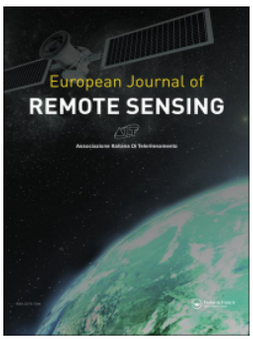

European Journal of Remote Sensing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20

Deep convolutional neural networks for surface coal mines determination from sentinel-2 images

L. Madhuanand, P. Sadavarte, A.J.H. Visschedijk, H.A.C. Denier Van Der Gon, I. Aben & F.B. Osei

To cite this article: L. Madhuanand, P. Sadavarte, A.J.H. Visschedijk, H.A.C. Denier Van Der Gon, I. Aben & F.B. Osei (2021) Deep convolutional neural networks for surface coal mines determination from sentinel-2 images, European Journal of Remote Sensing, 54:1, 296-309, DOI: 10.1080/22797254.2021.1920341

To link to this article: https://doi.org/10.1080/22797254.2021.1920341

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 10 May 2021.

Submit your article to this journal

Article views: 33

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tejr20

EUROPEAN JOURNAL OF REMOTE SENSING
2021, VOL. 54, NO. 1, 296–309
https://doi.org/10.1080/22797254.2021.1920341

Taylor & Francis
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

Deep convolutional neural networks for surface coal mines determination from sentinel-2 images

L. Madhuanand[a,b], P. Sadavarte[b,c], A.J.H. Visschedijk[b], H.A.C. Denier Van Der Gon[b], I. Aben[c] and F.B. Osei[a]

[a]Dept. Of Earth Observation Science, Faculty of Geo-Information Science and Earth Observation(ITC), University of Twente, Enschede, The Netherlands; [b]TNO, Dept. Climate, Air and Sustainability, Utrecht, The Netherlands; [c]SRON Netherlands Institute for Space Research, Utrecht,The Netherlands

ABSTRACT

Coal is a principal source of energy and the combustion of coal supplies around one-third of the global electricity generation. Coal mines are also an important source of CH_4 emissions, the second most important greenhouse gas. Monitoring CH_4 emissions caused by coal mining using earth observation will require the exact location of coal mines. This paper aims to determine surface coal mines from satellite images through deep learning techniques by treating them as a land use/land cover classification task. This is achieved using Convolutional Neural Networks (CNN) that has proven to be capable of complex land use/land cover classification tasks. With a list of known coal mine locations from various countries, a training dataset of "Coal Mine" and "No Coal Mine" image patches is prepared using Sentinel-2 satellite images with 13 spectral bands. Various pre-trained CNN network architectures (VGG, ResNet, DenseNet) are trained and validated with our prepared coal mine dataset of 3500 "Coal Mine" and 3000 "No Coal Mine" image patches. After several experiments with the VGG network combined with transfer learning is found to be an optimal model for this task. Classification accuracy of 98% has been achieved for the validation dataset of the pre-trained VGG architecture. The model produces more than 95% overall accuracy when tested on unseen satellite images from different countries outside the training dataset and evaluated against visual classification.

ARTICLE HISTORY
Received 26 October 2020
Revised 13 April 2021
Accepted 19 April 2021

KEYWORDS
Coal mine; supervised classification; deep learning; vgg; sentinel-2

INTRODUCTION

Methane (CH_4) is the second most important anthropogenic greenhouse gas after carbon dioxide, with a Global Warming Potential (GWP) of 84 compared to CO_2 over a 20-year time horizon (Myhre et al., 2013). Due to this high warming potential on a short timescale, reducing methane emissions is considered an essential and relatively easily achievable task to mitigate climate change in the near future. Consistent monitoring of anthropogenic methane emissions can aid in the understanding of the human-induced impact on atmospheric composition and subsequently the climate. The sources of methane include wetland rice fields, waste management, fossil fuel production and distribution, enteric fermentation and manure management from livestock. Among these, coal mines are one of the largest sources of methane emissions. The global emission inventory EDGAR v4.3.2 estimates fugitive CH_4 emissions of coal mines at ~13% of the global anthropogenic methane budget and as per U.S.Epa (2019) coal mines contribute to 11% of global methane emissions. Coal Mine Methane (CMM) is a common term used to refer to methane which is trapped in the coal bed (Karacan et al., 2011). In the case of underground mines, CMM is often actively vented to the atmosphere for safety reasons to avoid explosion danger. Even after a mine is abandoned the methane release from the coal mine may continue for a long period if unabated (Kholod et al., 2019). Most scholars suggest that the contribution of fossil fuel-related methane emissions is highly underestimated (Kholod et al., 2019). For coal mines, the emission estimates and associated mitigation potential vary based on mine type, coal type and mining practices from country to country.

Estimating the contribution of methane emissions from various sources and regions through different approaches is an important step towards mitigation of anthropogenic CH_4 emissions. The GALES (GAs LEaks from Space) project is aimed at the detection and quantification of localized CH_4 sources using the measurements from the Sentinal-5P satellite and its Tropospheric Monitoring Instrument (TROPOMI) at a global scale. The GALES project is led by SRON Netherlands Institute for Space Research in collaboration with TNO (Netherlands Organisation for Applied Scientific Research) and Free University (VU) to estimate CH_4 emissions from the fossil-fuel industry at a global scale. The TROPOMI onboard Sentinal-5P satellite is a state-of-the-art spectrometer that is

CONTACT L. Madhuanand l.madhuanand@uu.nl Department of Physical Geography, Utrecht University.Vening Meineszgebouw A, Princetonlaan 8a, Room no-4.42, 3584 CB Utrecht

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

launched by the European Space Agency (ESA) in October 2017. The TROPOMI instrument provides daily global coverage of atmospheric concentrations of CH_4 and other pollutants (CO and NO_2) at a ~ $7 \times 7\,km^2$ on-ground resolution. It was recently shown that large sources of methane due to oil and gas production can be identified and quantified using TROPOMI (Pandey et al., 2019; Varon et al., 2019). To accurately quantify the emissions from coal mines using TROPOMI, the exact locations of coal mines at a global scale are required. There is currently no open-source global dataset for the location of coal mines and databases at a national scale may have many missing locations and are difficult to access.

The main aim of this study is to identify surface coal mine locations at a global scale using remote sensing techniques. The underlying objective is that the obtained results can subsequently be used to identify potential coal mine methane emissions from TROPOMI CH_4 observations. This will reduce the large uncertainty of methane emissions from coal mines and can also be used to identify the location of mines with high methane emissions for future methane emission reduction strategies.

CLASSIFICATION

For this, the location of coal mines needs to be determined at a larger scale. Therefore, the identification of coal mines is approached as an image classification problem which is handled through remote sensing techniques. Image classification is a process of giving labels for each pixel based on their corresponding land use/land cover themes (Lillesand et al. (1994)). Different techniques are developed and widely used for identifying land use/land cover features. The classification can be per pixel or sub-pixel based, supervised or unsupervised, parametric or non-parametric, hard or soft classification (Mather and Tso (2009)). Supervised techniques require information on the land use/land cover labels to be used for the classification task while unsupervised approaches classify the images by clustering similar spectral classes together (Krizhevsky et al. (2017); Al-doski et al. (2013)). Classification algorithms follow either of the two approaches and vary based on the data availability or the features to be identified. One of the earliest and commonly used clustering algorithm is based on Gonzalez (1985)'s K-means clustering which groups pixels into clusters based on the nearest neighbourhood in an unsupervised way and the centres of the clusters are provided with a land use/land cover label. Many classification algorithms have proven successful in image classification tasks since then, like Support Vector Machines (SVM) (Richards (2013); Keuchel et al. (2003)), Decision Trees (Chen et al., (2002); Song and Lu (2015)), Object-Oriented Classification (Gamanya et al., 2009), Artificial Neural Networks (Sehgal, 2012).

Different approaches have been used in the past for the classification of minerals and geological features specifically. Gasmi et al. (2016) used principal component analysis with a combination of SVM for geological mapping of mineral deposits from Advanced Space Borne Thermal Emission and Reflection Radiometer (ASTER) data. Meanwhile, Hede Hawu et al. (2017) used various spectral indices to separate vegetation and mineral by splitting them based on spectral information. Tang and Spikes (2017) use neural networks to identify shale by training their network with high-resolution images from a Scanning Electron Microscope (SEM). Sun et al. (2017) compared techniques like support vector machines, artificial neural networks and random forests for mineral mapping, particularly copper ore.

