications, even with higher-dimensional spectral data from next-generation HS cameras as input. The main contributions of this paper are the introduction of CNN feature reconstruction into pixel-wise HSI classification and the discussion of which features contributed to the improvement of classification accuracy by reconstructed CNN features.

## 2. Related Work

### 2.1 Pixel-Wise HSI Classification

At present, most previous research focuses on pixel-wise HSI classification using CNN-based methods. Hu et al.<sup>9)</sup> proposed a 1D CNN which contains a convolutional layer, a pooling layer, and a fully connected layer, and showed better accuracy than two-layer neural network and support vector machine classifier. Sidike et al.<sup>16)</sup> proposed a deep Progressively Expanded Network (dPEN) that introduced progressive expansion to increase the number of network units to the limited training data. dPEN outperformed 1D CNN<sup>9)</sup> and other popular machine learning methods. However, there has been no attempt to extract the characteristic features from the CNN features to improve the classification accuracy.

### 2.2 Matrix Decomposition

Matrix decomposition factorizes the observed matrix into a product of several sub-matrices. Assuming the process of generating data from a product of sub-matrices, matrix decomposition also acts as an inverse operation. This indicates that the latent structure of the observed data can be recovered from the product of sub-matrices decomposed by matrix decomposition<sup>7)</sup>.

Suppose that the given data is a matrix  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ . A general assumption is that there is a low-dimensional subspace or a union of multiple subspaces hidden in  $\mathbf{X}$ . That is, there exists a dictionary matrix  $\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_r] \in \mathbb{R}^{d \times r}$  and corresponding  $\mathbf{C} \in \mathbb{R}^{r \times n}$  that  $\mathbf{X}$  can be expressed as

$$\mathbf{X} = \bar{\mathbf{X}} + \mathbf{E} = \mathbf{DC} + \mathbf{E}, \quad (1)$$

where  $\bar{\mathbf{X}} \in \mathbb{R}^{d \times n}$  is the output low-rank reconstruction, and  $\mathbf{E} \in \mathbb{R}^{d \times n}$  is the noise matrix to be discarded. Here we assume that the recovered matrix  $\mathbf{X}$  has the low-rank property, such that

$$\text{rank}(\bar{\mathbf{X}}) \leq \min(\text{rank}(\mathbf{D}), \text{rank}(\mathbf{C})) \leq r \ll \min(d, n). \quad (2)$$

The rank  $r$  of the factorization is generally chosen so that  $(d+n)r < dn^{10)}$ .

### 2.3 Deep Feature Reconstruction

Geng et al.<sup>7)</sup> proposed a feature reconstruction module (FRM) called Hamburger. This module can model global context by solving the low-rank matrix completion problem for deep features and guide extract of global features in CNN-based networks. Hamburger consists of one matrix decomposition and two linear transformations<sup>7)</sup>. Hamburger first maps the input  $\mathbf{Z} \in \mathbb{R}^{d_z \times n}$  into feature space with a linear transformation  $\mathbf{W}_l \in \mathbb{R}^{d_z \times d_z}$ , then uses matrix decomposition  $\mathcal{M}$  to solve a low-rank signal subspace, and finally transforms extracted signals into the output with another linear transformation  $\mathbf{W}_u \in \mathbb{R}^{d_z \times d}$ ,

$$\mathcal{H}(\mathbf{Z}) = \mathbf{W}_u \mathcal{M}(\mathbf{W}_l \mathbf{Z}), \quad (3)$$

where  $\mathcal{M}$  is matrix decomposition to recover the clear latent structure. Finally, this network outputs  $bmY$  via skip connections,

$$\mathbf{Y} = \mathbf{Z} + \mathcal{H}(\mathbf{Z}). \quad (4)$$

To fuse  $\mathcal{M}$  into the networks, a crucial issue is how the iterative algorithm back-propagates gradients. As the standard choice to differentiate the iterative process, there is the optimization of Recurrent Neural Network-like behavior. However, randomly initialized factor matrices require multiple iterations, which will degrade performance due to the vanishing gradient in practice. Therefore, we apply the one-step gradient<sup>7)</sup> to conquer the gradient vanishing during the optimization step for  $\mathcal{M}$ .

## 3. Methods

### 3.1 Process of Pixel-Wise HSI Classification

A pixel-wise HSI classification network  $f$  is trained using HSIs with labels given to each pixel.  $N$  pixels  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ , where each pixel  $\mathbf{x}_i$  has  $B$  bands ( $\mathbf{x}_i \in \mathbb{R}^B$ ), are selected from an HSI  $\mathbf{X} \in \mathbb{R}^{B \times H \times W}$  and used as input. Here,  $H$  is the height,  $W$  the width, and  $B$  the number of bands. Each  $N$  pixel has corresponding  $N$  labels  $\{y_1, y_2, \dots, y_N\}$  are given and  $\forall y_i \in \{1, 2, \dots, C\}$ , where  $C$  is the number of classes.

The pixel-wise HSI classification network  $f$  takes  $\mathbf{x}_i$  as input and extracts features  $\mathbf{z}_i$ . The extracted features  $\mathbf{z}_i$  output predictive labels  $\hat{y}_i$  of class number  $C$  by linear transformation and softmax function. In this paper, a one-dimensional convolutional neural network (1D-CNN) is used for the pixel-wise HSI classification

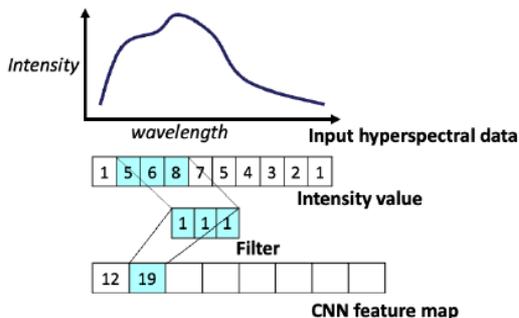


図1: 1D-CNNの処理例を示す。

network  $f$ .

一般に、1D-CNNは様々なフィルタを用いて入力スペクトルの特徴を抽出し、各フィルタの特徴マップを得る。これらのフィルタは入力スペクトルを横切って移動し、フィルタの重みとバンド値の間のドット積を計算する。複数のフィルタを用いることで、様々な特徴マップを得ることができる。図1は、スペクトルデータを入力とする1D-CNNの処理を視覚的に示したものである。

### 3. 2 CNNに基づく画素単位のHSI分類のためのCNN特徴再構成の導入

1D-CNNを用いて画素単位のHSI分類ネットワークにFRMを導入する。FRMはHamburger<sup>7)</sup>と同じ構造を持ち、特徴抽出部と1D-CNNの分類器の間に導入される。CNNベースのピクセル単位のHSI分類法として、最先端の手法(SOTA)であるdPEN<sup>16)</sup>と代表的な手法であるHuら<sup>9)</sup>を採用した。両手法とも1D-CNNを使用し、入力HSIの画素の1Dスペクトルデータである。図2は、1D-CNNへのFRMの導入を示したものである。

