

# REMOTE SENSING AND DEEP LEARNING FOR SUSTAINABLE MINING

*Pedram Ghamisi, Hao Li, Robert Jackisch, Behnood Rasti, and Richard Gloaguen*

Helmholtz-Zentrum Dresden-Rossendorf (HZDR),  
Helmholtz Institute Freiberg for Resource Technology, Machine Learning Group, Germany  
GIScience Chair, Institute of Geography,  
Heidelberg University, Germany

## ABSTRACT

This paper brings together advances in remote sensing and deep learning for mineral mapping in a sustainable way. In more detail, we propose a multisensor feature fusion approach to integrate heterogeneous RGB, multispectral, and hyperspectral images for sustainable mining. The proposed approach is composed of two main steps; Feature extraction and classification. In the feature extraction step, we develop a three-stream convolutional neural network to extract high-level information from the input multisensor data. In the classification step, we develop a multisensor composite kernel approach to perform fusion and mapping simultaneously. The proposed approach produces very high quality classification maps with exceptional results in terms of classification accuracies.

**Index Terms**— Multisensor data fusion, deep learning, composite kernels, convolutional neural network, sustainability, mineral mapping.

## 1. INTRODUCTION

Data analysis based on non-invasive remote sensing and geophysical survey techniques has the potential to provide a significant contribution for a sustainable exploration and development of mineral resources. By maximizing the efficiency of such techniques, the whole value chain from exploration for primary resources to recovery can benefit, and partly contribute, e.g., to meet the United Nations development goals of sustainability and ecological viability [1]. In this example, operation and maintenance of drone-borne hyperspectral mapping is cost-effective and feasible, with high productivity given proper application and training. Notably, an advantage of light-weight multi-sensor systems is the applicability under challenging environmental conditions, since those are met for example in the polar and equatorial regions. While unmanned aerial systems (UAS) of now might not fully replace conventional airborne methods, high data amounts require the integration of advanced data analysis and machine learning techniques to optimize the results and harness the full potential of multi-sensor, multi-layered datasets. Moreover,

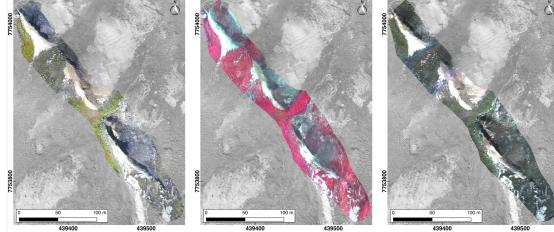
to create a sustainable development in the competitive mining sector, surveying and data analysis is applied repeatedly, since only a fraction of mineral discoveries is developed into a mine [2]. Finally, exploration is only one stage of the whole mineral development and processing chain. In that sense, the same technologies are valid to investigate secondary resources, for example those that are created along the resource production pipeline [3].

Advances in machine learning particularly in deep learning has shown successful performances for remote sensing classification [4]. Additionally, multisensor data fusion techniques provide tools to integrate data having different attributes and characteristics which has been shown to be also beneficial for different remote sensing applications including classification [5]. As a result, this paper brings together advances in deep learning and remote sensing for sustainable mining. In more detail, here, we develop a multisensor data fusion approach based on an innovative three-stream CNN and a multisensor composite kernel technique. The proposed approach effectively integrates three different data types including RGB, multispectral, and hyperspectral data with different spectral and spatial.

## 2. DATASET

Our acquired images were captured during the geologic field campaign under project MULSEDRO (Heincke et al., 2019) on the northern shores of Disko Island, West Greenland (69° 885' N, 52° 577' W). The target area features a part of the NW-SE striking Illukunnguaq dyke, a 5 m broad, paleocene lava intrusion through a cretaceous sedimentary formation. This basalt intrusion can be traced for roughly 800 m along the coast and is known for sulphide-iron-bearing mineralization, featuring e.g. pyrrhotite [6].