Specific studies of coal mine mapping are limited. Zeng et al. (2017) used Landsat-8 images to identify surface coal mines over a region in China. They used object-oriented decision trees as a classification algorithm which is a combination of decision trees and object-oriented classification techniques to identify the surface coal mines. The satellite images are segmented into smaller patches to form objects and three indices, Normalised Difference Coal Index (NDCI), Normalised Difference Vegetation Index (NDVI) and Built-up Area Index (BAI). These indices are applied to segregate the satellite images into clusters satisfying the given condition. This approach works on a trial and error basis to fix the appropriate spectral values of various indices. Similarly, Lobo et al. (2018) approached this detection problem of surface coal mines using classification and regression trees in Amazon forest regions. They suggested the use of satellite images with a medium spatial and spectral resolution for this purpose as the use of high-resolution images might increase the complexity of image processing techniques. Their method involved the use of Sentinel-2 images (10 m resolution) for pixel-based classification by applying a collection of decision trees. This is also implemented only in a particular region and the selected properties are based on the study area. On the other hand, L. Chen et al. (2019) approached this task using hyperspectral information which includes Hyperion data and a geological perspective by identifying its composition. In summary, such studies require spectral information and have the inherent limitation of being region-specific. Hence, these approaches may not be suitable for a global study.

With the developments in neural networks and deep learning techniques, deep Convolutional Neural Networks (CNN) have proved to be successful in different land use and land cover classification

tasks. Deep learning techniques have been widely applied and have proven performance in identifying fine features from satellite images in both supervised and unsupervised ways (Basu et al. (2019), Kroupi et al. (2019)), and Helber et al. (2019) used a patch-based classification technique with CNN by creating a EuroSAT dataset of 10 land use/land cover classes with Sentinel-2 images. The dataset consisted of around 2000 ~ 3000 images for each class covering around 34 European countries with a combination of 13 multi-spectral bands and RGB bands. This study proved that deep learning models can be used to identify a wide variety of features from satellite images. Sumbul et al. (2019) proposed the BigEarthNet dataset with multiple labels within a single image using the Sentinel-2 bands. This is a huge dataset with more than 20 classes and has a class which includes all the mineral extraction sites but does not specifically concentrate on coal mines. The identification of coal mines, which is a composition of complex land use/land cover features, requires similar high-level classification algorithms for extracting those features, making deep learning the preferred approach. There has been very limited or no study to determine the location of coal mines at a global scale from Sentinel-2 images using deep learning methods. This makes our study an important contribution towards evaluating the deep learning models for identifying features similar to coal mines.

The overall objective of our study is to identify locations of coal mines from satellite images using deep learning techniques. In this study, a supervised approach where the CNN network is trained with image samples that are labelled (classes) either "Coal Mines" or "No Coal Mines" is implemented in this study. The paper is structured as follows. Section 2 describes coal mine features, training dataset preparation and pre-processing. Section 3 explains the methodology followed and the deep learning network used. The results along with the different experiments performed are detailed in section 4 while section 5 summarizes the conclusions and provides suggestions for further research.

DATASET

This section explains the different types of coal mines and their composition along with their features. The pre-processing steps undertaken to prepare a training dataset for deep learning models to learn from are also explained.

COAL MINES

Coal is an important fossil fuel. It is a principal source of energy and supplies around one-third of the global electricity generation (Ramani & Evans, 2020). Therefore, coal mines are an important source for both local and global economies. Based on the extraction method coal mines can be classified as either surface mine or underground mine. This study will focus on the identification of surface coal mines as underground mines have a different composition of land use/land cover features. It is expected that higher resolution images than used in this study are required for capturing these features.

Surface mines or opencast mines deploy surface-based mining techniques where the mineral lies close to the surface, making it easier for extraction at relatively low costs (Drahansky et al., 2016). They are mainly composed of extracting areas, stripping areas and dumping areas. The appearance of surface coal mines varies from one region to another as shown in Figure 1. This makes the identification of coal mines at a global scale challenging. A typical opencast mine may also include different land use/land cover features like a barren land, water bodies and industrial complexes along with distinct stripping patterns. This composition of different features mandates the use of higher spectral bands for identification and classification. Considering these challenges, the training dataset is prepared by using training images of coal mines from various regions to make the dataset as globally representative as possible.

TRAINING IMAGES

The training dataset comprises images of the land use/land cover classes that are to be classified. Any feature that is not part of the training images cannot be predicted by the model. Moreover, the dataset should be unbiased for the model to be able to generalise. For this particular problem, the training images are prepared in a similar way as that of the EuroSAT dataset (Helber et al., 2019). The freely available, medium resolution Sentinel-2 (L1C) data is used for the preparation of training images. The resolution of Sentinel-2 varies between 10 m to 60 m with 13 spectral bands in visible, near-infrared and short-wave infrared spectrum. It has a global coverage of 56°S to 84°N and a revisit period of 5 days with a 290 km field of view. The frequent revisit time, its open-source nature and its global coverage make it a preferred choice to build the database. To prepare the training dataset, geographic locations of 400 opencast coal mines are manually identified as is shown in Figure 2. The number of locations identified from each country is listed in Table 1. For these reference locations, Sentinel-2 tile images are downloaded from the USGS Earth Explorer. At each location, the most recent image in which mines are clearly visible with minimal/no cloud cover has been selected.

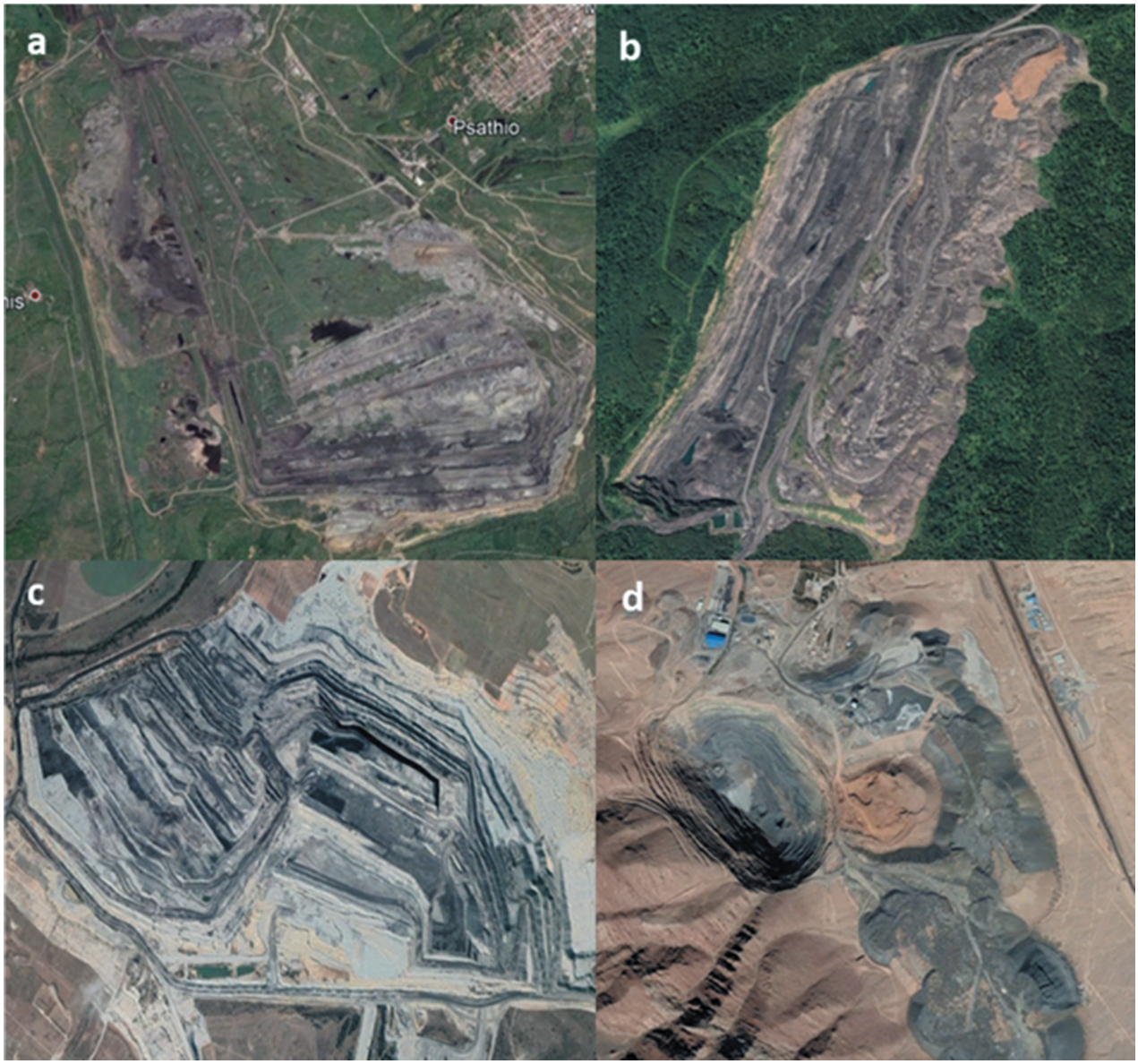

Figure 1. Opencast mines – a) Greece. b) Russia. c) Australia. d) Iran (Scale-1:20,000, Source – Google Earth,2018).

