

PAPER • OPEN ACCESS

Spectral-spatial neural network classification of hyperspectral vegetation images

To cite this article: V V Podlipnov *et al* 2023 *IOP Conf. Ser.: Earth Environ. Sci.* **1138** 012040

View the [article online](#) for updates and enhancements.

You may also like

- [Verification of stereotactic cranial radiotherapy treatments with MR-based gel dosimeters: practical aspects](#)

Filipa Costa, Evanthia Kousi, Anne Gasnier et al.

- [Design of Automatic Plant Areas Using Humidity Sensor Based On Internet of Thing](#)

G Gunawan, Marliana Sari, Benar Surbakti et al.

- [Collaborative artificial bee colony k-mean clustering algorithm for mixed data set](#)

C Nalini, J Sudeeptha, B S Charulatha et al.



244th Electrochemical Society Meeting

October 8 – 12, 2023 • Gothenburg, Sweden

50 symposia in electrochemistry & solid state science

Abstract submission deadline:
April 7, 2023

Read the call for
papers &
submit your abstract!

Spectral-spatial neural network classification of hyperspectral vegetation images

V V Podlipnov^{1,2}, N A Firsov¹, N A Iylyev^{1,2}, S V Mashkov³, P A Ishkin³,
R V Skidanov^{1,2} and A V Nikonorov^{1,2}

¹ Samara National Research University, 34, Moskovskoye shosse, Samara, 443086, Russia

² Image Processing Systems Institute of RAS – Branch of the FSRC “Crystallography and Photonics” RAS, Samara, 443001, Russia

³ Samara State Agrarian University, 2, Uchebnaya Str., Ust-Kinel, Kinel, 446442, Russia

E-mail: podlipnovvv@ya.ru

Abstract. The paper presents the results of a study of the method of classification of vegetation types based on hyperspectral survey data. A hyperspectrometer based on the Offner scheme was used as a shooting equipment. The spatial-spectral architecture of a convolutional neural network is used as a classifier. The experiments performed show the effectiveness of the proposed approach for the classification of vegetation types.

1. Introduction

Hyperspectral remote sensing of the Earth (remote sensing) allows to obtain multi-channel images of the underlying surface, each channel of which corresponds to a certain wavelength. Thanks to this, most of the information of the real-world scene is preserved in the visible and near infrared range. Each pixel of such an image is a spectral vector that can be used to characterize the object in the image [1]. Hyperspectral remote sensing of the Earth from Space and from unmanned aerial vehicles is successfully used in the tasks of agriculture, forestry, geology and, unlike RGB color shooting or multispectral, has several tens or even hundreds of channels corresponding to specified wavelengths and allows you to see even small differences in vegetation, up to sensitivity to the presence of chemical elements and minerals [2, 3]. The classification of such plots of land is important.

Previously, classical methods were used to classify hyperspectral images, such as the k-nearest neighbor method, resolution trees, and linear classifiers [4]. Their main problem was sensitivity to the Hughes effect [5]. The method of support vectors coped well with this problem, remaining the main method of classification of hyperspectral images [6].

In most cases, focusing only on the spectral properties of vegetation does not allow us to assert the reliability of the resulting classification of vegetation types. Many types of vegetation have quite similar spectra. The absorption lines of chlorophyll and carotenoids are located in similar parts of the spectrum and their intensities are similar. In addition, it is necessary to take into account the seasonal vegetative change in the color of vegetation. All these factors significantly complicate the task of classification.

For many purposes of the problem being solved, it may be useful to combine spectral data with any other additional ones that allow observing additional properties of plants. Combining hyperspectral



proximal sensing data with thermography makes it possible to increase the accuracy of water stress assessment. [7] Hyperspectral survey data can also be combined with data from fluorescence analysis, 3D scanning, etc. [8] To increase classification accuracy, a textural feature can be additionally considered. Textural features inherent in each type of vegetation can significantly increase the accuracy of classification [9]. Neural network-based algorithms can be used to analyze the texture.

In this paper, we conduct a study of the approach to the classification of high-resolution GIS in the applied problem of determining vegetation types, taking into account both the spatial and spectral components of the GIS. A convolutional neural network with a modified architecture is used as a classifier to better account for changes in scene lighting [6]. An automated algorithm based on an adaptive vegetation index is proposed for the formation of a training data set. The effectiveness of the proposed classification approach based on the data of surveys of agricultural land by a hyperspectral camera based on the Offner scheme is shown.