非負行列分解(NMF)<sup>10)</sup>をMとして、辞書行列と係数行列を更新する。その結果、Xは2つの非負行列DとCに分解される。数学的には、最適化プロセスは以下のように定式化できる:

$$\min_{D,C} \|X - DC\|^2, \quad s.t. D \geq 0, C \geq 0, \quad (5)$$

DとCは乗法更新ルール<sup>11)</sup>に従って更新される:

$$C_{a\mu} \leftarrow C_{a\mu} \frac{(D^T X)_{a\mu}}{(D^T DC)_{a\mu}}, \quad D_{ia} \leftarrow D_{ia} \frac{(XC^T)_{ia}}{(DCC^T)_{ia}}. \quad (6)$$

## 4. Experiments

### 4.1 Dataset

ボツワナ、インディアンパイン、パヴィア大学の3つのデータセットで実験を行う。これらのデータセットの詳細を以下に示す。

ボツワナ(BW)データセット<sup>14)</sup>は、1996年にボツワナのオカバンゴデルタ上空でNASAのEO-1衛星によって取得され、14の土地被覆クラスを含む規則的な洪水イベントと関連する植生応答を監視した。このデータセットは1496×256ピクセルで構成され、波長400~2500nmの範囲に242のスペクトルバンドがあるが、ノイズや吸水バンドを除去した後、145のスペクトルバンドのみが使用されている。

インドパイン(IP)データセット<sup>2)</sup>は、1992年にインディアナ州北西部のインドパイン試験場において、大気中の可視/赤外画像スペクトロメーター(AVIRIS)によって収集され、16の相互に排他的な植生クラスを含む土壌研究をサポートする。このデータセットは145×145ピクセルからなり、400~2500nmの範囲をカバーする224のスペクトルバンドがある。水の吸収領域を除去することで、200個のスペクトルバンドが保持される。

パヴィア大学(UP)のデータセット<sup>13)</sup>は、2001年に北イタリアのパヴィアで行われた大学キャンパスを飛行する際、反射光学システムイメージング分光計(ROSIS)センサーによって取得され、HySensプロジェクト<sup>8)</sup>による都市の土地被覆分析のためにハイパースペクトルデータを利用した。このデータセットには、都市部の9つの土地被覆クラスが含まれている。空間分解能は610×340ピクセルで、430~860nmの波長域に115のスペクトルバンドがある。水の吸収の影響を受けたスペクトルバンドは除去され、データセットには103のバンドが残った。

### 4.2 実験セットアップ

公正な比較のために、BW、IP、UPデータセットのクラスごとに、ラベル付きサンプルが少ない小さなクラスを除いて、それぞれ205、95、947サンプルをランダムに選択する。すべての実験は、トレーニング(60%)、検証(20%)、テスト(20%)の各クラスのサンプルの分割割合の下で行われた。すべての競合手法は同じ訓練、検証、テストサンプルを使用し、すべての入力データは各バンドをBandmax-min<sup>3)</sup>と呼ばれる[0, 1]に正規化し、次式で表されることに注意。

$$x_{ij}^* = \frac{x_{ij} - \min(x_{:j})}{\max(x_{:j}) - \min(x_{:j})},$$

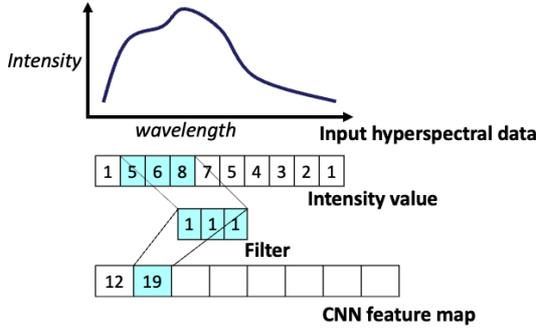


Fig. 1: A visual example of process of 1D-CNN.

network  $f$ .

In general, the 1D-CNN extracts a feature of an input spectral using various filters and obtains a feature map for each filter. These filters move across the input spectral, computing dot products between the filter weights and the band values. We can obtain various feature maps by using multiple filters. Fig. 1 shows a visual example of process of 1D-CNN with spectral data as input.

### 3.2 Introduction of CNN Feature Reconstruction for CNN-Based Pixel-Wise HSI Classification

We introduce the FRM into the pixel-wise HSI classification network  $f$  using 1D-CNN. The FRM has the same structure as Hamburger<sup>7)</sup> and is introduced between the feature extraction part and the classifier of 1D-CNN. As a CNN-based pixel-wise HSI classification method, we employed the method of state-of-the-art (SOTA) method, dPEN<sup>16)</sup>, and a representative method, Hu et al.<sup>9)</sup>. Both methods use 1D-CNN, and the input is the 1D spectral data of the pixel in the HSI. Fig. 2 illustrates the introduction of FRM into the 1D-CNN.

We use non-negative matrix factorization (NMF)<sup>10)</sup> as  $\mathcal{M}$  to update the dictionary and the coefficient matrices. Consequently,  $\mathbf{X}$  is decomposed into two non-negative matrices  $\mathbf{D}$  and  $\mathbf{C}$ . Mathematically, the optimization process can be formulated as follows:

$$\min_{\mathbf{D}, \mathbf{C}} \|\mathbf{X} - \mathbf{DC}\|^2, \quad s.t. \mathbf{D} \geq 0, \mathbf{C} \geq 0, \quad (5)$$

$\mathbf{D}$  and  $\mathbf{C}$  are updated according to following the Multiplicative Update rules<sup>11)</sup>:

$$\mathbf{C}_{a\mu} \leftarrow \mathbf{C}_{a\mu} \frac{(\mathbf{D}^T \mathbf{X})_{a\mu}}{(\mathbf{D}^T \mathbf{DC})_{a\mu}}, \quad \mathbf{D}_{ia} \leftarrow \mathbf{D}_{ia} \frac{(\mathbf{XC}^T)_{ia}}{(\mathbf{DCC}^T)_{ia}}. \quad (6)$$

## 4. Experiments

### 4.1 Dataset

We conduct experiments on the following three datasets which are Botswana, Indian Pines, and University of Pavia. The details of these datasets are described as follows.

The Botswana (BW) dataset<sup>14)</sup> was acquired by the NASA EO-1 satellite over the Okavango Delta, Botswana in 1996 to monitor regular flood events and associated vegetation responses, which contains 14 land-cover classes. This dataset comprises  $1496 \times 256$  pixels with 242 spectral bands in the wavelength range from 400 to 2500 nm, but only 145 spectral bands are used after removing noisy and water absorption bands.

The Indian Pines (IP) dataset<sup>2)</sup> was gathered by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over the Indian Pines test site in Northwestern Indiana in 1992 to support soil research, which contains 16 mutually exclusive vegetation classes. This dataset consists of  $145 \times 145$  pixels and 224 spectral bands covering the range of 400 – 2500 nm. By removing the water absorption region, 200 spectral bands are retained.

The University of Pavia (UP) dataset<sup>13)</sup> was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor during a flight campaign over the university campus at Pavia, Northern Italy in 2001 to exploit hyperspectral data for urban land cover analysis by HySens project<sup>8)</sup>. This dataset contains 9 land-cover classes from urban areas. The spatial resolution is  $610 \times 340$  pixels with 115 spectral bands in the wavelength range of 430 – 860 nm. The spectral bands influenced by water absorption were removed, and there were 103 bands remaining in the dataset.