The downloaded images with 13 separate bands are stacked and resampled together using bilinear sampling to form a single layer of 10 m resolution. Image patches of size 64 × 64 pixels (640 m × 640 m) are extracted from the stacked image covering the opencast coal mine location in each tile. From the same tile, randomly selected patches of size 64 × 64 pixels that do not contain coal mines are also extracted. It should be noted that the area of most of the surface coal mines is much larger than 640 m × 640 m and hence multiple patches are required to cover a single coal mine. In this way, a total of 3500 and 3000 image patches are extracted to form the "Coal Mine" and "No Coal Mine" class, respectively, for the training dataset. Some examples from the coal mine and no coal mine training data are shown in Figure 3.

VALIDATION IMAGES

The dataset consisting of the two classes, Coal Mines and No Coal Mines, is split into a training and a validation dataset. For each class, 70% of the image patches are used for training and 30% is used as a validation set. The total number of training image patches is around 2450 for Coal Mines and 2100 for No Coal Mines. The number of validation image patches is 1050 for Coal Mines and 900 for No Coal Mines. During training, images from the training dataset are given as input to the CNN for learning, and the model evaluates its performance by predicting the class of images from the validation dataset. The classification accuracy is defined as the percentage of correct predictions by the model when applied to the validation set.

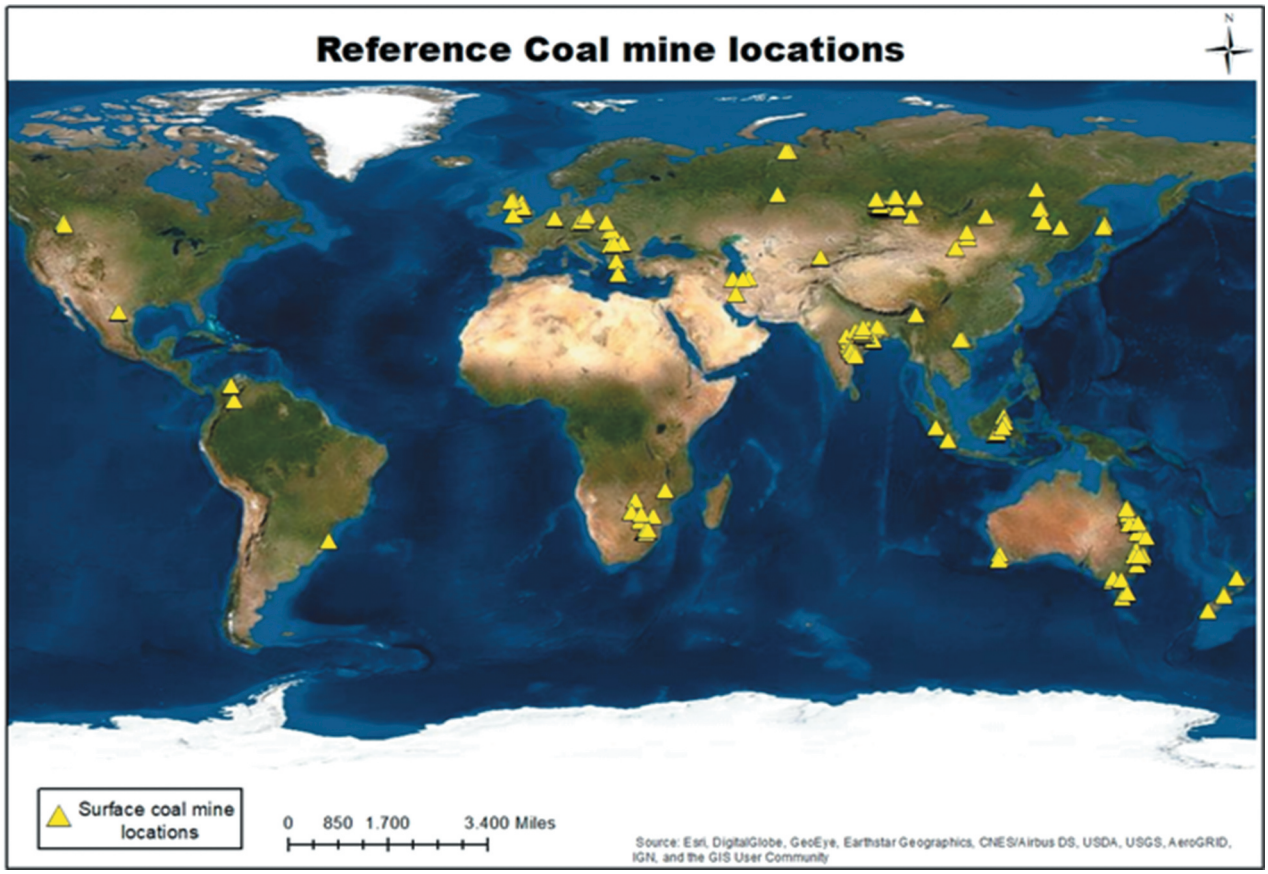

Figure 2. Manually identified reference coal mine locations.

Table 1. Number of coal mine locations identified manually from different regions.

Country	Number of coal mine points
India	62
Australia	72
Europe	34
South Africa	52
Russia	80
Other countries	100

TESTING IMAGES

The trained model is tested on a full or partial Sentinel-2 tile to understand its ability to identify coal mines on unseen tiles that did not contribute to the training dataset. Sentinel-2 tiles with coal mines are selected from different countries and classified with the trained model. Sentinel-2 tiles of size 40 × 40 km^2, one tile each from India, Australia, West Europe (Germany), East Europe (Poland) etc., are tested. The testing area consists of different land use/land cover features like barren land, waterbodies, built-up area, vegetation, agricultural fields, etc. The unseen tile is given as input to the trained model for classification of the entire image into two classes (Coal Mine or No Coal Mine). The classification results from the model are compared with visually obtained classification for estimation of accuracy.

METHODOLOGY

The methodology is shown in Figure 4. It is divided into three stages, dataset preparation, training the deep learning model and testing the model on different regions. The first step is to prepare a dataset for the intended classifying feature. In this case, the "Coal Mines" and "No Coal Mines" dataset is created as described in the previous section, for training the model. A Very Deep Convolutional Neural Networks (VGG) architecture is trained with the Coal Mine training dataset. The classification accuracy of the trained model is evaluated based on the model performance on the validation set. The trained model is then tested on unseen Sentinel-2 tiles with coal mines from different regions. The performance of the model on unseen tiles is evaluated through overall accuracy, user's accuracy, producer's accuracy and kappa by comparing model classification with visually identified coal mine classification (Richards, 2013). The following section explains the algorithm of deep learning model and VGG network architecture.

