DETERMINATION OF RELEVANT HYPERSPECTRAL BANDS USING A SPECTRALLY CONSTRAINED CNN

Ulrike Pestel-Schiller, Kai Hu, Daniel Gritzner, Jörn Ostermann

Institut für Informationsverarbeitung, Leibniz Universität Hannover, Appelstr. 9A, 30167 Hannover, Germany

ABSTRACT

For transmitting the large amount of hyperspectral image (HSI) data over a small data link from a small platform to the ground, an efficient data compression with low computational cost has to be done at the platform. Additionally, spectral band reduction interpreted as preprocessing of the compression is reasonable. We present a method for hyperspectral band reduction using a modified convolutional neural network (CNN) which retains the information about the spectral origin from layer to layer until it can be assigned directly to the classes to be classified. The relevant bands for each class are determined. Experimental verification shows that the network architecture using only the relevant bands has improved stability and results in a better overall performance.

Index Terms— hyperspectral band reduction, irrelevance reduction, CNN

1. INTRODUCTION

Materials can be distinguished by their spectral characteristics of absorption or reflectance. Based on this, hyperspectral image (HSI) sensing allows the detection and identification of specific materials, e.g. for geological mapping, for monitoring agriculture and forest status, for environmental studies, for search and rescue services, for disaster managment or for surveillance. For each hyperspectral image pixel, a HSI sensor provides more than one hundred narrow and spectrally contiguous channels with a bandwidth of a few nanometers, ranging from the visible to infrared regions of the electromagnetic spectrum.

Having airborne applications in mind, we have to transmit the large amount of data from a platform to the ground. That requires either a large data link or a data compression already done at the platform. We strive for a small data link. For powerful data compression, standardized coding algorithms, e.g. High Efficiency Video Coding (HEVC) [1] can be applied: the two spatial dimensions in HSI are similar to a video image; the third dimension of the HSI, the spectral dimension, can be handled like the temporal dimension of a video codec in order to reduce the spectral correlation [2]. Furthermore, having in mind small platforms, we have to look for solutions with low computational cost. On the ground a classification, semantic segmentation or anomaly detection can be applied.

Typically, classification algorithms first of all reduce the spectral dimensionality [3]. Thus, it is very efficient to start with a spectral band reduction before the transmission. Hence, spectral band reduction can be interpreted as a preprocessing of the coding. Coding using HEVC is like a general purpose redundancy and irrelevance reduction; whereas spectral band reduction is a task specific irrelevance reduction which can not be carried out by the codec. The band reduction deals with the determination of the relevant bands for a specific classification task and can be done offline.

A general overview on hyperspectral band reduction is given in [3]. In the area of ranking-based methods for classification, on the one hand many unsupervised methods like the energy-based PCA in [4] and the feature-selection-based method in [5] are known, on the other hand many supervised methods e.g. based on feature extraction [6] or statistical dependencies [7] exist. Typically, all these approaches try to maximize the classification performance.

We propose a supervised ranking-based method where we look for a solution which selects relevant bands for broadly defined classes. Primarily, we want to omit definitly irrelevant bands. We modify a simple convolutional neural network (CNN) [8] such that the information about spectral wavelength is retained until the class decision determining the relevant bands. The method is applied to natural and man-made materials which is intended to be visually indistinguishable.

The remainder of the paper is organized as follows: In Sec. 2 we propose the band reduction method by describing and analyzing the spectrally constrained CNN together with the method for determining the relevant bands. We present and discuss the experimental verification in Sec. 3 where the applied data is described and the determination of the relevant bands is presented. Furthermore, we compare the classification performance using the determined relevant bands to that using all bands. Sec. 4 concludes the paper.

2. BAND REDUCTION METHOD

We propose a method for spectral band reduction using a spectrally constrained CNN. The designed network architec-

ture is introduced in Subsec. 2.1. In Subsec. 2.2 the network architecture is analyzed. The determination of the relevant bands is derived in Subsec. 2.3.

2.1. Spectrally constrained CNN

The network architecture comprises the following layers:
A **3D input layer** expects one hyperspectral pixel as CNN input data of size $1 \times 1 \times N$ where N is the number of spectral bands. The output is filtered in spectral dimension by a **3D convolutional layer** and applying M filters each with three coefficients, that means that only three neighboring spectral bands are considered for one output value of the 3D convolutional layer. Therefrom, we call the CNN spectrally constrained (scCNN). The output of the convolutional layer results in a matrix of size $1 \times 1 \times N \times M$.