Unmanned aerial vehicles (UAV) were used to obtain the RGB and HSI scans. A fixed-wing UAV, the eBee plus, captured high-resolution GNSS-tagged orthophotos of the coast-line and the target with the 20 MP SODA camera. Multi-spectral data was taken with the Sequoia 4 band multispectral camera during a consecutive flight with the



**Fig. 1:** Input data - From left to right: RGB, MSI (R-735 nm, G-660 nm, and B-550nm), and HSI (R-639 nm, G-551nm, and B-503nm)

same fixed-wing over the same area.] Structure-from-motion stereo-photogrammetry within Agisoft Photoscan was used to create a precise, georeferenced orthomosaic with a 5 cm ground sampling distance (GSD) for the RGB, and 14 cm for the multispectral mosaic.

To capture the HSI image mosaic, the Senop Rikola HS Imager, featuring 0.6 MP and 50 bands in flight-mode was used. The camera takes single image frames that require the UAV (Tholeg Octocopter Tholeg Tho-R-PX-8/12) to follow a pre-programmed flight path while performing a stop-scan-and-motion pattern. Image pre-processing including geometric and radiometric corrections were conducted in the MEPHYSTO Toolbox [7]. The final product is an orthorectified image mosaic measuring 350x50 m with a resampled GSD of 14 cm. Data was acquired under sunny and low-wind conditions. In general, the polar regions create challenging circumstances for drone-borne remote sensing, due to low sun angles, prominent shadows, atmospheric water vapor and cold climate. Fig. 1 shows the input RGB, multispectral, and hyperspectral data.

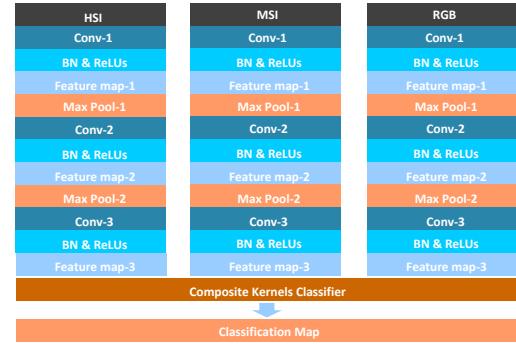
### 3. METHODOLOGY

Fig. 2 illustrates the general workflow of the proposed method. As can be seen, this method consists of two main parts: (1) Three-stream convolutional neural network to extract informative features from heterogeneous multisensor RGB, multispectral, and hyperspectral images and (2) multisensor composite kernels (MCKs) to classify the extracted features. We describe these two steps in more detail in the following subsections.

#### 3.1. Multisensor Convolutional Neural Network

In general, CNN architectures contain three parts: convolutional layers, pooling layers, and nonlinear transformations [8]. Convolutional layers are defined as follows:

$$\mathbf{x}_i^l = f \left( \sum_{j=1}^P \mathbf{x}_j^{l-1} * k_{ij}^l + b_i^l \right), \quad (1)$$



**Fig. 2:** Flowchart of the proposed classifier

where  $P$  indexes the feature map numbers,  $\mathbf{x}_j^{l-1}$  is the  $j$ th feature map of the  $(l-1)$ th layer, and  $\mathbf{x}_i^l$  is the  $i$ th feature map of the current  $(l)$ th layer.  $k_{ij}^l$  refers to the weight of  $(l)$ th layer, which connects the  $i$ th and  $j$ th feature maps.  $b_i^l$  stands for the bias of the  $j$ th feature map in the  $i$ th layer. The function  $f$  is a nonlinear activation function, and  $*$  is the convolution operation.

A pooling layer can be placed after a convolution layer to extract invariant features and reduce the size of the feature maps. Here, we use a max pooling layer that combines a small  $n \times n$  patch of the corresponding convolution layer as follows:

$$\mathbf{x}_{max} = \max_{n \times n} (\mathbf{x}_i^{n \times n} u(n, n)), \quad (2)$$

where  $u(n, n)$  is a window function to the patch of the convolution layer and  $\mathbf{x}_{max}$  is the maximum in the neighborhood.