DEEP LEARNING MODELS

CNNs are advantageous for classification tasks as they consider the spatial context of an image. CNNs work through moving windows over images, having fewer

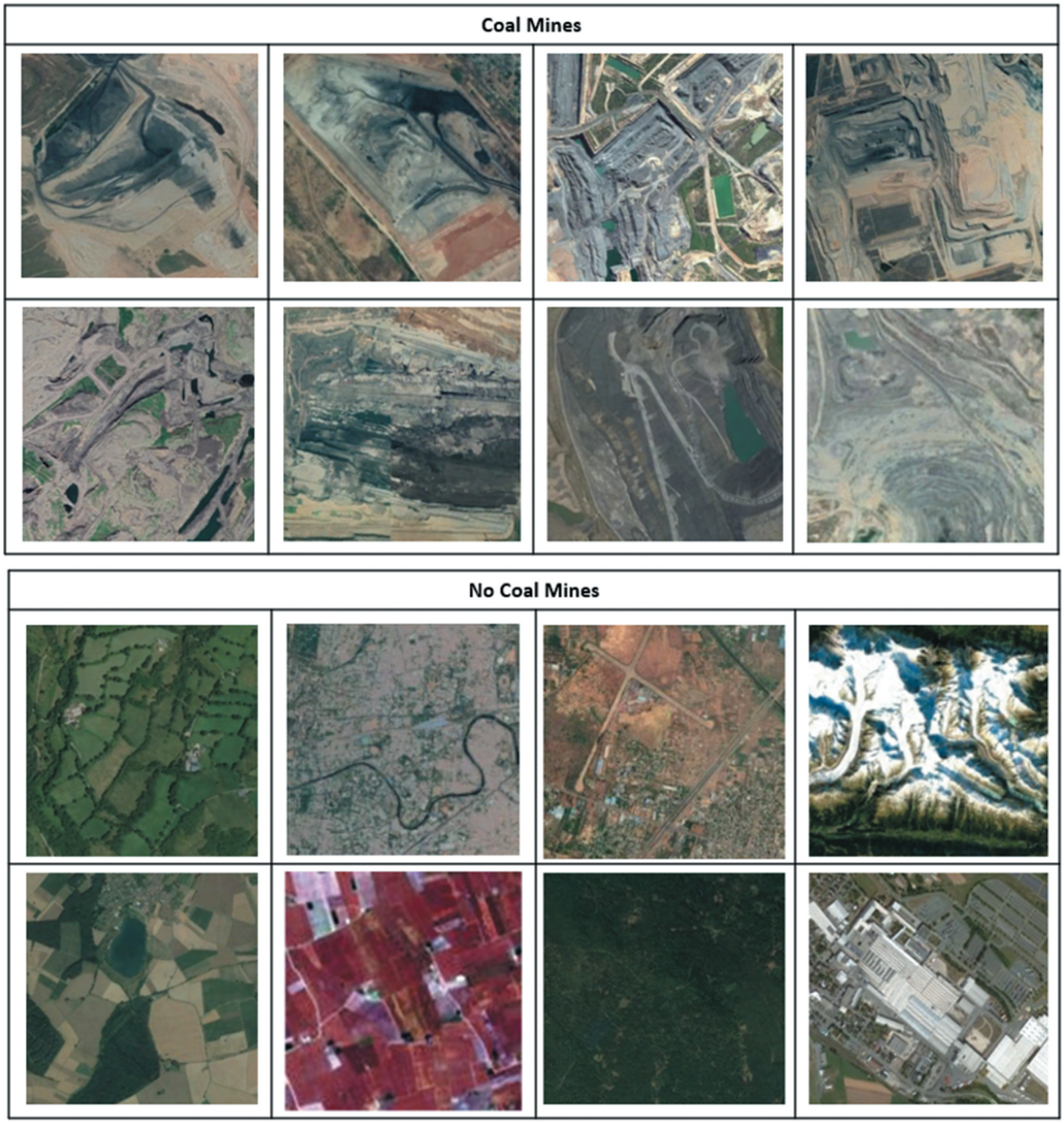

Figure 3. A) Coal Mines image patches b) No Coal Mines image patches.

parameters makes it easier for training (Krizhevsky et al., 2017). The filters are trained to learn to extract different features (like gradients, edges, etc.) at each layer by sliding over the images in small-sized windows. As the training image is passed through the deep learning model, the filters at each layer extract important information and pass it through the subsequent layers. The output from each convolutional layer is passed through a max-pooling layer which identifies the maximum value in each patch to form a downsampled image with dominant features. The process will be repeated with similar convolutional filters with various window sizes followed by a Rectified Linear Unit (ReLU) as activation function to convert into non-linear features and max-pooling layers for down-sampling. The whole image along with different band information is downsampled to form a single fully connected network of single pixels. This is passed through a softmax function layer which provides a label based on the land use/landcover layer learnt through training. The output label is compared with the reference label and the difference between them is used as a loss to improve the network in learning weights and bias. This forms the overall architecture of the convolutional deep learning classification models. There are different deep learning models which vary based on the architecture and the functions used for the learning process. This includes popular models like Very Deep Convolutional Networks by Visual Geometry Group (VGGNet) (Simonyan & Zisserman, 2015), Deep Residual Learning network (ResNet) (He et al., 2016) and

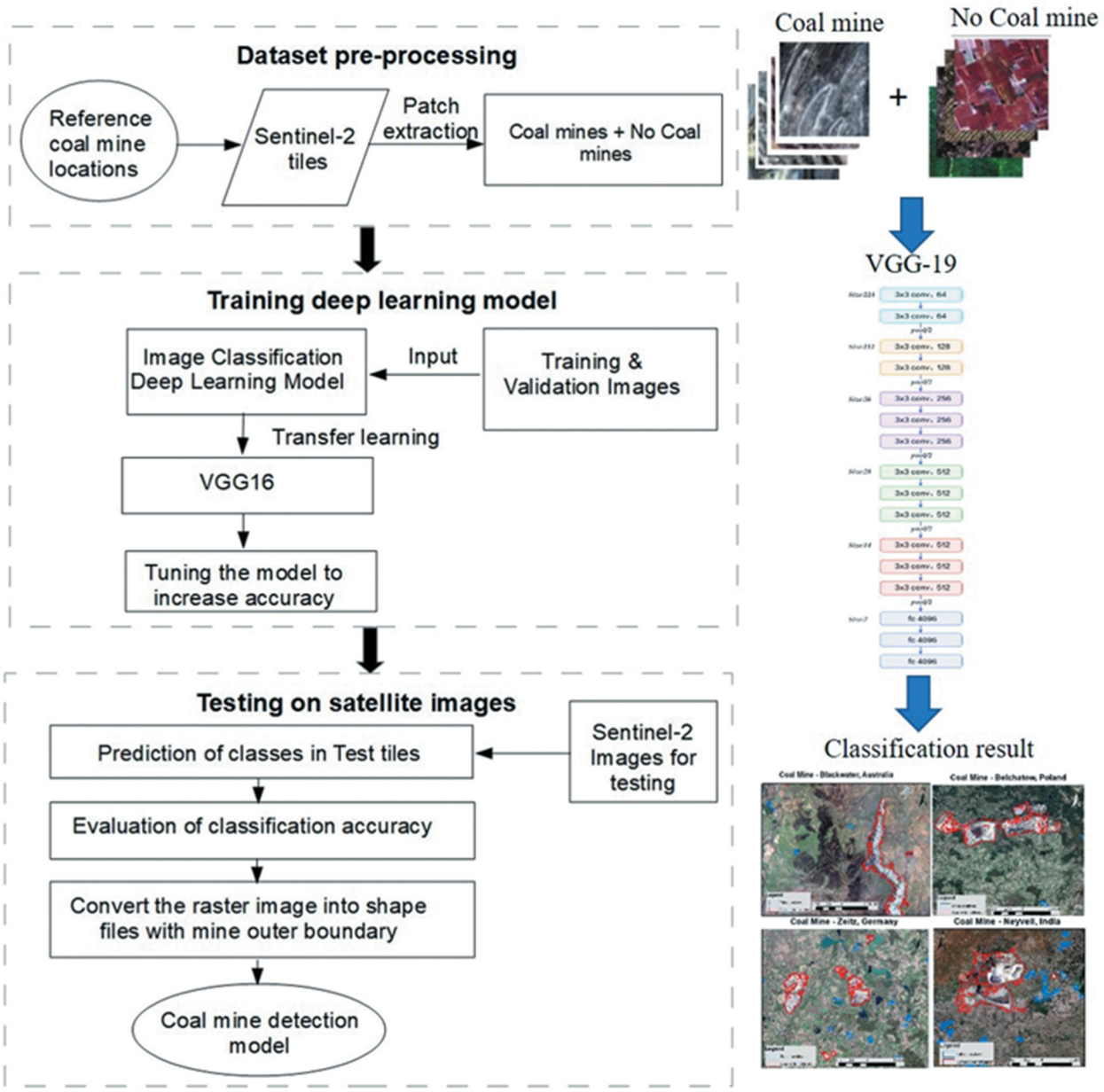

Figure 4. Overall Methodology.

Densely connected convolutional networks (DenseNet) (Huang et al., 2012), all of which have proved their performance in various classification tasks. This study uses a simple VGG architecture that reduces the training time by decreasing the number of learning parameters.