### 4.2 Experimental Setup

For a fair comparison, we randomly select 205, 95, and 947 samples per class of the BW, IP, and UP datasets, respectively, except for those small classes containing few labeled samples. All the experiments were conducted under a split proportion of the samples from each class for training (60%), validation (20%), and testing (20%). Note that all competing methods used the same training, validation, and testing samples, and all input data normalized each band into  $[0, 1]$  called Bandmax-min<sup>3)</sup>, expressed as

$$x_{ij}^* = \frac{x_{ij} - \min(x_{:j})}{\max(x_{:j}) - \min(x_{:j})},$$

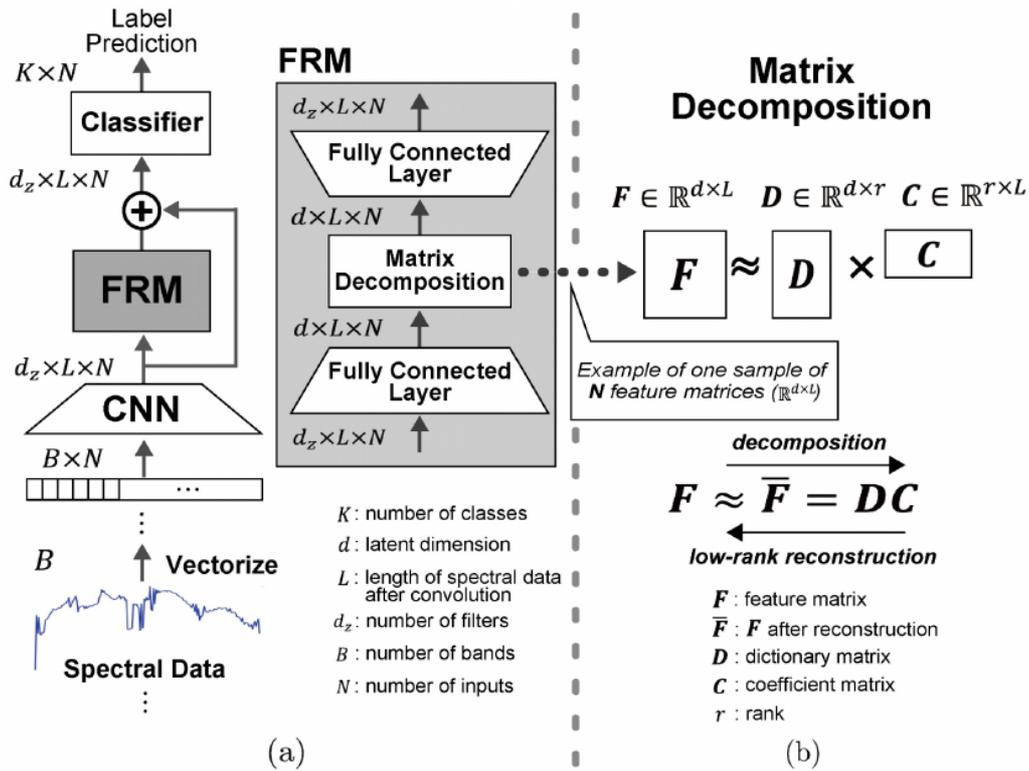


図2: 1D-CNNへのFRMの導入を描いたネットワークアーキテクチャの説明図。(a) CNN層、FRM、分類器の位置を示す全体的なネットワークアーキテクチャ。(b) FRMにおける行列分解のプロセス。行列がどのように分解されるかを示す。

ここで、 $x_i$ は元の画素ベクトル、 $x_{ij}$ はj番目の帯域のi番目の画素のスカラー値、 $x_{ij}^*$ は正規化画素値、 $\max(x_{.j})$ は帯域jの最大値、 $\min(x_{.j})$ は帯域jの最小値である。すなわち、各画素ベクトル $x_i$ は帯域的に正規化される。

Huら<sup>9)</sup>やSidikeら<sup>16)</sup>のようなFRMを導入したCNNベースのモデルは、隠れ層の活性化関数を双曲線正接関数から整流線形ユニット(ReLU)に変更することに注意する。FRMの行列分解としてNMFを採用したため、NMFへの中層層入力の出力は非負の制約を満たす必要がある。HuらとSidikeらのCNNベースのネットワークは、隠れ層の活性化関数として入力値を-1から1に変換する双曲線正接関数を用いているため、FRMで無視される負の値は分類に影響を与える可能性がある。したがって、ReLUは $\max(u, 0)$ のように値 $u < 0$ に対して入力値 $u$ を $u = 0$ に変換するので、負の値を避けるための活性化関数として使用される。

ネットワーク学習では、エポック数を1000エポック、バッチサイズ64とし、

表1: 分類精度(%)の比較分類精度は、

Methods	Dataset		
	BW	IP	UP
Hu et al. <sup>9)</sup>	93.23	84.35	92.34
- Introduction of FRM <sup>7)</sup>	<b>93.98</b>	<b>88.62</b>	<b>93.57</b>
dPEN (Sidike et al. <sup>16)</sup> )	93.61	83.33	92.87
- Introduction of FRM <sup>7)</sup>	<b>96.99</b>	<b>86.38</b>	<b>94.44</b>

学習率0.001、重み減衰 $10^{-5}$ で確率的勾配降下法を採用し、損失関数を減少させる。

#### 4. 3 分類精度の評価

FRMの有効性を評価するために、BW、IP、UPデータセットを用いて、CNNベースのネットワークにFRMを導入する前と後の分類精度を比較した。実験結果を表1に示す。FRMの導入は全てのデータセットで最良の結果を達成している。FRMを導入するだけで、SOTA法であるdPEN<sup>16)</sup>の分類精度が向上することがわかる。例えば、BW、IP、UPデータセットにおいて、dPENにFRMを導入する方法とオリジナルのdPEN<sup>16)</sup>を比較すると