After applying a sequence of convolution and pooling layers, we can add a fully connected layer before the last softmax layer, where the input feature maps are further reshaped into a vector feature.

During the back-propagation, an optimization approach needs to be used to tune the trainable parameters (i.e., weights  $k_{ij}^l$  and biases  $b_i^l$ ). Here, we use a stochastic gradient descent algorithm to achieve a minimized loss. Once the CNN training is finished, the trained networks can further perform as feature extractors by cutting off the coefficients of the last softmax layer [9]. As the next step, we use a multisensor CK classifier (SVM) to classify features extracted by the three-stream CNN. For detailed explanation, see [10].

#### 3.2. Multisensor Composite Kernel (MCK)

The concept of MCK was firstly introduced in [11]. Let  $\mathbf{x}_i^{HSI}$ ,  $\mathbf{x}_i^{MSI}$  and  $\mathbf{x}_i^{RGB}$  be the output of CNN-based feature extraction approach applied to the hyperspectral, multispectral, and RGB images in their original feature spaces  $\mathcal{H}$ , respectively, which correspond to three nonlinear feature mapping functions  $\Phi_1(\cdot)$ ,  $\Phi_2(\cdot)$ , and  $\Phi_3(\cdot)$  into Hilbert space  $\mathcal{H}_1$ ,  $\mathcal{H}_2$ , and  $\mathcal{H}_3$ . Then, the following transformation can be obtained:

INPUT	[28*28*P]
CONV-1	Kernel size:5*5*P kernel:32 weights:(5*5*P)*32+32
Feature map-1	24*24*32
POOL-1	12*12*32
CONV-2	Kernel size:5*5*32 kernel:64 weights:(5*5*32)*64+64
Feature map-2	8*8*64
POOL-2	4*4*64
CONV-3	Kernel size:4*4*64 kernel:128 weights:(4*4*64)*128+128
Feature map-3	1*1*128
Full connections	Kernel size:1*1*128 weights:(1*1*128)*M+M
Feature map-3	1*1*M
Softmaxloss	
Output label	Probability vector:1*M

**Fig. 3:** Detailed information on CNN designs, which lists the feature map size of each layer. Kernels size, number, and weights are also explained in detail. P is the input feature map number and M is the number of classes.

$$\Phi(\mathbf{x}_i) = \{\Phi_1(\mathbf{x}^{\text{HSI}}), \Phi_2(\mathbf{x}^{\text{MSI}}), \Phi_3(\mathbf{x}^{\text{RGB}})\}. \quad (3)$$

A three-stream composite kernels could be then calculated with the dot product of  $\Phi(\mathbf{x}_i)$ :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = K_{\text{HSI}}(\mathbf{x}_i^{\text{HSI}}, \mathbf{x}_j^{\text{HSI}}) + K_{\text{MSI}}(\mathbf{x}_i^{\text{MSI}}, \mathbf{x}_j^{\text{MSI}}) + K_{\text{RGB}}(\mathbf{x}_i^{\text{RGB}}, \mathbf{x}_j^{\text{RGB}}). \quad (4)$$

Here, the CK is formulated as the sum of positive definite matrices [11] independent from the spectral and spatial. As for the dimensionality,  $\dim(\mathbf{x}_i^w) = N_w$ ,  $\dim(\mathbf{x}_i^s) = N_s$ , and  $\dim(\mathbf{x}_i^e) = N_e$ ,  $\dim(K) = \dim(K_w) = \dim(K_s) = \dim(K_e)$ .

## 4. EXPERIMENTAL RESULTS

### 4.1. Algorithm Setup

Three networks with exactly similar architecture have been trained to extract features from RGB, multispectral, and hyperspectral images. Fig. 3 gives detailed information about the investigated network.

For the composite kernel network, an SVM with three composite radial basis function kernels was used. The optimal hyperplane parameters  $C$  and  $\sigma$  were tuned in the range of  $C = -4, -3, -2, \dots, 1$  and  $\sigma = -12, -11, \dots, 3$  using the fivefold cross-validation algorithm.