VGGNet

The VGG model can learn directly from scratch but this requires large amounts of training data for the model and much higher training duration. Since we have a limited number of images for training, transfer learning is preferred. Transfer learning is a commonly used method where the model is first trained with a very large dataset to learn the initial weights and bias. Here, the model is initially trained with ImageNet dataset and the weights learnt by the model at different convolutional layers are stored. The learnt weights are loaded for certain fixed number of layers only while excluding the input layer, the first convolutional layer and the final layers. In this study, the frozen weights and biases are initialised from the ImageNet dataset and the initial layer, along with the final classification output layer are trained with the coal mine dataset. This reduces computational time as it helps the model in learning the images faster due to pre-defined weights.

RESULTS AND DISCUSSIONS

EXPERIMENTAL ANALYSIS

Many different experiments are carried out with the model to determine the optimal approach that provides good accuracy. Initially, the problem was approached by adding the coal mine dataset as the 11th class in the EuroSAT dataset which already has

10 land use/land cover classes. The results from such an 11-class classification are noisy and deviate from the primary objective of the study, which is to identify coal mines only. To focus only on the identification of coal mines, it is decided to use just two classes "coal mines" and "no coal mines". Also, the size of the image patches is doubled to 128 × 128 pixels from the original 64 × 64 pixels to understand whether more spatial information can increase accuracy. However, this increased the computation time drastically and also puts very high demands on the system memory capacity. Due to these practical difficulties, the size of image patches is kept at 64 × 64 pixels. Experiments are also conducted to find the CNN network architecture with the best performance from three selected CNN models. Along with VGG, two other network architectures, ResNet (He et al., 2016) and DenseNet (Huang et al., 2012), are trained and tested on the same dataset. Both ResNet and DenseNet have a larger number of layers and a deeper architecture than VGG. These three models are chosen based on many other image classification studies and their performance in it. Whereas this led to increased computational time for ResNet and DenseNet, there is no noticeable improvement in accuracy to justify the increased computational costs. Hence VGG is preferred. To understand the role of spectral combinations, the model was trained with the entire 13 bands and then with selected bands with spectral cues for coal mine identification based on Zeng et al. (2017). The results show that although the 13-band combination requires more memory, the accuracy is significantly better than the 5-band combination and hence 13-bands are used. Based on these experiments, a two-class classification using 64 × 64 pixel image patches and 13-band combination trained with a VGG architecture is performed and the results are presented in the next section.

RESULTS

The results from the VGG deep learning model trained for identifying two classes, Coal Mine and No Coal Mine, are shown below. The model took around 6 ~ 7 hours for training with a system of NVIDIA Quadro GPU of 4 GB memory capacity. The VGG19 model showed a classification accuracy of 98.3% on our validation dataset of two classes with 13 spectral bands.

For testing the model performance on unseen (by the model) satellite images, Sentinel-2 tiles with coal mines are selected from different countries and classified with the trained model. Sentinel-2 tiles of size 40 × 40 km^2, one tile each from India, Australia, West Europe (Germany), East Europe (Poland) etc., are tested. The testing area consists of different land use/land cover features like barren land, waterbodies, built-up area, vegetation, agricultural fields, etc. In addition to testing on unseen images from countries that contributed to the training dataset, the model is also tested on an image from China where no image patches are included during training. The test image selected from China is similar to the one used by Zeng et al. (2017) to compare their results with our approach. This is done to explore whether the dataset and the model are applicable on a global scale or limited to the regions from which training image patches are chosen. To find the accuracy of the classified map, coal mines are identified visually and mapped manually using high-resolution images from Google Earth. Results from the model classification for various regions along with manually identified coal mines and comparison between the two are shown in Figure 6. The following accuracy metrics are used for measuring the accuracy of the classification. The overall accuracy gives us the proportion of pixels of coal mine sites accurately classified. The producer accuracy refers to the percentage of pixels from the ground truth that has been correctly classified and the user accuracy refers to the percentage of pixels identified as coal mines by the model that are actually coal mines. The kappa coefficient determines how well the classification has been performed in comparison with an arbitrary assigning of values. The confusion matrix along with overall accuracy, user accuracy, producer accuracy and kappa is given in Table 2.A-2.E.

The classification results from Figure 6 show the classification by the deep learning model and its comparison with visually identified coal mine classification. Moreover, the model is capable of identifying coal mines not just in regions from which training image patches are used (Australia, India, West Europe and East Europe) but also in regions on which the model was not trained (China). This

Table 2: a) Confusion matrix for accuracy assessment of Coal mine and No coal mine-Blackwater, Australia

		Reference (km2)		
Classified	**Blackwater, Australia**	**Mine**	**No Mine**	**Row Total**
	Mine	80.02	19.1	99.12
	No Mine	0.65	500.1	500.75
	Column Total	80.67	519.2	599.87
Overall accuracy=95.2%; Producer accuracy=99.1%; User Accuracy=80.7%; Kappa=0.81				

Table 2: b) Confusion matrix for accuracy assessment of Coal mine and No coal mine-Zeitz, Germany

		Reference (km2)		
Classified	**Blackwater, Australia**	**Mine**	**No Mine**	**Row Total**
	Mine	50.04	3.32	53.36
	No Mine	0.245	356.6	356.845
	Column Total	50.285	359.92	410.21
Overall accuracy=98.8%; Producer accuracy=99.4%; User Accuracy=93.7%; Kappa=0.96				

Table 2: c) Confusion matrix for accuracy assessment of Coal mine and No coal mine-Neyveli, India

		Reference (km2)		
Classified	Blackwater, Australia	Mine	No Mine	Row Total
	Mine	100.59	13.6	114.19
	No Mine	1.16	326	327.16
	Column Total	101.75	339.6	441.35
Overall accuracy=96.6%; Producer accuracy=98.4%; User Accuracy=88.1%; Kappa=0.91				

Table 2: d) Confusion matrix for accuracy assessment of Coal mine and No coal mine-Belachatow, Poland

		Reference (km2)		
Classified	Blackwater, Australia	Mine	No Mine	Row Total
	Mine	68.7	1.74	70.44
	No Mine	1.05	415.68	416.73
	Column Total	69.75	417.42	487.17
Overall accuracy=99.2%; Producer accuracy=98.4%; User Accuracy=97.5%; Kappa=0.97				

Table 2:e). Confusion matrix for accuracy assessment of Coal mine and No coal mine-Ordos, China

		Reference (km2)		
Classified	Blackwater, Australia	Mine	No Mine	Row Total
	Mine	94.8	11.73	106.53
	No Mine	0.885	332.6	333.485
	Column Total	95.685	344.315	440
Overall accuracy=97.1%; Producer accuracy=99%; User Accuracy=88.9%; Kappa=0.91				

shows that the training dataset is representative of the coal mines around the world and hence the model can be applied at a global scale. However, it is still possible that deep learning models may not be able to identify coal mines that are drastically different from those found in the training dataset.

It can be observed that there is much higher misclassification of No Coal Mine class as Coal Mine class (false positives) whereas there is little or no misclassification of a Coal Mine class as No Coal Mine class (false negatives). The false positives occur systemically in and around the external boundaries of the surface coal mine area. This may be due to the low resolution of the classification, which is based on image patches of size 64 × 64 pixels. Also, small areas of land that are spectrally similar to surface coal mines are misclassified as coal mines. The land use/land cover features which are mostly misclassified are barren lands and dried river beds. To try to reduce the false positives, the model is retrained by adding patches that are misclassified as coal mines to that of the no coal mine class dataset. While this improved the performance slightly, including more and more misclassified patches in no coal mine class led to overfitting of the model and adversely affected the accuracy. Fortunately, there are other ways to reduce the misclassification. For example, the false positives are further reduced by filtering out all coal mine classifications that are less than 64 × 64 pixels (640 m × 640 m) in size as it is unrealistic to have a surface coal mine to be of such a small size.

The accuracy of the model classification is evaluated by considering the visually identified coal mines as ground truth. All the tested images showed an overall accuracy of more than 95%, ranging from 95.2% for Blackwater, Australia to 99.2% for Belachatow, Poland. The producer accuracy varies from 98.4% to 99.4% while the user accuracy varies between 80.7% and 97.5%. The high values of producer accuracy and relatively low values of user accuracy bring out the fact that there is a predominance of false positives over false negatives in the model misclassification. It is possible to filter some of these out using additional constraints. On the other hand, if there are coal mines that are completely missed by the model then the model is not fulfilling its objective. Hence, the false positives are more tolerable in this case as they can be removed by imposing further constraints. The kappa for all tested images ranged from 0.81 to 0.97. The region shown in Figure 5(e) is classified for surface coal mine areas by Zeng et al. (2017) using an object-oriented decision tree approach based on three spectral indices whose results are comparable with that of the results produced by our deep learning model.