Also, as an addition, we use the analysis of vegetation indices in the algorithm, which allows us to make a primary classification of vegetation cover from other areas of the underlying surface without vegetation or with dry plant residues.

2. Materials and Methods

2.1. Obtaining hyperspectral images

For the research, a field site was selected with two crops growing relatively close: corn and oats. Also, the site for conducting experiments on the border of agricultural crops contains weed plants. This is mainly the rubbing of the ordinary schiritsa (amaranth). The studied site for the survey is located in the Samara region, Privolzhsky district, Privolzhye, fields of the farm Tsiruleva E.P. in the vicinity of coordinates latitude 52.81, longitude 48.61. The date of the survey is July 10–20, 2021. A hyperspectrometer based on the Offner scheme was used for shooting [10] The hyperspectrometer has a spatial resolution of 2040 pixels along the slit or 1020 pixels in binning shooting mode. During the shooting, the slit was positioned perpendicular to the scanning vector, and the scanning was carried out by rotation using a tripod with a rotary mechanism, forming a panorama of the shooting area. The camera itself is located at an angle relative to the horizon, the installation angle varied in the range of 30–60°. The height of the tripod is about 2m. To obtain panoramic hyperspectral images, the shooting speed was selected by the rotation drive synchronously with the shooting of frames, trying to carry out spectral shooting without gaps, which allows the spatial component to be transmitted as accurately as possible. Figure 1 shows the experimental survey scheme.



(a)



(b)

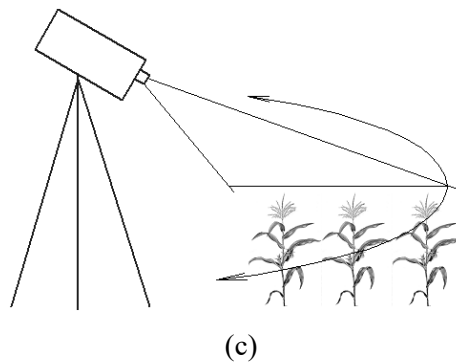


Figure 1. Obtaining hyperspectral images of vegetation using a scanning hyperspectrometer on a rotary platform: (a) – field research area; (b) – X-Rite ColorChecker location in the scanning area of the camera; (c) – space scanning scheme.

The exact exposure time was selected according to the histogram during shooting with the camera fixing the frame aimed at the white area of the calibrated X-Rite ColorChecker Video color scale. The reconstruction of the GSI from the resulting video sequence was performed according to previously developed algorithms described in [11–12]. The obtained GSI have 250 spectral channels evenly distributed in the range of 420–980 nm. The frame rate in all scenes is fixed and corresponds to 15 fps, which ensures the constancy of spatial resolution on all received GPS. Figure 2 shows an example of one of a number of similar color-synthesized images obtained from a hyperspectrometer-captured hypercube panorama of a field section.

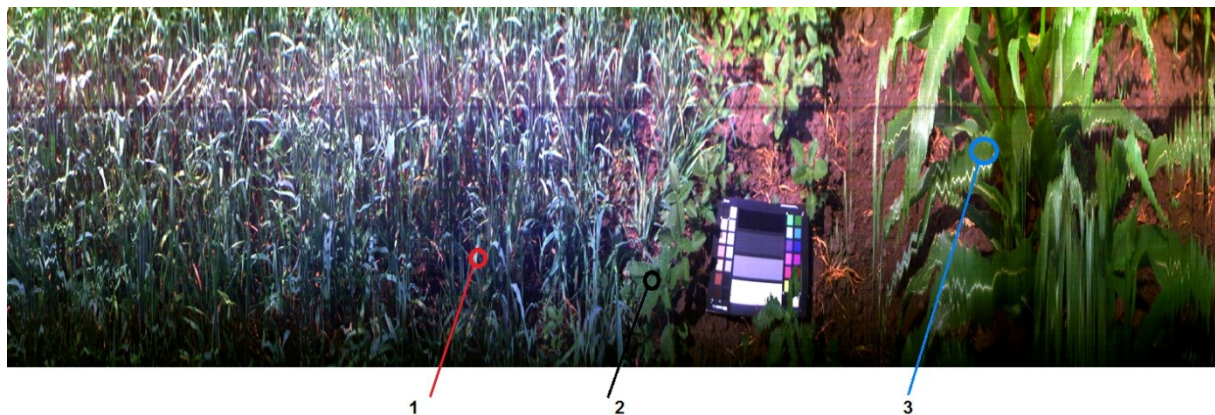


Figure 2. RGB representation of a hyperspectral cube and vegetation growth areas from which spectral profiles were taken: 1 – oats; 2 – schiritsa; 3 – corn.