### 4.2. Discussions

Table 2 provides information about the obtained classification accuracies on the input data. In the table, overall accuracy (OA), average accuracy (AA), Kappa coefficient (K), and class accuracies are used to compare the performance of the proposed method against other studied techniques. **RGB**,

**MSI**, **HSI**, and **PC<sub>HSI</sub>** refer to a single CNN applied to RGB, multispectral image, hyperspectral image, and the first 6 principal components (due to the availability of 6 classes of interest) of the hyperspectral image, respectively. **DeepMCK** refers to the results obtained by the proposed method. Fig. 4 shows the corresponding classification maps.

As can be seen, **DeepMCK** provides the best classification accuracy compared to other studied approaches. In this context, **DeepMCK** improves **RGB**, **MSI**, **HSI**, and **PC<sub>HSI</sub>** by 37%, 15%, 31%, and 14%, respectively, in terms of overall accuracy. Additionally, for Basalt and Sandstone-Basalt **DeepMCK** provides 100% accuracy. In other words, those classes can be detected successfully using **DeepMCK**. Moreover, Debris and Vegetation are also classified successfully with over 91% and 94% accuracy, respectively. We should note that the accuracy obtained for the Sulphide class is not satisfactory (47%), however, the improvement obtained by **DeepMCK** is substantial compared to the other method used. This significant improvement is due to the fact that the proposed approach makes the most of multisensor datasets with complementary characteristics. In this context, CNN can model nonlinear and complex features effectively, which can improve the performance of the subsequent classifier. Furthermore, based on the kernel concept, MCK is able to implicitly fuse diverse types of data in a high-dimensional feature space, which mainly addresses the issue of high dimensionality with the availability of a limited number of training samples.

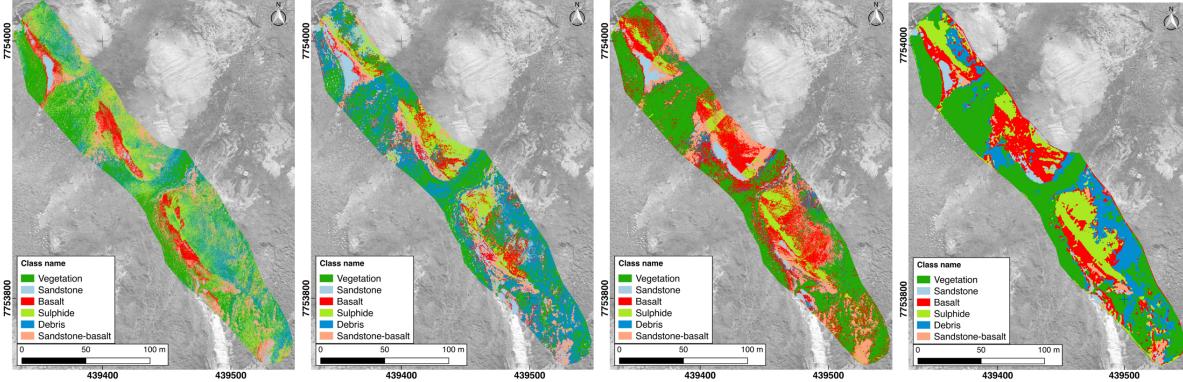
By comparing the classification maps shown in Fig. 4, it is evident that the proposed fusion framework can produce more precise and homogeneous classes, which is of great importance in accurate classification, particularly in complex mining scenes.

**Table 1:** Numbers of training and test samples

Classes	Name	Train	Test
1	Vegetation	55	104
2	Sandstone	104	62
3	Basalt	72	46
4	Sulphide	105	157
5	Debris	56	68
6	Sandstone-basalt	91	40
Total		483	477

## 5. CONCLUSIONS AND REMARKS

In this paper, we proposed a multisensor data fusion approach by integrating three different types of data including RGB, multispectral, and hyperspectral images with different spectral and spatial resolutions. The proposed fusion approach is based on an innovative three-stream CNN and a multisensor composite kernel technique. The experimental results confirmed that the proposed fusion approach can accurately classify different minerals in a sustainable way. The proposed



**Fig. 4:** Classification maps - From left to right: RGB, MSI, HSI, and the proposed method

**Table 2:** Classification accuracies for per class (in %), OA, AA, kappa coefficient (is of no unit), and processing time.