Comparison with a traditional maximum likelihood classifier

To compare our results with traditional supervised classification techniques, the classification task for the same test images is also performed using a maximum likelihood classification technique. For each test region, around 10 training samples for both Coal Mine and No Coal Mine are selected and used for maximum likelihood classification. To maintain consistency, the classified coal mine patches of size less than 64 × 64 pixels are removed from the maximum likelihood classifier also as is the case with the results of the deep learning model. Figure 7 shows the results from the two methods and Table 3 compares the overall accuracies for the tested images obtained from these methods.

It can be seen from Table 3 that the results from maximum likelihood have lower overall accuracy than the deep learning method for all the tested images. The deep learning model is better at distinguishing between coal mines and other spectrally similar land use/land cover features than the traditional maximum likelihood classifier approach. The results show that the deep learning model produces classification results that are closer to visually identified classification and is also able to generalise this with a representative dataset.

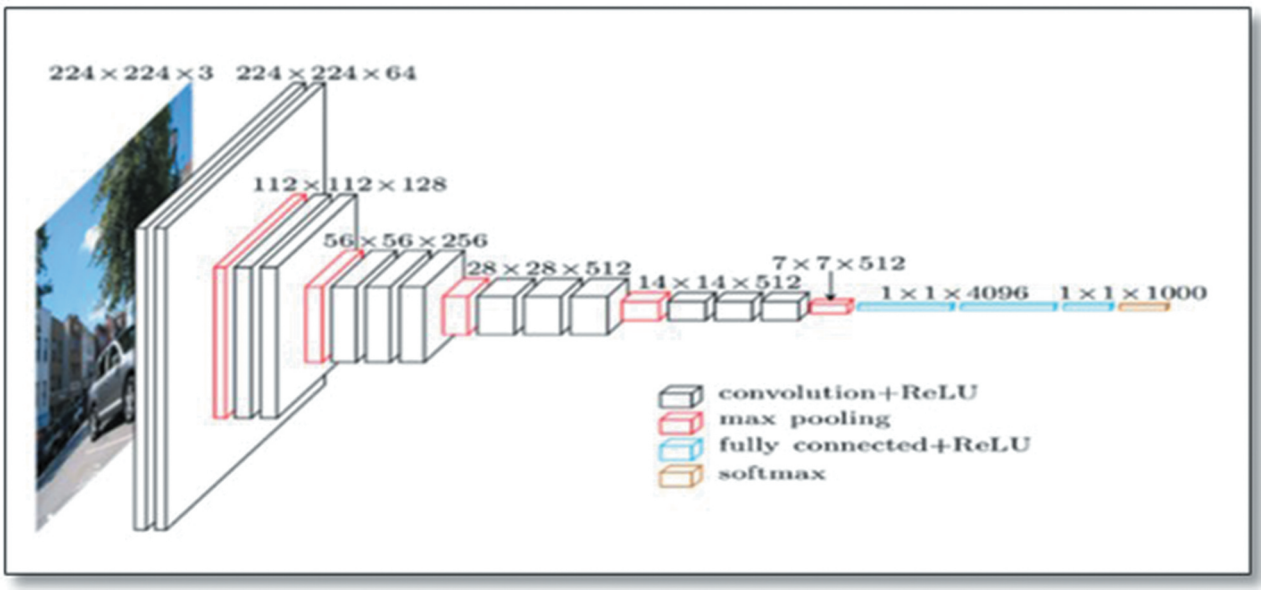

Figure 5. VGG architecture. Adapted from "Simonyan and Zisserman (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition 1–14".

Table 3. Classification accuracy of each region by deep learning model and maximum likelihood classifier.

Country	Deep learning model - Overall accuracy	Maximum likelihood classifier-Overall accuracy
Blackwater, Australia	95.2%	90%
Zeitz, Germany	98.8%	87%
Neyveli, India	96.6%	91%
Belchatow, Poland	99.2%	93%
Ordos, China	97.1%	89%

DISCUSSIONS AND CONCLUSIONS

The main aim of this study is to determine whether a deep learning model can be used for identifying surface coal mines from Sentinel-2 data. This study gives an outline about the dataset preparation for surface coal mine identification, CNN architectures suitable for our task and the evaluation over various regions along with a traditional classifier comparison.

In our study, we created a coal mine dataset with Sentinel-2 image patches of size 64 × 64 pixels from various regions are created to train deep learning models to identify coal mines. The resulting VGG-based deep learning model trained with the coal mine dataset can be applied over different regions for identification of surface coal mines. Adding misclassified patches into the training dataset and retraining and filtering out patches that are much smaller than a typical coal mine are done to increase the performance of the model. The model produced an overall accuracy of more than 95% for all the tested images. The model also produces better overall accuracy than traditional methods like the maximum likelihood classifier approach.

The model consistently gives higher producer accuracy than user accuracy, meaning that the misclassification is mostly of No Coal Mine class into Coal Mine class which are false positives. For our purpose, being the preparation of a dataset for future screening of high methane concentrations in TROPOMI images, false positives, although unwanted, are less of a problem than false negatives. The reason is that false positives simply add a limited number to the locations to be screened but will not lead to a less complete identification of methane hotspots. On the other hand, false negatives may cause an important surface coal mine to be missed and can lead to losing out some coal mine methane emissions.

It should be noted that other mineral mines like copper mines, which have similar spectral and textural features as coal mines, with a large open pit and industrial settings may possibly be misclassified. In order to overcome this, a pre-screening mechanism can be applied whereby the testing of Sentinel-2 tiles is limited to areas of geological formations that have the potential for coal mine extraction (Facts about coal and minerals, 2016) or, in other words, where coal layers are present in the subsurface.

This study contributes towards understanding the performance of deep learning models in identifying or classifying spectrally complex objects similar to coal mines at a global scale.

Outlook

We aim to apply our selected Coal mine detection model over a larger area, e.g., a continent or large country like India. The identified coal mines from our model will be used for identifying methane