The 3D convolutional layer is followed by a **batch normalization layer**, a **rectified linear unit layer** and a **fully connected layer**. At its output we get K logits used for classification.

2.2. Analyzing the network architecture

The elements of the third dimension of the CNN input can be assigned to the wavelengths of N bands. After the convolution we get an output of size $1 \times 1 \times N \times M$, The number of filter coefficients is only three which means that only three neighboring spectral bands are taken into account by each filter. This corresponds to a spectral selection. Hence, we call our network architecture spectrally constrained. Because of this spectral constraint, the third dimension still represents the wavelength and the network architecture retains this information up to the input F of the fully connected layer. On the other hand, the classification can be retraced directly to the output H of the fully connected layer. Its output $H = (H_1, \cdots, H_k, \cdots, H_K)$ is directly assigned to the K classes.

The output H_k of class k is computed by the fully connected layer as

$$H_k = \sum_{n=1}^{N} \sum_{m=1}^{M} (f_{mn} \cdot w_{kmn}) \tag{1}$$

where the parameters f_{mn} represent the input F over the N spectral bands and M features, and the parameters w_{kmn} represent the parameters $W = (W_1, \cdots, W_k, \cdots, W_K)$ of the fully connected layer. The parameters of class k are given by

$$W_k = \begin{bmatrix} w_{k11} & \dots & w_{k1n} & \dots & w_{k1N} \\ \vdots & & \ddots & & \vdots \\ w_{km1} & \dots & w_{kmn} & \dots & w_{kmN} \\ \vdots & & \ddots & & \vdots \\ w_{kM1} & \dots & w_{kMn} & \dots & w_{k\,M\,N}. \end{bmatrix} \tag{2}$$

2.3. Determination of the relevant spectral bands

By analyzing and calculating the arithmetic operation of the fully connected layer, the relevant bands for each class can be determined as explained in the following. For a better understanding, without loss of generality, we consider one class k only.

Eq. 1 can be written as

$$H_k = \sum_{n=1}^{N} \left\{ \sum_{m=1}^{M} f_{mn} \cdot w_{kmn} \right\}. \tag{3}$$

In a first step the Hadamard product of the input matrix F with the parameter matrix W_k results in a matrix

$$F \cdot W_k = \begin{bmatrix} f_{11} w_{k11} & \dots & f_{1n} w_{k1n} & \dots & f_{1N} w_{k1N} \\ \vdots & & \ddots & & \vdots \\ f_{m1} w_{km1} & \dots & f_{mn} w_{kmn} & \dots & f_{mN} w_{kmN} \\ \vdots & & \ddots & & \vdots \\ f_{M1} w_{kM1} & \dots & f_{Mn} w_{kMn} & \dots & f_{MN} w_{kMN} \end{bmatrix}. \tag{4}$$

In a second step, the inner summation over M sums up the features contained in a column n of Eq. 4 to

$$g_{k\,n} = \sum_{m=1}^{M} (f_{mn} \cdot w_{kmn}). \tag{5}$$

As a reminder, each sum g_{kn} at position n is assigned to the wavelength of the spectral band n.

The column sums g_{kn} can be grouped into a vector and normalized to

$$G_{k,norm} = \frac{(g_{k1}, \dots, g_{kn}, \dots, g_{kN})}{max(g_{k1}, \dots, g_{kn}, \dots, g_{kN})} \tag{6}$$

of length N where high values indicate that they contribute significantly to the class decision.

From $G_{k,norm}$ we derive a class relevance vector $R_k = (r_{k1}, \dots, r_{kn}, \dots, r_{kN})$ of class k by considering only high values above a threshold T_k by

$$r_{kn} = \begin{cases} 1 & \text{for } \frac{(g_{kn})}{max(g_{k1}, \dots, g_{kn}, \dots, g_{kN})} > T_k \\ 0 & \text{else} \end{cases}. \tag{7}$$

We can calculate R_k for all k classes and derive the relevance vector $R = (r_1, \dots, r_n, \dots, r_N)$ by

$$r_n = \begin{cases} 1 & \text{for } \sum\limits_{k} r_{kn} \geq 1 \\ 0 & \text{else} \end{cases} \tag{8}$$

where for all relevant bands R is equal to 1.