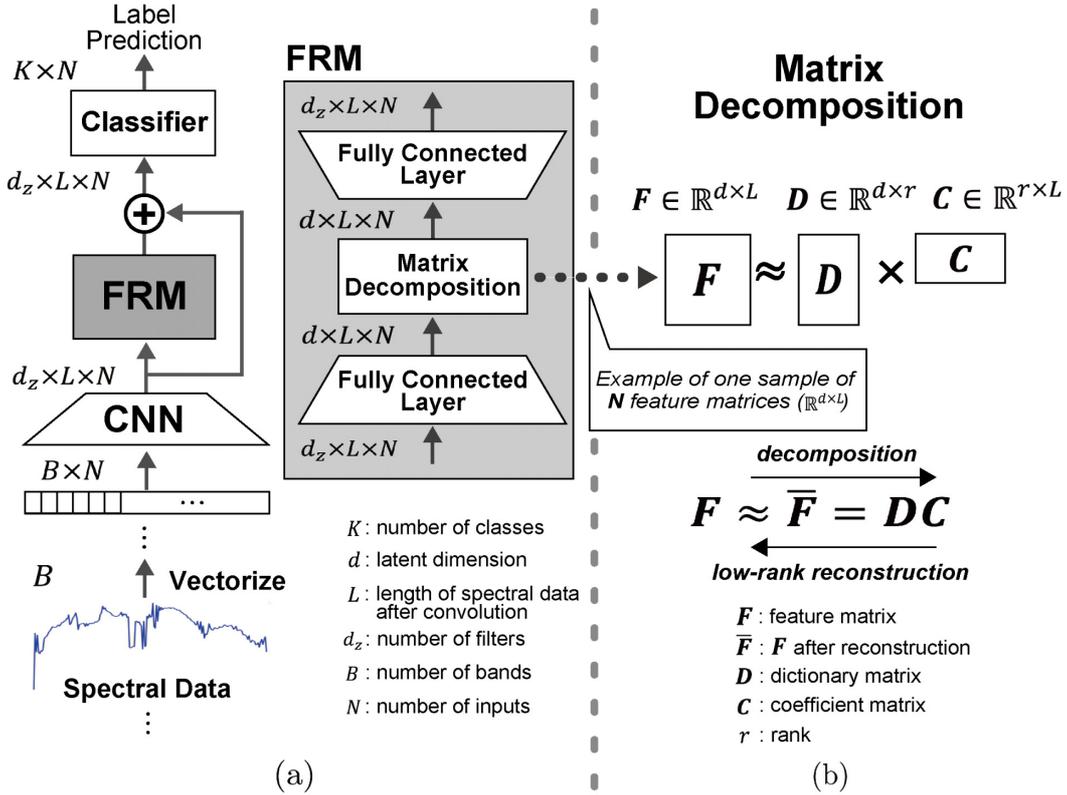


Fig. 2: Illustration of the network architecture, which depicts the introduction of FRM into 1D-CNN. (a) The overall network architecture, which indicates the positions of CNN layers, FRM, and the classifier. (b) The process of matrix decomposition in FRM, which shows how matrices are decomposed.

where  $\mathbf{x}_i$  is the original pixel vector,  $x_{ij}$  is the scalar value of the  $i$ -th pixel in the  $j$ -th band,  $x_{ij}^*$  is the normalized pixel value,  $\max(x_{:j})$  is the maximum value in band  $j$ , and  $\min(x_{:j})$  is the minimum value in band  $j$ . That is, each pixel vector  $\mathbf{x}_i$  is normalized band-wise.

We note that CNN-based models such as Hu et al.<sup>9)</sup> and Sidike et al.<sup>16)</sup>, which introduce FRM, change the activation function of the hidden layer from a hyperbolic tangent function to a Rectified Linear Unit (ReLU). Since we adopted NMF as FRM's matrix decomposition, the output of the middle layer input to NMF must satisfy the non-negative constraints. CNN-based networks of Hu et al. and Sidike et al. use a hyperbolic tangent function that converts input values from  $-1$  to  $1$  as the activation function of the hidden layer, so negative values ignored by FRM may affect classification. Therefore, ReLU is used as the activation function to avoid negative values because ReLU converts the input value  $u$  to  $u = 0$  for values  $u < 0$ , as in  $\max(u, 0)$ .

During the network training, the number of epochs is set to 1000 epochs with a batch size of 64, and the

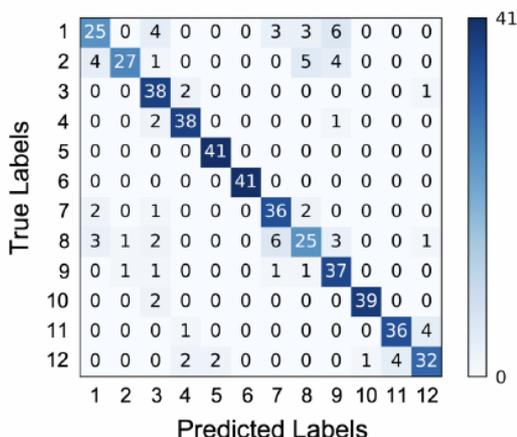
Table 1: Comparison of classification accuracy (%)

Methods	Dataset		
	BW	IP	UP
Hu et al. <sup>9)</sup>	93.23	84.35	92.34
- Introduction of FRM <sup>7)</sup>	<b>93.98</b>	<b>88.62</b>	<b>93.57</b>
dPEN (Sidike et al. <sup>16)</sup> )	93.61	83.33	92.87
- Introduction of FRM <sup>7)</sup>	<b>96.99</b>	<b>86.38</b>	<b>94.44</b>

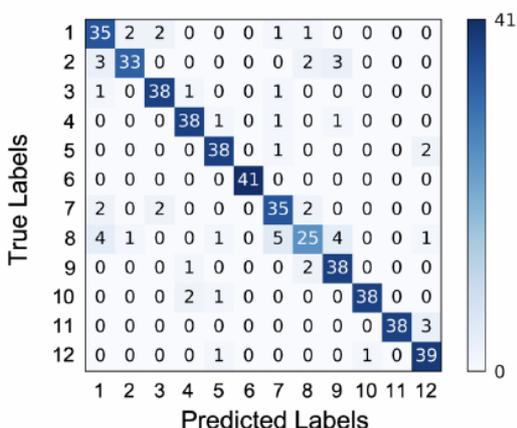
stochastic gradient descent is adopted to reduce the loss function with a learning rate of 0.001 and a weight decay of  $10^{-5}$ .

### 4.3 Classification Accuracy Evaluation

To evaluate the effectiveness of FRM, we compared the classification accuracy before and after introducing the FRM to CNN-based networks with the BW, IP, and UP datasets. The experimental results are shown in Table 1. Introduction of FRM achieves the best results in all datasets. It is noticeable that only introducing the FRM can improve the classification accuracy of dPEN<sup>16)</sup>, the SOTA method. For instance, comparing the methods introducing the FRM into dPEN and original dPEN<sup>16)</sup> on the BW, IP, and UP datasets, the



(a) Confusion matrix of 1D-CNN of Hu et al. Input is the test data of the IP dataset.



(b) Confusion matrix of 1D-CNN of Hu et al. with FRM. Input is IPデータセットのテストデータである。

図3:混同行列の比較。

それぞれ3.38%、3.05%、1.57%向上している。したがって、ピクセル単位のHSI分類ネットワークにFRMを導入することで、分類精度の面で効果的に性能を向上させることができることがわかる。

IPデータセットにおいて、Huらの1D-CNNにFRMを導入する前と後の混同行列を比較し、それぞれ図3(a)と図3(b)に示した。最も正しい予測が増加したクラス1のF1スコアは、特徴量の導入前と導入後にそれぞれ0.6667から0.8140に改善した。このように、FRMの導入により、偽陽性と偽陰性の数が減少し、より正確な評価が可能となった。

FRMの再構成に関する議論のための特徴マップの比較 4 分類精度向上のためのFRMの

#### 4.4 分類精度向上のための

表2:学習・推論時間とパラメータ数の比較。

Methods	Training (s)	Inference (s)	# Param.
Hu et al. <sup>9)</sup>	86.4	5.33	$4.562 \times 10^5$
- Introduction of FRM <sup>7)</sup>	194.4	5.69	$4.592 \times 10^5$
dPEN	228	5.74	$1.827 \times 10^7$
- Introduction of FRM <sup>7)</sup>	429	6.40	$1.828 \times 10^7$