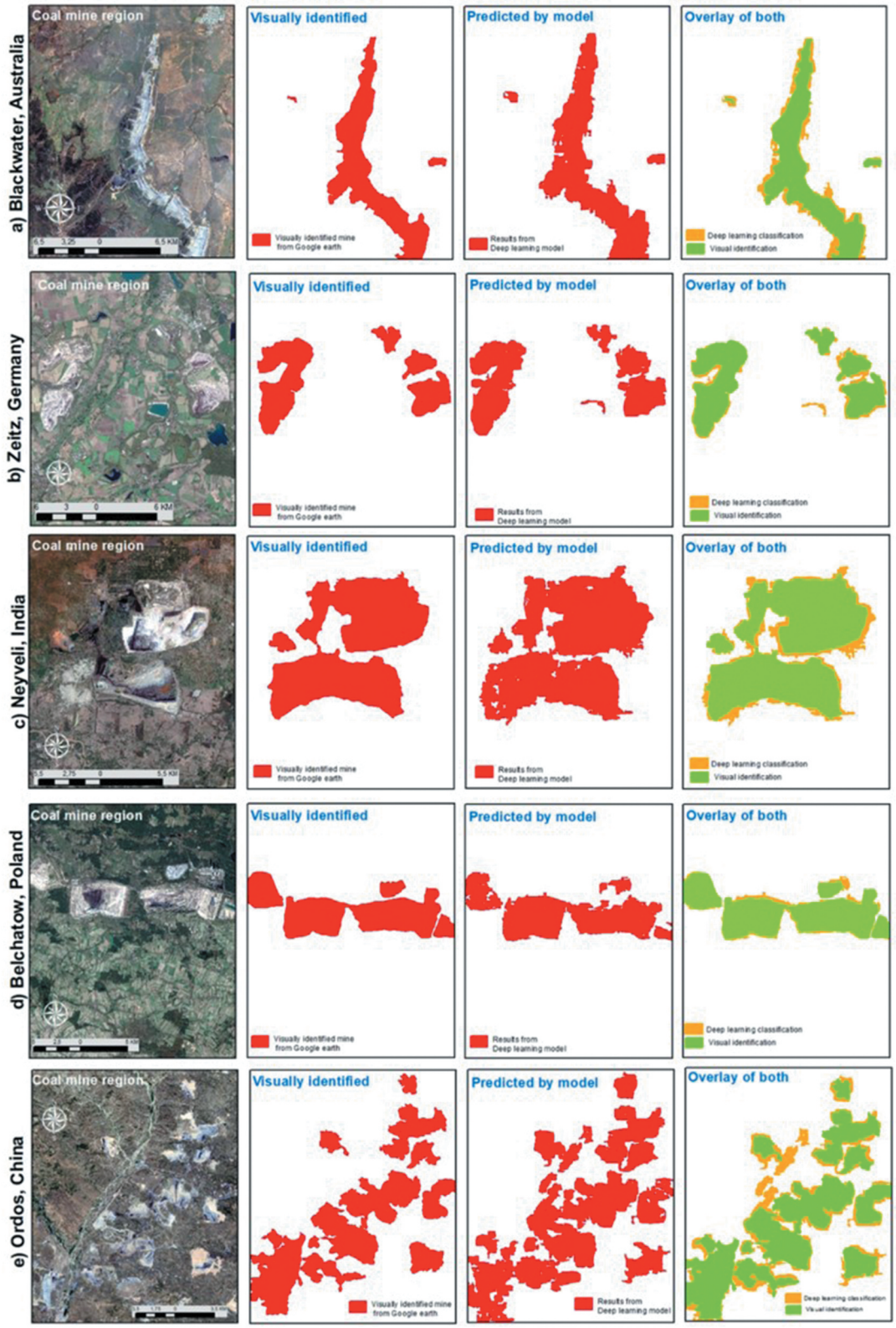

Figure 6. Classification results tested over different regions.

emissions from coal mines from space. A major question to be answered in this next step will be if the opencast mines contribute significantly to methane emission or if mainly underground mines are important emitters. Regardless, a robust and reliable identification of surface coal mines using satellite images will be relevant for other research applications such as land use change detection and development.

In preparation of the above, a preliminary further test is done where the model was applied on a larger scale for a part of India. Screening entire Sentinel-2 image tiles is computationally demanding. In order to

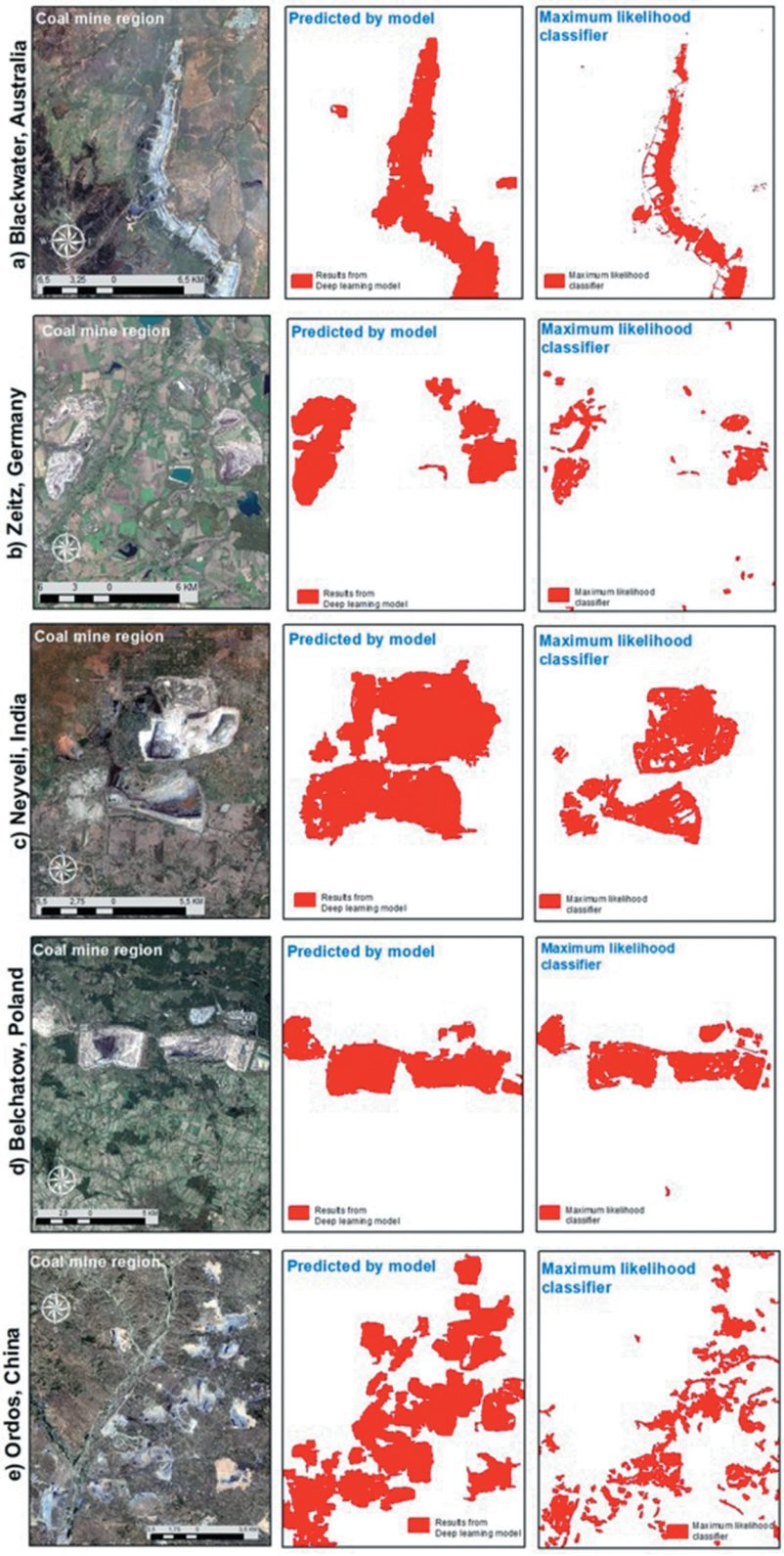

Figure 7. Different Classification results tested over various regions.

reduce the number of images to be tested, a pre-selection based on the available coal bed resource (Trippi & Tewalt, 2011) is applied which identified around 50 image tiles as potential coal mine sites in India. These tiles are downloaded and the model is then tested over the entire tile. This was done for three tiles. We compared the result with a test image over Germany and Poland and observed that the number of false positives in the temperate region is significantly less than that of a tile from India. Satellite images from India tend to have a lot of dry, barren lands which contribute to false positives due to similar spectral properties as that of coal mine stripping areas. To reduce such false positives in dry regions we are planning to make use of the fact that opencast coal mines generally have a steep slope. Using a Digital Elevation

Model (DEM), an additional constraint can be imposed which filters out areas that do not have a steep gradient. This constraint is decided based on the initial results from India, where the overwhelming majority of false positives are dry, barren, flatland. Similarly, while testing on a global scale, the approach can be modified based on the region and other region-specific constraints can be developed to reduce false positives.

Acknowledgments

This research is partly supported through the GALES (Gas Leaks from Space) project (Grant 15597) by the Dutch Technology Foundation, which is part of the Netherlands Organisation for Scientific Research (NWO) and is partly funded by the Ministry of Economic Affairs, The Netherlands. We thank Dr Janot Tokaya (TNO) for helping with the preliminary application over a larger part of India and discussion.