Figure 3 shows the spectral profiles of hyper pixels corresponding to the areas marked in Figure 2.

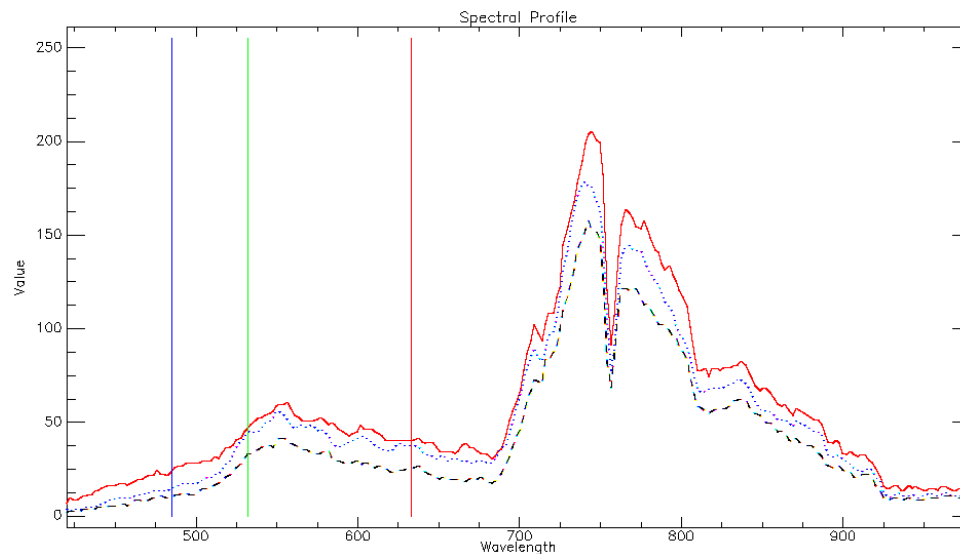


Figure 3. Spectral profiles of hyperpixels of the plot image. The red line is oats, the blue dotted line is corn, the black dotted line is ordinary schiritsa.

2.2. Hyperspectral data processing

For the problem of binary segmentation of soil and vegetation, two vegetation indices were considered: NDVI (normalized difference vegetation index) [13] and EVI (enhanced vegetation index) [14]. These indexes are used to find photosynthetically active biomass:

$$NDVI = \frac{NIR-RED}{NIR+RED} \quad (1)$$

$$EVI = 2.5 \cdot \frac{NIR-RED}{NIR+6 \cdot RED-7.5 \cdot BLUE+1} \quad (2)$$

where NIR is the near infrared region of the spectrum; RED is the red region of the spectrum; BLUE is the blue region of the spectrum. The specific spectral range in (1) and (2) is selected based on the specifics of the recording equipment and the problem being solved [15].

The calculated index masks are subjected to threshold processing in such a way that all pixels belonging to the “soil” class are zeroed (Figures 2 and 3), leaving values greater than zero only in the “plants” class.

Since the difference in the results is minimal for the hyperspectral images presented, NDVI (figure 4) was chosen as the vegetation index as simpler for calculations.



Figure 4. NDVI image.

Next, the quality of binary segmentation is verified. To do this, hyper pixels are taken from the hyperspectral image in turn and distributed, depending on the value of the index mask in a pixel, into two arrays of vectors – “soil” and “plants”.

For them, t-statistics are considered, which in this case is 0.00011, which is significantly less than $p\text{-value} = 0.05$, and from this it can be argued that vectors belonging to the classes “soil” and “plants” are segmented correctly with a high probability.

Further segmentation by vegetation classes is carried out manually and t-statistics are used for its validation (Figure 4).

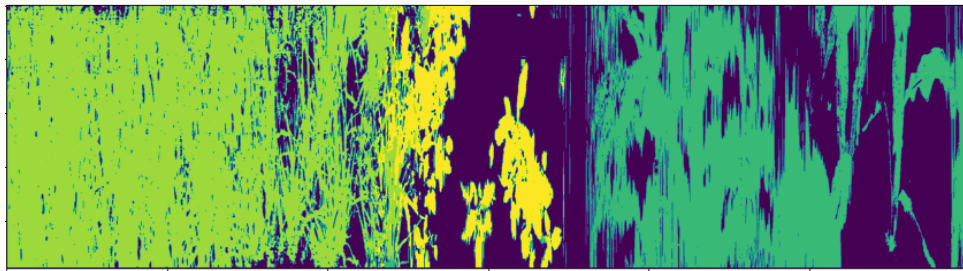


Figure 5. Result of manual segmentation.