再構成に関する議論のための特徴マップの比較 Huらの1次元CNNから得られた特徴マップを可視化した<sup>9)</sup> BW、IP、UPデータセットのFRMを導入し、図4に示すように可視化した。

BWデータセットの特徴マップを図4(a2), (a3)に示す。CNN特徴量の可視化(図4(a2))では、ほとんどの特徴マップが12~40の範囲を強調し、一部は全範囲を強調している。FRMの出力特徴量の可視化(図4(a3))では、全範囲を強調する特徴マップが減少し、12~40の範囲を強調する特徴マップが残った。行列分解による行列再構成もノイズ<sup>6)19)</sup>を除去する効果がある。FRMの再構成時にノイズと判定されたため、全範囲を強調する特徴マップは減少したと考えられる。

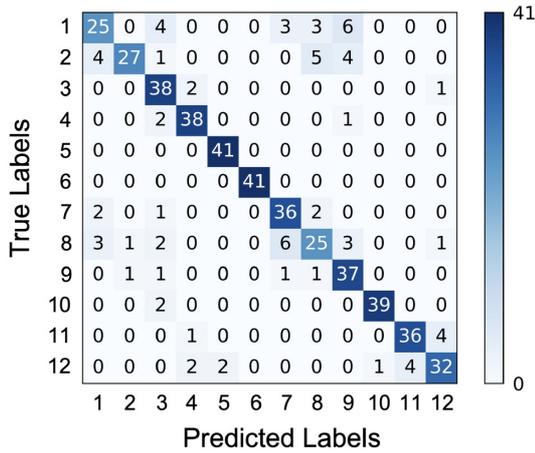
IPデータセットの特徴マップを図4(b2), (b3)に示す。CNN特徴量の可視化(図4(b2))では、ほとんどの特徴マップが全範囲を強調している。FRMの出力特徴量の可視化(図4(b3))では、ほとんどの特徴マップが全範囲を強調しているが、一部の特徴マップは値が減少している。FRMは特徴マップの特定の範囲を強調するように再構成するため、FRMは特徴マップのいくつかの値を減少させたと考えられる。

UPデータセットの特徴マップを図4(c2), (c3)に示す。このデータセットの可視化結果は、BWデータセットと同様であると考えられる。したがって、いくつかの特徴マップは再構成中にノイズであると判断されたと考えられる。

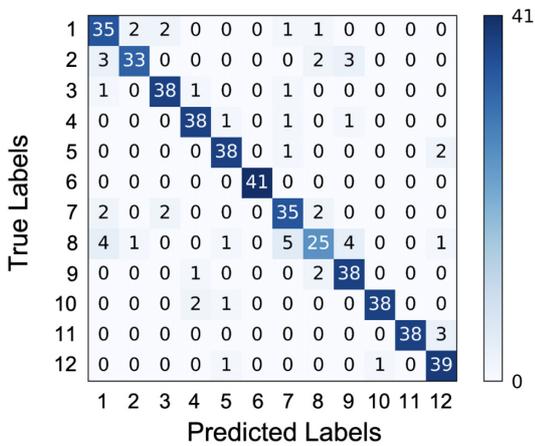
これらの結果は、FRMによって再構成された特徴マップは、元の特徴マップよりも特定の範囲を強調していることを示している。したがって、FRMは特定の範囲の特徴マップを強化し、これらの特徴マップは分類精度の向上に寄与したと考えられる。

#### 4.5 計算時間とパラメータ数の比較

FRM<sup>7)</sup>の導入により、計算時間とパラメータ数の変動を比較する。



(a) Confusion matrix of 1D-CNN of Hu et al. Input is the test data of the IP dataset.



(b) Confusion matrix of 1D-CNN of Hu et al. with FRM. Input is the test data of the IP dataset.

Fig. 3: Comparison of confusion matrices.

classification accuracy is 3.38%, 3.05%, 1.57% increase, respectively. Therefore, it indicates that the introduction of the FRM into the pixel-wise HSI classification network can effectively improve performance in terms of classification accuracy.

On the IP dataset, the confusion matrices before and after the introduction of the FRM into the 1D-CNN of Hu et al. were compared and shown in Fig. 3(a) and Fig. 3(b), respectively. The F1 score for Class 1, which had the most increase in correct prediction, improved from 0.6667 to 0.8140 before and after the introduction of the feature, respectively. Thus, the introduction of the FRM reduced the number of false positives and false negatives, indicating a more accurate evaluation.

#### 4.4 Comparison of Feature Maps for Discussion of Improved Classification Accuracy

In order to discuss what features were affected by the

Table 2: Comparison of training and inference times and number of parameters.

Methods	Training (s)	Inference (s)	# Param.
Hu et al. <sup>9)</sup>	86.4	5.33	$4.562 \times 10^5$
- Introduction of FRM <sup>7)</sup>	194.4	5.69	$4.592 \times 10^5$
dPEN	228	5.74	$1.827 \times 10^7$
- Introduction of FRM <sup>7)</sup>	429	6.40	$1.828 \times 10^7$

reconstruction of the FRM improving the classification accuracy, we visualized the feature maps obtained from 1D CNN of Hu et al.<sup>9)</sup> introducing the FRM on the BW, IP, and UP datasets, as shown in Fig. 4.

The feature maps for the BW dataset are shown in Fig. 4 (a2), (a3). In the visualization of the CNN feature (Fig. 4 (a2)), most feature maps highlighted the 12 to 40 range, and some highlighted the entire range. In the visualization of FRM's output feature (Fig. 4 (a3)), the feature maps that highlighted the entire range were reduced, leaving the feature maps that highlighted the 12 to 40 range. Matrix reconstruction by matrix decomposition also has the effect of removing noise<sup>6)19)</sup>. We consider that the feature maps that highlighted the entire range were reduced since they were determined to be noise during the reconstruction of the FRM.

The feature maps for the IP dataset are shown in Fig. 4 (b2), (b3). In the CNN feature visualization (Fig. 4 (b2)), most feature maps highlighted the entire range. In the visualization of FRM's output feature (Fig. 4 (b3)), most feature maps highlighted the entire range, but some of the feature maps had decreased values. We consider that the FRM reduced some values in the feature maps since the FRM reconstructed them to emphasize a specific range of feature maps.

The feature maps for the UP dataset are shown in Fig. 4 (c2), (c3). We consider the visualization results for this dataset to be similar to the BW dataset. Therefore, we consider that some feature maps were determined to be noise during the reconstruction.

These results indicate that the feature map reconstructed by the FRM emphasizes a specific range more than the original feature map. Therefore, we consider that the FRM enhanced the feature maps of a specific range, and these feature maps contributed to the improvement of classification accuracy.