Data Availability

The data that support our results can be obtained from the corresponding author based on request. The codes used are available in GitHub – https://github.com/logsanand.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is partly supported through the GALES (Gas Leaks from Space) project (Grant 15597) by the Dutch Technology Foundation, which is part of the Netherlands Organisation for Scientific Research (NWO) and is partly funded by Ministry of Economic Affairs, The Netherland

ORCID

L. Madhuanand http://orcid.org/0000-0002-1534-5234

References

Al-doski, J., Mansor, S. B., Zulhaidi, H., & Shafri, M. (2013). Image Classification in Remote Sensing. Journal of Environment and Earth Science, 3(10), 141–148. Retrieved from https://core.ac.uk/download/pdf/234663192.pdf

Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., & Nemani, R. (2015). DeepSat - A learning framework for satellite imagery. In SIGSPATIAL '15: Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems (Vol. 37, pp. 1–10). https://doi.org/10.1145/2820783.2820816

Chen, L., Yang, X., & Zhen, G. (2019). Potential of Sentinel-2 data for alteration extraction in coal-bed methane reservoirs. *Ore Geology Reviews, 108*, 134–146. https://doi.org/10.1016/j.oregeorev.2017.10.009

Chen, P., Lu, X. X., Liew, S. C., & Kwoh, L. K. (2002). Quantification of land cover change and its impact on hydro-geomorphic processes in the upper yangtze using multi-temporal Landsat imagery: An example of the Minjiang area. In IEEE International Geoscience and Remote Sensing Symposium (Vol. 2, pp. 1216–1218). https://doi.org/10.1109/igarss.2002.1025893

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. In IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255). https://doi.org/10.1109/cvprw.2009.5206848

Drahansky, M., Paridah, M. ., Moradbak, A., Mohamed, A. ., Owolabi, F. abdulwahab taiwo, Asniza, M., & Abdul Khalid, S. H. . (2013). Reclamation of Degraded Landscapes due to Opencast Mining. In In Advances in Landscape Architecture (Vol. i, pp. 13). Intech. https://doi.org/http://dx.doi.org/10.5772/57353

Facts about coal and minerals. (2016). National Mining Association

Gamanya, R., De Maeyer, P., & De Dapper, M. (2009). Object-oriented change detection for the city of Harare, Zimbabwe. *Expert Syst. Appl*, *36*(1), 571–588. https://doi.org/10.1016/j.eswa.2007.09.067

Gasmi, A., Gomez, C., Zouari, H., Masse, A., & Ducrot, D. (2016). PCA and SVM as geo-computational methods for geological mapping in the southern of Tunisia, using ASTER remote sensing data set. *Arab. J. Geosci*, *9*(20), 1–12. https://doi.org/10.1007/s12517-016-2791-1

Gonzalez, T. F. (1985). Clustering to minimize the maximum intercluster distance. Theoretical Computer Science, 38, 293–306. https://doi.org/10.1016/0304–3975(85)90224–5

He, K., Zhang, X., Ren, S., & Sun, J. 2016. Deep residual learning for image recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit 770–778. https://doi.org/10.1109/CVPR.2016.90.2016-Decem

Hede Hawu, N. A., Koike, K., Kashiwaya, K., Sakurai, S., Yamada, R., & Singer, D. A. (2017). How can satellite imagery be used for mineral exploration in thick vegetation areas? *Geochemistry, Geophysics, Geosystems, 18*(2), 584–596. https://doi.org/10.2113/2.4.476

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*(7), 2217–2226. https://doi.org/10.1109/IGARSS.2018.8519248

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2261–2269). https://doi.org/10.1109/CVPR.2017.243

Karacan, C. Ö., Ruiz, F. A., Cotè, M., & Phipps, S. (2011). Coal mine methane: A review of capture and utilization practices with benefits to mining safety and to greenhouse gas reduction. *Int. J. Coal Geol*, *86*(2–3), 121–156. https://doi.org/10.1016/j.coal.2011.02.009

Keuchel, J., Naumann, S., Heiler, M., & Siegmund, A. (2003). Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. *Remote Sens. Environ*, *86*(4), 530–541. https://doi.org/10.1016/S0034-4257(03)00130-5

Kholod, N., Evans, M., Pilcher, R. C., Roshchanka, V., Ruiz, F., Coté, M., & Collings, R. (2020). Global methane

emissions from coal mining to continue growing even with declining coal production. *Journal of Cleaner Production, 256*(120489). https://doi.org/10.1016/j.jclepro.2020.120489

Krizhevsky, A. (2010). Convolutional deep belief networks on cifar-10. Retrieved from http://scholar.google.com/scholar?

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2007). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1201/9781420010749

Kroupi, E., Kesa, M., Navarro-Sánchez, V. D., Saeed, S., Pelloquin, C., Alhaddad, B., Moreno, L., Soria-Frisch, A., & Ruffini, G. (2019). Deep convolutional neural networks for land-cover classification with Sentinel-2 images. *J. Appl. Remote Sens*, *13*(2), 1. https://doi.org/10.1117/1.jrs.13.024525

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (1994). Remote Sensing and Image Interpretation. Wiley International Edition (5th ed., Vol. 53),550–552. Wiley. https://doi.org/10.1017/CBO9781107415324.004

Lobo, F., De, L., Souza-Filho, P. W. M., Novo, E. M. L., De, M., Carlos, F. M., & Barbosa, C. C. F. (2018). Mapping mining areas in the Brazilian amazon using MSI/Sentinel-2 imagery (2017). *Remote Sens*, *10*(8), 1–15. https://doi.org/10.3390/rs10081178

Mather, P. M., & Tso, B. (2009). *Classification Methods for Remotely Sensed Data, Second Edi*. ed. Taylor & Francis.

Myhre, G., Shindell, D., Bréon, F.M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., Zhan, H., 2013. Anthropogenic and Natural Radiative Forcing. Clim. Chang. 2013. *Phys. Sci. Basis Work. Gr. I Contrib. to Fifth Assess. Rep. Intergov. Panel Clim. Chang. 9781107057*, 659–740. https://doi.org/10.1017/CBO9781107415324.018

Pandey, S., Gautam, R., Houweling, S., Van Der Gon, H. D., Sadavarte, P., Borsdorff, T., Hasekamp, O., Landgraf, J., Tol, P., Van Kempen, T., Hoogeveen, R., Van Hees, R., Hamburg, S. P., Maasakkers, J. D., & Aben, I. (2019). Satellite observations reveal extreme methane leakage from a natural gas well blowout. *Proc. Natl. Acad. Sci. U. S. A*, *116*(52), 26376–26381. https://doi.org/10.1073/pnas.1908712116

Ramani, R. V., & Evans, M. A. (2020). Coal mining - Underground mining. Retrieved December 5, 2019, from https://www.britannica.com/technology/coal-mining https://www.britannica.com/technology/coal-mining

Richards, J. A. (2013). Remote sensing digital image analysis: An introduction. Springer (Vol. 9783642300). https://doi.org/10.1007/978–3–642–30062–2

Sehgal, S. (2012). Remotely Sensed LANDSAT Image Classification Using Neural Network Approaches. *Computer Science, 2*(5), 43–46. Retrieved from https://www.semanticscholar.org/paper/Remotely-Sensed-LANDSAT-Image-Classification-Using-Sehgal/ff66711f110389327c7aa3b34b81c664866d8d25

Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science, abs/1409.1*, 1–14. Retrieved from http://arxiv.org/abs/1409.1556

song, Y. Y., & Lu, Y. (2015). Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry*, *27*(2), 130–135. https://doi.org/10.11919/j..1002-0829.215044

Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019). BigEarthNet: A Large-Scale Benchmark Archive For Remote Sensing Image Understanding. IGARSS 2019–2019. *IEEE International Geoscience and Remote Sensing Symposium,* 5901–5904. https://doi.org/10.1109/IGARSS.2019.8900532

Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. *Ore Geology Reviews, 109*, 26–49. https://doi.org/10.1016/j.oregeorev.2019.04.003

Tang, D. G., & Spikes, K. T. (2017). Segmentation of shale SEM images using machine learning. In T. Houston (Ed.), SEG Internation Exposition and Annual Meeting (pp. 3898–3902). Retrieved from https://onepetro.org/SEGAM/proceedings-abstract/SEG17/All-SEG17/SEG-2017-17738502/102531

Trippi, M. H. H., & Tewalt, S. J. J. (2011). Geographic information system (GIS) representation of coal-bearing areas in India and Bangladesh: Open-File Report (2011–1296). https://doi.org/https://doi.org/10.3133/ofr20111296

U.S.EPA. (2019). Global Anthropogenic Non-CO2 Greenhouse Gas Emissions: 2015–2050. Retrieved from http://www.epa.ie/pubs/reports/air/airemissions/ghgprojections2017-2035/2018_Seminar_GHG_Projections_to_2035.pdf f

Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S., Aben, I., Scarpelli, T., & Jacob, D. J. (2019). Satellite Discovery of Anomalously Large Methane Point Sources From Oil/Gas Production. *Geophys. Res. Lett*, *46*(22), 13507–13516. https://doi.org/10.1029/2019GL083798

Zeng, X., Liu, Z., He, C., Ma, Q., & Wu, J. (2017). Detecting surface coal mining areas from remote sensing imagery: An approach based on object-oriented decision trees. *J. Appl. Remote Sens*, *11*(1), 015025. https://doi.org/10.1117/1.JRS.11.015025