3. Results and Discussion

To calculate the t-statistics for this mask, 1000 vectors belonging to certain classes were randomly selected from the hyperspectral image (Table 1).

Table 1. Values of p-value t-statistics between classes of plants.

Class_1 / Class_2	t-statistics
oats / weed	$4.1 * 10^{-137}$
weed / corn	$1.7 * 10^{-79}$
oats / corn	$1.81 * 10^{-5}$

Based on the data obtained, it can be argued that manual segmentation of plants is highly likely correct.

The change in the spectral composition of illumination has a significant effect on the result of hyperspectral shooting [16]. The survey of agricultural vegetation is carried out during the day in conditions of open sunlight, so the two main factors affecting the spectral composition of illumination are the angle of elevation of the sun and the shading caused by clouds.

To compensate for the effect of illumination on the spectral composition of the survey, two approaches were used to normalize the spectral composition of the GSI. Following the hypothesis of gray contours [17], normalization to the spectrum of the brightest part of the image can be performed at the preprocessing stage: $HSI_{LightNorm}^{(i)} = \frac{HSI^{(i)}}{MAX(HSI^{(i)})} \times 255$.

Also, compensation for the effects of lighting changes can be performed using an additional layer of batch normalization [17] in the neural network architecture. Within the framework of this work, both approaches are implemented, their effectiveness is investigated in the experimental section.

3.1. Neural network classification

In this paper, it is proposed to use a modification of the neural network classifier from [6]. This network is a combination of convolutional layers for the analysis of 1D and 2D slices of hyperspectral images, thereby allowing you to proceed to the analysis of 3D sections (Figure 5). A hypercube is fed to the network input, the dimensions of which are determined by the number of channels B and spatial

coordinates (height and width) H and W respectively. This approach allows not only to identify the features of one hyperpixel, but also to find its dependencies with neighboring ones.

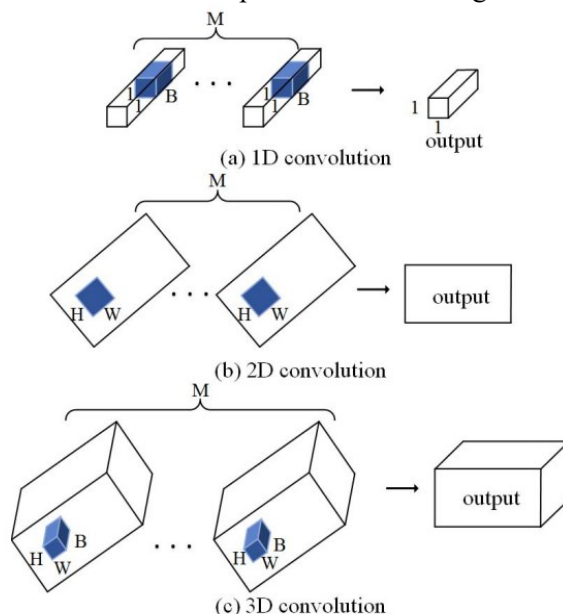


Figure 6. Transition from 1D and 2D analysis to 3D.

The network consists of 10 convolutional layers and one fully connected with a dropout (Table 2), the network depth is 5. The dropout layer avoids retraining, the exclusion coefficient in it is 0.6 [6]. To compensate for the effect of illumination, a batch normalization layer was added after each convolutional layer.

To form a training sample, sections from several hyperspectral images were used (Figure 7). The final size was 900×2730 hyperpixels, which is about a third of all data, after which 90% of pixels were taken into the training set and 10% into the validation set. This approach minimizes the difference in illumination when analyzing various hyperspectral images.

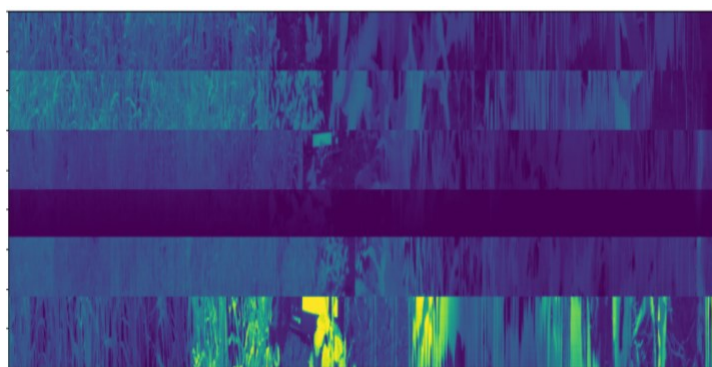


Figure 7. Example of a training set made up of slices of several hyperspectral images.