#### 4.5 Comparison of Computation Time and Number of Parameters

We compare the variation in computation time and number of parameters with the introduction of the

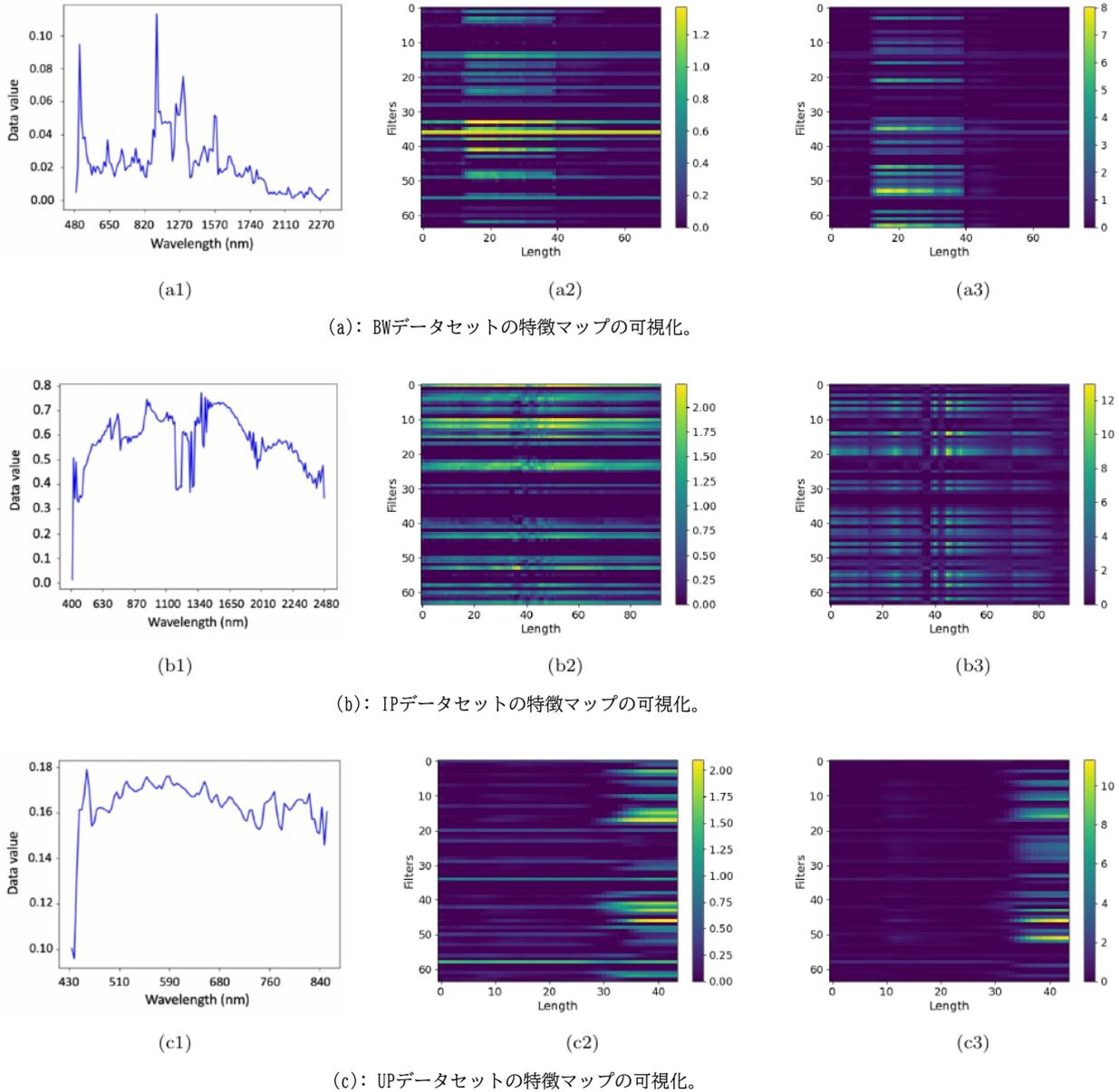


図4:3つのHSIデータセットにおけるFRMを導入したHuらの特徴マップの可視化。最初の列はピクセルスペクトルを示す。2列目はFRM入力前のcnn特徴マップである。最後の列は、FRMを出力するフェチュアマップである。

表 2 に FRM 導入前後の計算時間とパラメータ数を示す。学習時間は100エポックの学習に要する時間である。推論時間は、テストデータのラベルを予測するのに必要な時間である。

表2より、FRMの導入によりパラメータ数は増加したが、FRMは元のネットワークよりもパラメータ数が少ないため、その増加幅は小さいことがわかる。Huら<sup>9)</sup>の学習時間は約2.25倍、dPEN<sup>16)</sup>の学習時間は約1.88倍であった。

FRMを用いたモデルは、学習中に行列分解が更新されるため、計算時間が長くなったと考えられる。推論時間は、FRMを導入した場合、Huらでは約1.07倍、dPENでは約1.12倍となった。FRMの導入は、いずれの方法においてもパラメータ数の有意な増加にはつながらなかった。車載処理では推論のみを行い、学習<sup>1)</sup>は行わないため、学習時間よりも推論時間の増加の方が問題となる。

