

LITHOLOGICAL CLASSIFICATION USING MULTI-SENSOR DATA AND CONVOLUTIONAL NEURAL NETWORKS

M. Brandmeier^{1,*}, Y. Chen^{1,2}

¹ Esri Deutschland GmbH, Ringstr. 7, D-85402 Kranzberg – m.brandmeier@esri.de

² Technical University of Munich, Arcisstr. 21, D-80333 München - chenyz61@hotmail.com

KEY WORDS: Deep Learning, ArcGIS, Lithological Classification, Geology, ASTER, Sentinel-2

ABSTRACT:

Deep learning has been used successfully in computer vision problems, e.g. image classification, target detection and many more. We use deep learning in conjunction with ArcGIS to implement a model with advanced convolutional neural networks (CNN) for lithological mapping in the Mount Isa region (Australia). The area is ideal for spectral remote sensing as there is only sparse vegetation and besides freely available Sentinel-2 and ASTER data, several geophysical datasets are available from exploration campaigns. By fusing the data and thus covering a wide spectral range as well as capturing geophysical properties of rocks, we aim at improving classification accuracies and support geological mapping. We also evaluate the performance of the sensors on their own compared to a joint use as the Sentinel-2 satellites are relatively new and as of now there exist only few studies for geological applications. We developed an end-to-end deep learning model using Keras and Tensorflow that consists of several convolutional, pooling and deconvolutional layers. Our model was inspired by the family of U-Net architectures, where low-level feature maps (encoders) are concatenated with high-level ones (decoders), which enables precise localization. This type of network architecture was especially designed to effectively solve pixel-wise classification problems, which is appropriate for lithological classification. We spatially resampled and fused the multi-sensor remote sensing data with different bands and geophysical data into image cubes as input for our model. Pre-processing was done in ArcGIS and the final, fine-tuned model was imported into a toolbox to be used on further scenes directly in the GIS environment. The tool classifies each pixel of the multiband imagery into different types of rocks according to a defined probability threshold. Results highlight the power of using Sentinel-2 in conjunction with ASTER data with accuracies of 75% in comparison to only 70% and 73% for ASTER or Sentinel-2 data alone. These results are similar but examining the different classes shows that there are significant improvements for classes such as dolerite or carbonate sediments that are not that widely distributed in the area. Adding geophysical datasets reduced accuracies to 60%, probably due to an order of magnitude difference in spatial resolution. In comparison, Random Forest (RF) and Support Vector Machines (SVMs) that were trained on the same data only achieve accuracies of 46 % and 36 % respectively. Most insecurity is due to labelling errors and labels with mixed lithologies. However, results show that the U-Net model is a powerful alternative to other classifiers for medium-resolution multispectral data.

1. INTRODUCTION

Remote sensing has a long history for geological applications and spectral data has been used in many studies to support geological mapping, exploration targeting and lithological classification, mainly by applying band-ratio techniques as well as different machine-learning approaches for classification (e.g. Brandmeier et al. 2013; Hewson 2005; Ninomiya 2005; Rowan 2005). Deep Learning has been successfully used in different fields of research, being named one of the 10 break-through technologies in 2013 (MIT 2013). For computer vision tasks, deep-learning has been pushed by companies such as Google, Baidu, Microsoft and Facebook (Zhu et al. 2017) and is becoming increasingly popular in remote sensing tasks such as object detection and classification (e.g. Kamilaris and Prenafeta-Boldú 2018; Mahdianpari et al. 2018; Zhu et al. 2017 and many more). However, deep learning approaches and an end-to-end integration into a GIS environment are still scarce for geological applications.

Thus, in the present study we evaluate a deep-learning approach for lithological classification and compare results to classical machine-learning algorithms using different multispectral and geophysical datasets. Our goal is to develop a workflow for geological mapping by integrating a trained convolutional neural network (CNN) for automated lithological classification into the ArcGIS Platform to support field-based data collection and map construction.

2. GEOLOGY OF THE STUDY AREA

Mount Isa (Australia) is well-known to host world-class base-metal deposits and due to the great economic and scientific interest, there exist many detailed studies of area (e.g. Murphy et al. 2011; Wilde 2011).