The training was conducted using the Adagrad optimizer. The learning rate parameter was 0.01, and the momentum parameter of the optimizer was set to 0.01. The size of the mini-batch was 40. The block size was $7 \times 7 \times 250$. The process of training the network took 50 epochs to complete stabilization.

Four different approaches to the classification problem were investigated: the classical architecture of M3D-DCNN [6] without preprocessing and with it, the modification of M3D-DCNN proposed in this paper without preprocessing and with it. (Figures 8–9).

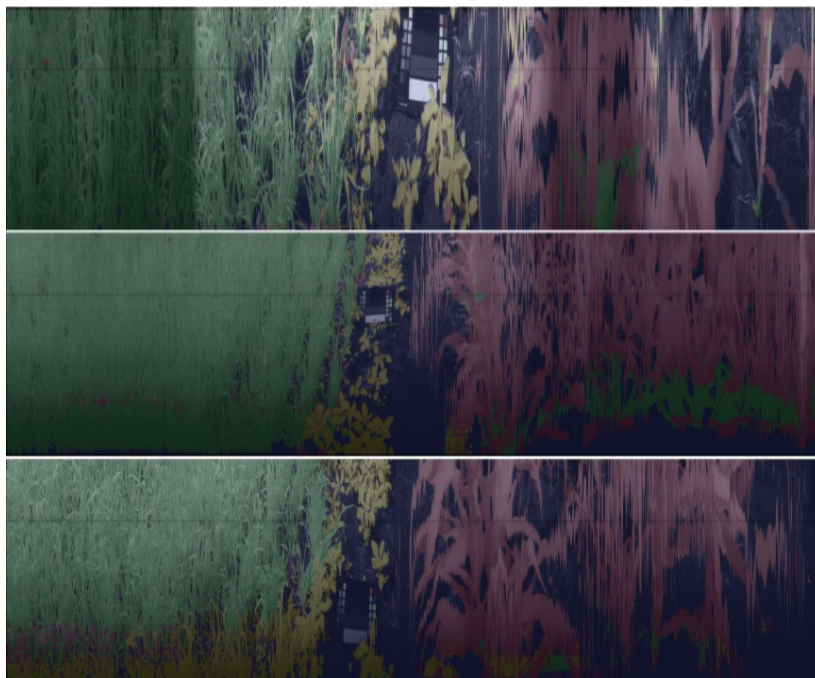


Figure 8. Classification with M3D-DCN architecture with preprocessing.

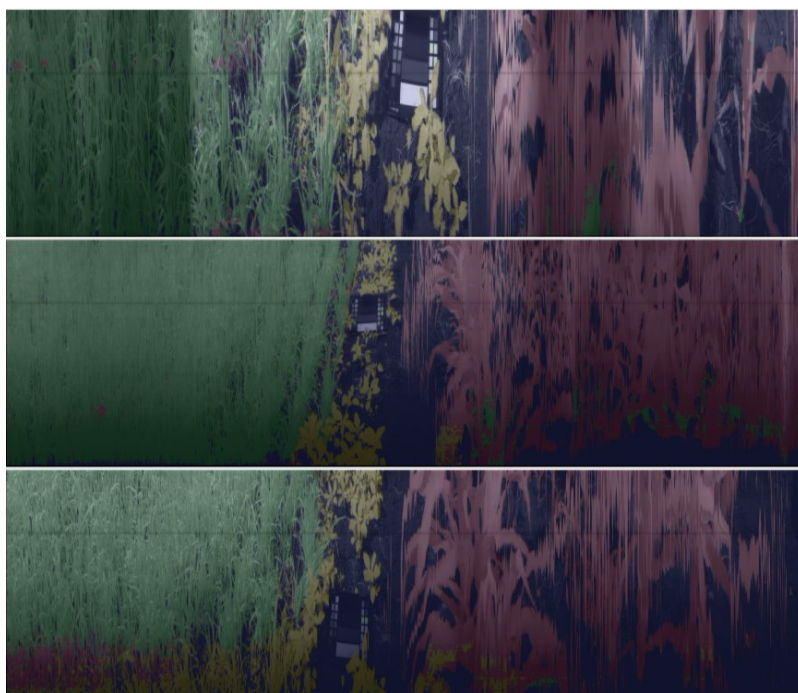


Figure 9. Classification of the modified M3D-DCNN architecture with preprocessing.