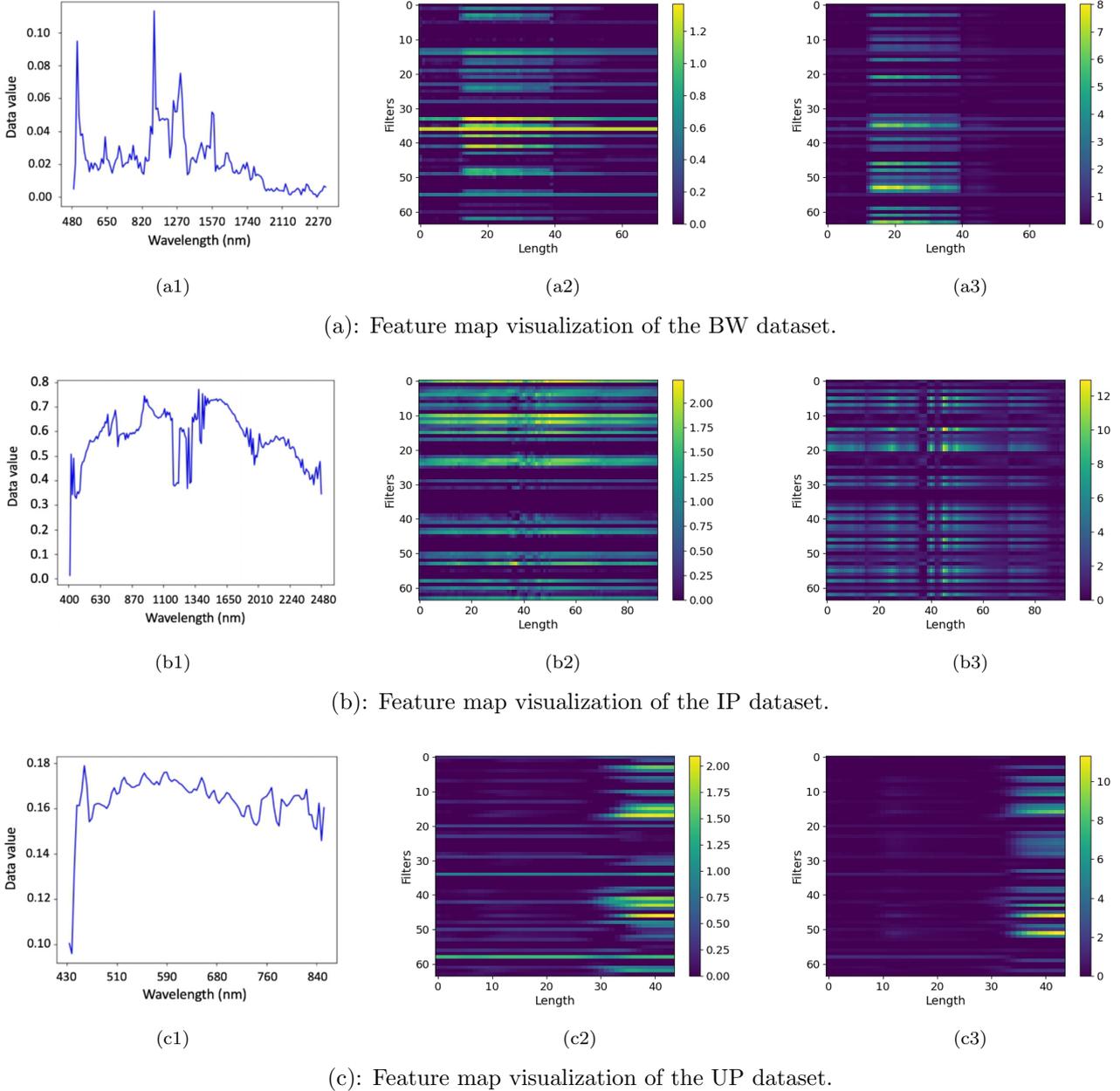


Fig. 4: Feature map visualization of Hu et al. introducing the FRM on three HSI dataset. The first column show the pixel spectra. The second column show the cnn feature map before input the FRM. The last column show the feture map of output the FRM.

FRM<sup>7)</sup>. Table 2 shows the computation time and the number of parameters before and after the introduction of the FRM. The training time is the time required to train 100 epochs. The inference time is the time required to predict the labels of the test data.

Table 2 shows that the number of parameters increased with the introduction of the FRM, but the increase was small since the FRM has fewer parameters than the original network. The training time required for Hu et al.<sup>9)</sup> was about 2.25 times longer, and for dPEN<sup>16)</sup> was about 1.88 times longer. We consider

that the models with the FRM took longer computation time since the matrix decomposition was updated during training. The inference time was about 1.07 times longer for Hu et al. and about 1.12 times longer for dPEN when the FRM was introduced. Introduction of the FRM did not lead to a significant increase in the number of parameters in either method.

The on-board processing only performs inference, not training<sup>1)</sup>, so the increase in inference time is more problematic than the training time. Since the increase in the number of parameters was suppressed while the

パラメータ数の増加は抑制され、推論時間の増加も抑制されたことから、FRMの導入はリモートセンシングコミュニティのオンボード処理にも適していると考えられる。

## 5. Conclusion

本論文では、CNNベースのピクセル単位のHSI分類ネットワークにおけるFRM<sup>7)</sup>の有効性を、分類精度の観点から評価した。実験では、3つのHSIデータセットにおいて、代表的なCNNベースの手法とSOTA手法にFRMを導入し、FRM導入前後の分類精度を比較した。FRMの導入により、全てのデータセットにおいて両手法の分類精度が向上することがわかった。したがって、FRMの導入は、ピクセル単位のHSI分類の分類精度を向上させるのに有効である。また、特徴マップの可視化、計算時間、パラメータ数を比較することで、FRM導入の効果を評価した。今後の課題としては、FRMによって再構成された特徴をさらに分析し、各タスクにより適した特徴を再構成する新しいFRMを開発することである。

## References

- Alcolea, A., Paoletti, M. E., Haut, J. M., Resano, J. and Plaza, A.: Inference in Supervised Spectral Classifiers for On-Board Hyperspectral Imaging: An Overview, *Remote Sensing*, Vol. 12, No. 3 (2020).
- Baumgardner, M. F. et al.: 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3 (2015).
- Cao, F., Yang, Z., Ren, J., Jiang, M. and Ling, W.-K.: Does Normalization Methods Play a Role for Hyperspectral Image Classification?, *ArXiv*, Vol. abs/1710.02939 (2017).
- Chehdi, K., Soltani, M. and Cariou, C.: Pixel classification of large-size hyperspectral images by affinity propagation, *Journal of applied remote sensing*, Vol. 8, No. 1, pp. 083567–083567 (2014).
- Chen, F., Jiang, H., Van de Voorde, T., Lu, S., Xu, W. and Zhou, Y.: Land cover mapping in urban environments using hyperspectral APEX data: A study case in Baden, Switzerland, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 71, pp. 70–82 (2018).
- Chen, Y., Huang, T.-Z., Zhao, X.-L. and Deng, L.-J.: Hyperspectral image restoration using framelet-regularized low-rank non-negative matrix factorization, *Applied Mathematical Modelling*, Vol. 63, pp. 128–147 (2018).
- Geng, Z. et al.: Is Attention Better Than Matrix Decomposition?, *Proc. of ICLR* (2021).
- Holzwarth, S., Muller, A., Habermeyer, M., Richter, R., Hausold, A., Thiemann, S. and Strobl, P.: HySens-DAIS 7915/ROSIIS imaging spectrometers at DLR, *Proceedings of the 3rd EARSeL workshop on imaging spectroscopy*, pp. 3–14 (2002).
- Hu, W. et al.: Deep convolutional neural networks for hyperspectral image classification, *Journal of Sensors*, Vol. 2015, ID.258619 (2015).
- Lee, D. D. and Seung, H. S.: Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol. 401, No. 6755, pp. 788–791 (1999).
- Lee, D. D. and Seung, H. S.: Algorithms for non-negative matrix factorization, *Advances in Neural Information Processing Systems (NIPS)* (2000).
- Melgani, F. and Bruzzone, L.: Classification of hyperspectral re-

- mote sensing images with support vector machines, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 8, pp. 1778–1790 (2004).
- Mueller, A. A., Hausold, A. and Strobl, P.: HySens-DAIS/ROSIIS imaging spectrometers at DLR, *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology*, Vol. 4545, pp. 225–235 (2002).
- Neuenschwander, A. L., Crawford, M. M. and Ringrose, S.: Results from the EO-1 experiment—A comparative study of Earth Observing-1 Advanced Land Imager (ALI) and Landsat ETM+ data for land cover mapping in the Okavango Delta, Botswana, *International Journal of Remote Sensing*, Vol. 26, No. 19, pp. 4321–4337 (2005).
- Sharma, A., Rani, S., Sah, D. K., Khan, Z. and Boulila, W.: HOMLC-Hyperparameter Optimization for Multi-Label Classification of Intrusion Detection Data for Internet of Things Network, *Sensors*, Vol. 23, No. 19 (2023).
- Sidike, P., Vasit, S., Maitiniyazi, M., Matthew, M., Nadia, S., Joel, B., Todd, M. and Felix, B. F.: dPEN: Deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery, *Remote sensing of environment*, Vol. 221, pp. 756–772 (2019).
- Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P. and Li, X.: A review of deep learning used in the hyperspectral image analysis for agriculture, *Artificial Intelligence Review*, Vol. 54, No. 7, pp. 5205–5253 (2021).
- Yang, J. M., Yu, P. T. and Kuo, B. C.: A Nonparametric Feature Extraction and Its Application to Nearest Neighbor Classification for Hyperspectral Image Data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 3, pp. 1279–1293 (2010).
- Zhang, H., He, W., Zhang, L., Shen, H. and Yuan, Q.: Hyperspectral Image Restoration Using Low-Rank Matrix Recovery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52, No. 8, pp. 4729–4743 (2014).