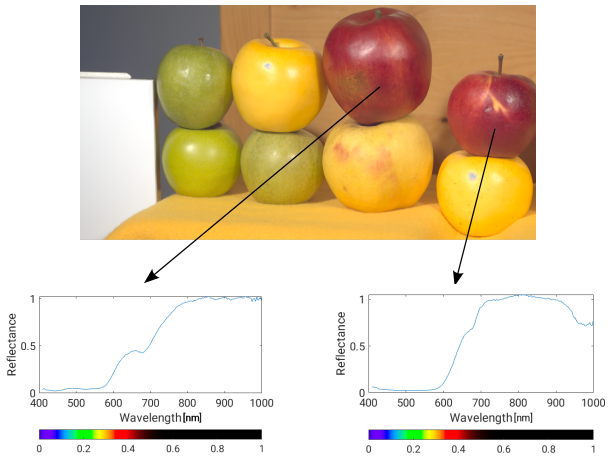
3. EXPERIMENTAL VERIFICATION

In the experimental verification we considered the two classes 'natural fruits' and 'man-made fruits' which had to be classified. The recorded natural fruits are green, yellow, red and multicolor apples, lemons, quince, melon, banana, green pepper, cucumber, pumpkin and avocado; the recorded man-made fruits are green, yellow, red and multicolor apples, lemon, green and orange pepper (Sec. 3.1).

We applied the proposed network (Sec. 2.1) architecture. Determining the relevant spectral bands is described in Sec. 3.2. In Sec. 3.3 test accuracies using the determined relevant spectral bands are compared with those using all spectral bands.

3.1. Recording and preparation of data

Several HSI were recorded using a NEO VNIR-1800 sensor with 186 spectral bands ranging from 400 to 1000 nanometers and 1800 samples/line. The natural and man-made fruits were recorded under different light conditions using day-light lamps, halogen lamps and natural daylight. Each recorded HSI was converted to HSI reflectance data using a 100% reflecting calibration target as shown in the image in Fig. 1 on the left side.



(a) man-made red apple (b) natural red apple

Fig. 1. Spectral signatures of HSI shown with RGB image

HSI pixels of both classes were manually picked from 4 recorded HSIs resulting in 7200 HSI pixels in the class 'natural fruits' and 4080 HSI pixels in the class 'man-made fruits' as shown in Table 1 (available under [9]). Associated spectral signatures are shown in Fig. 1. The HSI was partitioned into blocks of 8x8 HSI pixels, each block was averaged resulting in one preprocessed HSI pixel with $N = 186$ each. The averaging smoothes the spectral signature by reducing noise.

Table 1. TNT Hyperspectral Test Set Fruits

Data Set (No.of pixel)	
natural (7200)	man-made (4080)
green apple (960)	green apple (960)
yellow apple (480)	yellow apple (480)
red apple (480)	red apple (480)
multicolor apple (480)	multicolor apple (960)
lemon (1600), banana (400)	lemon (400)
yellow pepper (400)	orange pepper (400)
green pepper (400)	green pepper (400)
melon (400), quince (400)	
pumpkin (400), avocado (400)	
cucumber (400)	

3.2. Relevant spectral bands

The preprocessed HSI pixel constitutes the CNN input data. By varying the number of features M in the convolution layer from 2 to 20 in steps of 2, it was found out that this hyperparameter combines high accuracy and small complexity best at $M = 8$. The initial learning rate was set to 0.01.

The proposed scCNN was trained with 1760 pixels of natural fruits (green and yellow apples, banana, lemon) and 960 pixels of man-made fruits (green and yellow apples). The data of the remaining fruits was used for testing. As soon as the test data showed that the network parameters were well trained, we started to determine the relevant bands.

For the determination of relevant bands we used *all* training and test data to ensure that the resulting bands are valid for broad classes. Using Eq. 7 we chose the thresholds $T_1 = 0.015$ and $T_2 = 0.025$ to eliminate noise in $G_{k,\,norm}$. This results in the class relevances R_1 for 'natural fruits' and R_2 for 'man-made fruits', the associated $G_{k,\,norm}$ is illustrated in Fig. 2 for both classes.

Fig. 2 shows that essentially the spectral bands belonging to the VNIR spectrum are relevant for the classification. The bands in the electro-optical spectrum do not influence the classification crucially. This is not surprising since the man-made fruits have been designed to look like real ones. Differences in the electro-optical spectrum are being ignored, accordingly.

With Eq. 8 we get the relevant bands 90 to 152 and 158 to 185.