As you can see from the presented figures, the preprocessing of hyperspectral images does not significantly improve the quality of classification. Using batch normalization as illumination normalization reduces the size of areas of erroneous classification.

For stricter quality control, classification accuracy was calculated for three hyperspectral images for each proposed approach (Table 2).

Table 2. Accuracy of the classifier from [4] and its modifications for initial and preprocessed hyperspectral images.

	M3D-DCNN		Modified M3D-DCNN	
	–	Preprocessing	–	Preprocessing
HSI_1	95.3%	93.7%	95.7%	95%
HSI_2	94.4%	92.5%	95.1%	96%
HSI_3	91.4%	90%	91.5%	90.7%
Average	93.7%	92%	94.1%	93.9%

Based on the data obtained, it can be concluded that preprocessing in the form of normalization of hyperspectral images by illumination does not give an increase in classification accuracy. Using batch normalization for this purpose, on the contrary, gives the best results among those presented.

4. Conclusion

This paper presents an algorithm for classifying high-resolution hyperspectral images taking into account both spatial and spectral components. A convolutional neural network with an architecture based on [6], modified to better account for changes in the lighting of the scene, is used as a classifier. To prepare the training data, it is proposed to use vegetation indices that allow for primary binary segmentation of the hyperspectral image. The effectiveness of the proposed approach is shown. The resulting classification accuracy was more than 94%.

It is shown that the presented model has a classification accuracy higher than that of existing methods and allows using this algorithm for agricultural tasks.

Acknowledgments

The theoretical part of the work and the development of neural network models was carried out with the support of the Russian Science Foundation (RSF 20-69-47110).

References

- [1] Sharma V, Diba A, Tuytelaar T, and Gool LV 2016 *Technical report KUL/ESAT/PSI/1604, KU Leuven, ESAT, Leuven, Belgium*
- [2] Claudio H C, Cheng Y, Fuentes D A, Gamon J A, Luo H, Oechel W, and Sims D A 2006 *Remote Sensing of Environment* **103** (3) 304–311
- [3] Mahajan G R, Sahoo R N, Pandey R N, Gupta V K, and Kumar D 2014 *Precision agriculture* **15** (5) 499–522
- [4] Bing Liu, Xuchu Yu, Pengqiang Zhang, Xiong Tan, Anzhu Yu, Zhixiang Xue 2018 *Remote Sensing Letters* **8** 839–848
- [5] Bioucas-Dias J M, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi N, and Chanussot J 2013 *IEEE Geoscience and Remote Sensing Magazine* **1** 6–36
- [6] Mingyi He, Bo Li, Huahui Chen 2017 *IEEE International Conference on Image Processing (ICIP)* 3904–3908
- [7] Elmetwalli A H, El-Hendawy S, Al-Suhaibani N, Alotaibi M, Tahir M U, Mubushar M, Elsayed S 2020 *Sensors* **20** (22) 6569
- [8] Mahlein A K 2016 *Plant Disease* **100** (2) 241–251
- [9] AlSuwaidi A, Grieve B, Yin H 2018 *Measurement Science and Technology* **29** (10) 104001

- [10] Podlipnov V, Shchedrin V, Babichev A, Vasilyev S, and Blank V 2018 *Computer optics* **42** (5) 877–884
- [11] Nikonorov A, Petrov M, Bibikov S, Kutikova V, Yakimov P, Morozov A. 2018 *10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS) IEEE* 1–9
- [12] Karpeev S, Khonina S, Murdagulov A, Petrov M. 2016 *Vestnik of the Samara State Aerospace University* **15** (1) 197–206
- [13] Jung A, Kardeván P, and Tőkei L 2006 *Progress in agricultural engineering sciences* **2** (1) 95–117
- [14] Kwan C, Gribben D, Ayhan B, Li J, Bernabe S, Plaza A 2020 *Remote Sensing* **12** (23) 3880–3888
- [15] Hu W, Huang Y, Wei L, Zhang F, Li H 2015 *Journal of Sensors* **2015** 30–42
- [16] Manea D, Calin M A 2015 *The Imaging Science Journal*, **63** (4) 214–219
- [17] Van De Weijer J and Gevers T 2005 *IEEE International Conference on Image Processing* **2** 722