岡田 純京は工学部卒業。2022年東京電機大学大学院情報システム工学専攻修士課程修了、2024年東京電機大学大学院情報システム工学専攻修士課程修了。現在、博士号取得を目指している。東京電機大学大学院工学系研究科機械学習・コンピュータビジョン専攻研究テーマは物体認識とハイパースペクトルセンシング。



小篠 裕子 2015年神戸大学大学院工学研究科博士課程修了。2015年、産業技術総合研究所(AIST)博士研究員。2016年から2020年まで、慶応大学大学院理工学研究科研究員。2020年から2022年まで、東京電機大学大学院システムデザイン・技術研究科助教。2023年より東京電機大学大学院システムデザイン・技術研究科助教。大学。2024年より学院大学客員教授。研究テーマは、物体認識と接地、視覚認識、ハイパースペクトルセンシング。

increase in inference time was also suppressed, we believe that the introduction of the FRM is also suitable for on-board processing for the remote sensing community.

## 5. Conclusion

In this paper, we evaluated the effectiveness of the FRM<sup>7)</sup> in a CNN-based pixel-wise HSI classification network in terms of classification accuracy. In the experiment, we introduced the FRM to a representative CNN-based method and a SOTA method on three HSI datasets and compared the classification accuracy before and after the introduction of the FRM. We found that the introduction of the FRM improved the classification accuracy of both methods on all datasets. Therefore, the introduction of the FRM is effective in improving classification accuracy for pixel-wise HSI classification. We also evaluated the effects of introducing the FRM by comparing feature map visualization, computation time, and number of parameters. Our future work is to further analyze the features reconstructed by the FRM and to develop a novel FRM that reconstructs features to be more appropriate for each task.

## References

- 1) Alcolea, A., Paoletti, M. E., Haut, J. M., Resano, J. and Plaza, A.: Inference in Supervised Spectral Classifiers for On-Board Hyperspectral Imaging: An Overview, *Remote Sensing*, Vol. 12, No. 3 (2020).
- 2) Baumgardner, M. F. et al.: 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3 (2015).
- 3) Cao, F., Yang, Z., Ren, J., Jiang, M. and Ling, W.-K.: Does Normalization Methods Play a Role for Hyperspectral Image Classification?, *ArXiv*, Vol. abs/1710.02939 (2017).
- 4) Chehdi, K., Soltani, M. and Cariou, C.: Pixel classification of large-size hyperspectral images by affinity propagation, *Journal of applied remote sensing*, Vol. 8, No. 1, pp. 083567–083567 (2014).
- 5) Chen, F., Jiang, H., Van de Voorde, T., Lu, S., Xu, W. and Zhou, Y.: Land cover mapping in urban environments using hyperspectral APEX data: A study case in Baden, Switzerland, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 71, pp. 70–82 (2018).
- 6) Chen, Y., Huang, T.-Z., Zhao, X.-L. and Deng, L.-J.: Hyperspectral image restoration using framelet-regularized low-rank non-negative matrix factorization, *Applied Mathematical Modelling*, Vol. 63, pp. 128–147 (2018).
- 7) Geng, Z. et al.: Is Attention Better Than Matrix Decomposition?, *Proc. of ICLR* (2021).
- 8) Holzwarth, S., Muller, A., Habermeyer, M., Richter, R., Hausold, A., Thiemann, S. and Strobl, P.: HySens-DAIS 7915/ROSIS imaging spectrometers at DLR, *Proceedings of the 3rd EARSeL workshop on imaging spectroscopy*, pp. 3–14 (2002).
- 9) Hu, W. et al.: Deep convolutional neural networks for hyperspectral image classification, *Journal of Sensors*, Vol. 2015, ID.258619 (2015).
- 10) Lee, D. D. and Seung, H. S.: Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol. 401, No. 6755, pp. 788–791 (1999).
- 11) Lee, D. D. and Seung, H. S.: Algorithms for non-negative matrix factorization, *Advances in Neural Information Processing Systems (NIPS)* (2000).
- 12) Melgani, F. and Bruzzone, L.: Classification of hyperspectral remote sensing images with support vector machines, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 8, pp. 1778–1790 (2004).
- 13) Mueller, A. A., Hausold, A. and Strobl, P.: HySens-DAIS/ROSIS imaging spectrometers at DLR, *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology*, Vol. 4545, pp. 225–235 (2002).
- 14) Neuenschwander, A. L., Crawford, M. M. and Ringrose, S.: Results from the EO-1 experiment—A comparative study of Earth Observing-1 Advanced Land Imager (ALI) and Landsat ETM+ data for land cover mapping in the Okavango Delta, Botswana, *International Journal of Remote Sensing*, Vol. 26, No. 19, pp. 4321–4337 (2005).
- 15) Sharma, A., Rani, S., Sah, D. K., Khan, Z. and Boulila, W.: HOMLC-Hyperparameter Optimization for Multi-Label Classification of Intrusion Detection Data for Internet of Things Network, *Sensors*, Vol. 23, No. 19 (2023).
- 16) Sidike, P., Vasisit, S., Maitiniyazi, M., Matthew, M., Nadia, S., Joel, B., Todd, M. and Felix, B. F.: dPEN: Deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery, *Remote sensing of environment*, Vol. 221, pp. 756–772 (2019).
- 17) Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P. and Li, X.: A review of deep learning used in the hyperspectral image analysis for agriculture, *Artificial Intelligence Review*, Vol. 54, No. 7, pp. 5205–5253 (2021).
- 18) Yang, J. M., Yu, P. T. and Kuo, B. C.: A Nonparametric Feature Extraction and Its Application to Nearest Neighbor Classification for Hyperspectral Image Data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 3, pp. 1279–1293 (2010).
- 19) Zhang, H., He, W., Zhang, L., Shen, H. and Yuan, Q.: Hyperspectral Image Restoration Using Low-Rank Matrix Recovery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52, No. 8, pp. 4729–4743 (2014).



**Junkei Okada** Junkei Okada received the B.Eng. and M.Sc.Eng. degree in Information System Engineering from Tokyo Denki University, Japan, in 2022 and 2024, respectively. He is currently pursuing the Ph.D. degree in Machine Learning and Computer Vision at Tokyo Denki University. His research interests include object recognition and hyperspectral sensing.



**Yuko Ozasa** Yuko Ozasa received her PhD degree in engineering from Kobe University in 2015. She was a postdoctoral researcher at National Institute of Advanced Industrial Science and Technology (AIST) in 2015. She had been a research associate at Graduate School of Science and Technology, Keio University from 2016 to 2020. She had been an assistant professor at Graduate School of System Design and Technology, Tokyo Denki University from 2020 to 2022. Since 2023, she has been an associate professor at the university. She has also been a guest professor at Gakushuin University since 2024. Her research interests include object recognition and grounding, visual perception, and hyperspectral sensing.