3.3. Applications

We applied the proposed network architecture in an experimental setup (Sec. 3.3.1) using the data of Sec. 3.1. The hyperparameter N, number of spectral bands, contains either the relevant bands or is equal to all bands (Sec. 3.3.2).

For a more rigorous testing the training data was reduced to 960 HSI pixels of natural respective man-made apples (green and yellow).

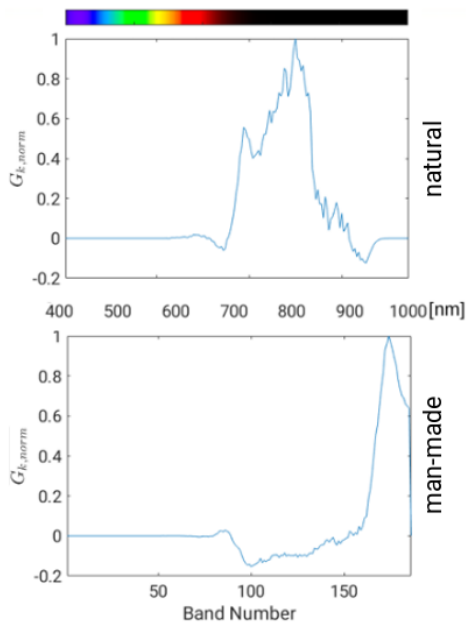


Fig. 2. Spectral band contribution to class decision

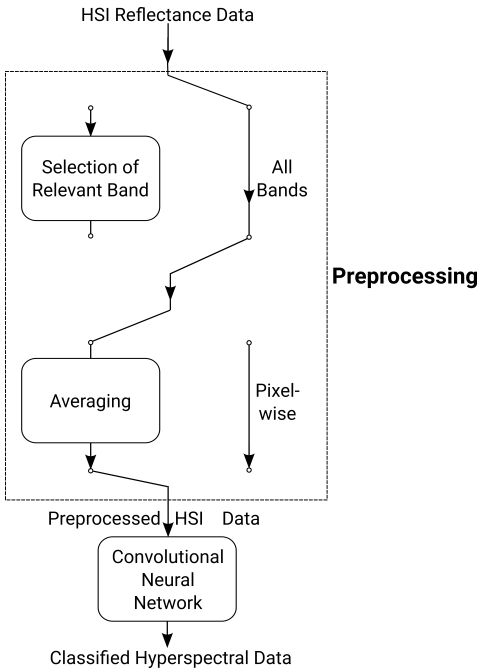


Fig. 3. Experimental setup

Table 2. Modes of the experimental setup

Mode	Training	Test
1	averaged	averaged
2	averaged	pixelwise
3	pixelwise	pixelwise

Starting with 10 different randomly initialized network parameter sets, we trained the CNN 10 times for each hyperparameter N resulting in 10 different sets of trained network parameters $P_A, P_B \ldots, P_E, \ldots, P_J$ each.

3.3.1. Experimental Setup

Fig. 3 shows the block diagram of the experimental setup. First, the HSI reference data is preprocessed. The preprocessing comprises an optional spectral band reduction to the relevant bands and an optional averaging as in 3.1. The output of the preprocessing represents the data set of the CNN. After the preprocessing, one HSI pixel or an averaged HSI pixel represent the N spectral values of the CNN input data. The number of features is $M = 8$ as in 3.2.

Three different modes were investigated by varying the averaging in the preprocessing. The different modes are shown in Table 2. We used the same network architecture for all modes. Only the hyperparameter of the spectral input dimension varied depending on the reduction of bands in the preprocessing.

The data set had to be partitioned into either training & validation data or test data for using a CNN. The training & validation process contained 1920 pixels, the test process the remaining pixel.

3.3.2. Classification

First, let us consider the relevant spectral bands determined in Sec. 3.2. Fig. 4 shows the resulting test accuracy of all three modes for 10 different sets of network parameters.

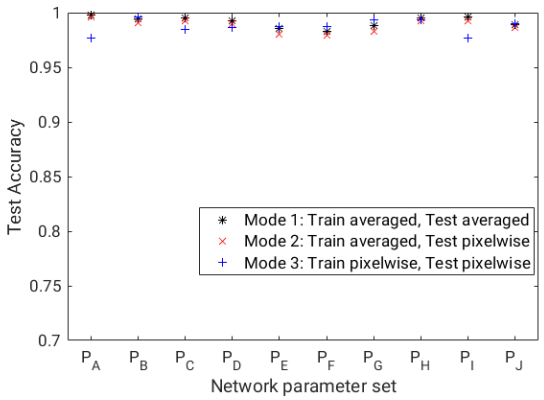


Fig. 4. Test accuracy using relevant bands

Next, let us consider all 186 spectral bands. Fig. 5 shows the resulting test accuracies of all three modes for 10 different sets of network parameters each.