Class	HSI	PC <sub>HSI</sub>	RGB	MSI	DeepMCK
Vegetation	68.27	92.31	58.65	82.69	94.23
Sandstone	100.00	91.94	70.97	87.10	87.10
Basalt	86.96	91.30	4.35	36.96	100.00
Sulphide	4.46	39.49	31.21	38.72	47.77
Debris	14.71	27.94	30.88	72.06	91.18
Sandstone-Basalt	87.50	82.50	50.00	92.50	100.00
Overall (%)	47.17	64.78	41.30	63.52	78.62
Average (%)	51.70	60.78	35.15	58.50	74.32
Kappa	0.3934	0.3421	0.5814	0.5528	0.7422
Processing Time (s)	44.3214	9.4506	21.5658	20.6137	35.5986

approach can be generalized to other rasterized data for classification and fusion.

## 6. REFERENCES

- [1] S.H. Ali, D. Giurco, N. Arndt, E. Nickless, G. Brown, A. Demetriadis, R. Durrheim, M.A. Enriquez, J. Kinnaid, A. Littleboy, and et al., “Mineral supply for sustainable development requires resource governance,” *Nature*, vol. 543, pp. 367–372, 03 2017.
- [2] J. Prno and D. S. Slocombe, “Exploring the origins of ‘social license to operate’ in the mining sector: Perspectives from governance and sustainability theories,” *Resources Policy*, vol. 37, no. 3, pp. 346 – 357, 2012.
- [3] J. S. Salazar and L. Tavares, “Sustainability in the minerals industry: Seeking a consensus on its meaning,” *Sustainability*, vol. 10, 05 2018.
- [4] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson, “Feature extraction for hyperspectral imagery: The evolution from shallow to deep (overview and toolbox),” *IEEE Geoscience and Remote Sensing Magazine*, pp. 0–0, 2020.
- [5] P. Ghamisi, B. Rasti, N. Yokoya, Q. Wang, B. Hofle, L. Bruzzone, F. Bovolo, M. Chi, K. Anders, R. Gloaguen, P. M. Atkinson, and J. A. Benediktsson, “Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 1, pp. 6–39, March 2019.
- [6] A. Pedersen, L. M. Larsen, and G. K. Pedersen, “Lithostratigraphy, geology and geochemistry of the volcanic rocks of the maligat formation and associated intrusions on disko and nuussuaq, paleocene of west greenland,” *Geological Survey of Denmark and Greenland Bulletin*, vol. 40, pp. 1–239, 2018.
- [7] S. Jakob, R. Zimmermann, and R. Gloaguen, “The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: MephistoÀä toolbox for pre-processing drone-borne hyperspectral data,” *Remote Sensing*, vol. 9, no. 1, 2017.
- [8] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, “Deep feature extraction and classification of hyperspectral images based on convolutional neural networks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232–6251, Oct 2016.
- [9] Y. Chen, C. Li H. Jiang, X. Jia, and P. Ghamisi, “Deep feature extraction and classification of hyperspectral images based on convolutional neural networks,” *IEEE Trans. Geos. Remote Sens.*, vol. 54, no. 10, pp. 6232–6250, Oct 2016.
- [10] H. Li, P. Ghamisi, U. Soergel, and X. X. Zhu, “Hyperspectral and lidar fusion using deep three-stream convolutional neural networks,” *Remote Sensing*, vol. 10, no. 10, 2018.
- [11] P. Ghamisi, B. Rasti, and J. A. Benediktsson, “Multi-sensor composite kernels based on extreme learning machines,” *IEEE Geos. Remote Sens. Let.*, vol. 16, no. 2, pp. 196–200, Feb 2019.