For a better comparison we evaluated the mean m and standard deviation σ of the test accuracies over the sets of network parameters of each mode for both investigated band

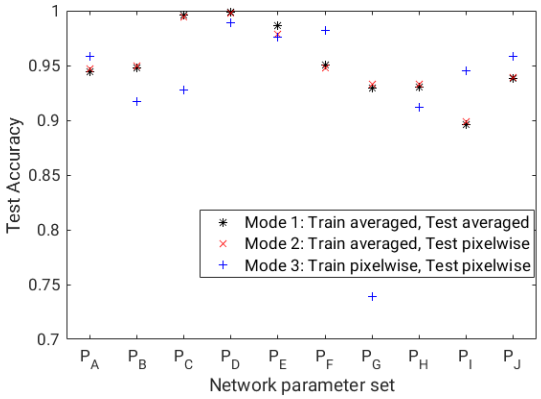


Fig. 5. Test accuracy using all bands

Table 3. Evaluation of scCNN

Mode	m_{scCNN}	m_{all}	σ_{scCNN}	σ_{all}
1	0.992	0.952	0.005	0.033
2	0.988	0.952	0.006	0.031
3	0.987	0.930	0.007	0.072

numbers shown in Table 3.

For the scCNN all three modes result in very high test accuracies for all network parameters. The mean accuracy in all three modes is higher compared to that of the all-band-CNN (Table 3). Fig. 4 and Fig. 5 show that the scCNN works more steadily as supported by comparison of the standard deviations in Table 3. Thus, the small training data set works well with the reduced number of bands.

When using a smoothed spectral signature (mode 1 and 2) in the training process, the results are more steady compared to training processes on a noisy spectral signature (mode 3). This shows that the scCNN is noise sensitive. Otherwise, using averaged or pixelwise test data (mode 1 or 2) influences the test accuracy only slightly. Hence, mode 2 is appropriate as a starting point for semantic segmentation where a standard segmentation can be applied spatially and the proposed approach orthogonally in the spectral dimension.

4. CONCLUSION

We presented a method for hyperspectral band reduction using a spectrally constrained CNN which we modified such that it retains the information about the spectral origin from layer to layer until it can be assigned directly to the classes to be classified. The relevant bands for each class were determined. Experimental verification showed that the network architecture using only the relevant bands results in a better and more steady overall performance.

5. REFERENCES

[1] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[2] Ulrike Pestel-Schiller, Karsten Vogt, Jörn Ostermann, and Wolfgang Groß, "Impact of hyperspectral image coding on subpixel detection," in *Proceedings of 32nd Picture Coding Symposium*, Dec. 2016, vol. http://dx.doi.org/10.1109/PCS.2016.7906396.

[3] Weiwei Sun and Qian Du, "Hyperspectral band selection: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, pp. 118–139, 06 2019.

[4] Craig Rodarmel and Jie Shan, "Principal component analysis for hyperspectral image classification," *Surv Land inf Syst*, vol. 62, 01 2002.

[5] Patrick Erik Bradley, Sina Keller, and Martin Weinmann, "Unsupervised feature selection based on ultrametricity and sparse training data: A case study for the classification of high-dimensional hyperspectral data," *Remote Sensing*, vol. 10, no. 10, 2018.

[6] Rui Huang and Mingyi He, "Band selection based on feature weighting for classification of hyperspectral data," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 156–159, 2005.

[7] S. Feng, Y. Itoh, M. Parente, and M. F. Duarte, "Hyperspectral band selection from statistical wavelet models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 2111–2123, 2017.

[8] Asifullah Khan, Anabia Sohail, Umme Zahoora, and Aqsa Saeed Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *CoRR*, vol. abs/1901.06032, 2019.

[9] Inst. f. Informationsverarb. (TNT), Leibniz Universität Hannover, "TNT Hyperspectral Test Set Fruits," https://www.tnt.uni-hannover.de/staff/pestel/HyperspectralTestSetFruits.zip 2019-2020.