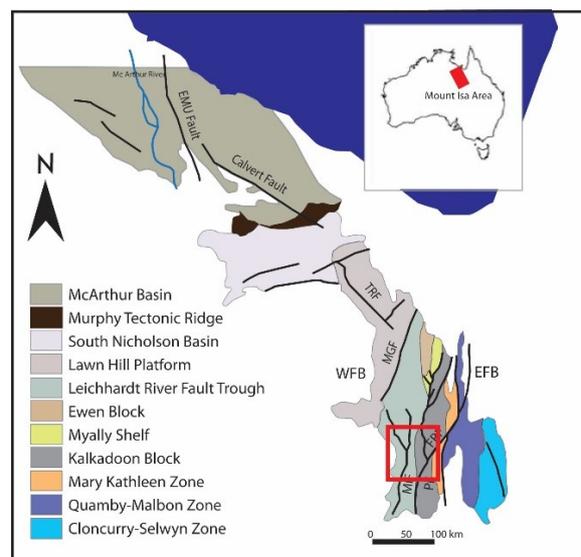


Figure 1. Schematic map of the Mt Isa and Mc Arthur River regions with major domains and faults. MIF: Mt Isa Fault, MGF= Mt Gordon Fault, PF= Pilgrim Fault, FRF= Fountain Range Fault, TRF= Termite Range Fault, EFB= Eastern Fold Belt, WFG = Western Fold Belt. Red square = study area. (after Murphy et al. 2011)

The Mount Isa Inlier is divided into the Western and Eastern Fold Belts and the central Kalkadoon-Leichhardt Belt (Fig. 1). In terms of stratigraphy and architecture, three different periods are distinguished: basement rocks (>1900-1820 Ma), Leichhardt Superbasin (1752-1740 Ma), Calvert Superbasin (1740-1670 Ma) and Isa Superbasin (1670 -1640 Ma). Our study area is located in the Western Fold belt and stretches into the Kalkadoon-Leichhardt Belt and, thus, covers a wide range of rock units in a complex geodynamic setting (red square in Fig. 1). For more information about the geodynamic history and the stratigraphy, the reader is referred to the respective publications as cited above. For our purpose, the lithological units were classified according to the dominant rock type. This is necessary as we can only capture major physical properties of rocks/minerals by using spectral data at medium-resolution remote sensing data. Thus, we can hardly distinguish between, for example, two carbonatites from different stratigraphic units. Figure 2 shows the corresponding lithological map of our study area.

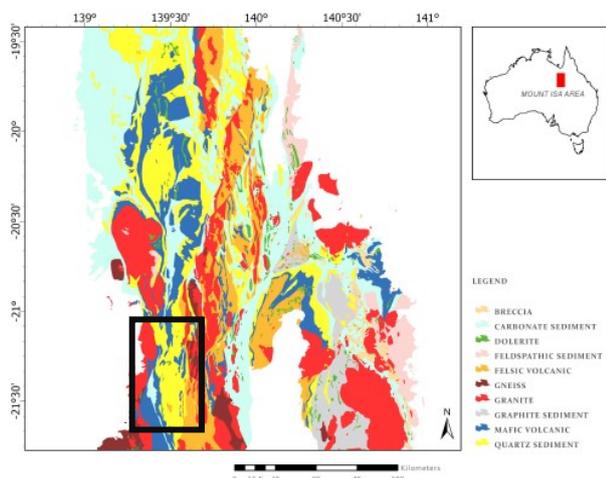


Figure 2. Lithological map of the Mount Isa area classified with respect to the dominant rock type. Black square: study area. Data: (Queensland Department of Mines and Energy 2000)

3. DATA AND METHODOLOGY

The overall workflow and data used are shown in Figure 3. A detailed description of the data and the methods used is provided in the following paragraphs.

3.1 Data

One goal of our study was to evaluate Sentinel-2 data for geological mapping. As the Sentinel-2 mission is relatively young (Sentinel-2A and B were launched in 2015 and 2017 respectively), there are not yet many studies investigating its potential for geological applications (e.g. Van der Werff and Van der Meer 2015, 2016). Thus, we downloaded a Sentinel-2 atmospherically corrected (L2A) dataset covering the study area. The scene was collected on April 3rd, 2018 and was selected to

have a cloud coverage of 0%. Bands 1,9 and 10 were excluded from analysis as they only have a spatial resolution of 60m compared to the 10m and 20m spatial resolutions of the other bands and are not important for our purpose. Most spectral bands of Sentinel-2 are in the visible and near infrared (VNIR) (Fig. 4). However, many rock-forming and alteration minerals such as pyroxenes, carbonate minerals and micas have absorption features located in the shortwave infrared (SWIR) of the electromagnetic spectrum. Thus, we also used ASTER data that has a long history in geological remote sensing as there are several spectral bands located in the SWIR area (Figure 4). ASTER data was acquired as atmospherically corrected data product. The scene is from October 3rd, 2004 and was also selected according to the criteria of 0% cloud cover. In addition to the spectral datasets, we also used radiometric data available from exploration campaigns in the area (Rogers 2009). In contrast to the ASTER and Sentinel-2 datasets with a spatial resolution of 15/30 and 10/20m respectively (compare Figure 4), the geophysical data is at 80m grid resolution. The grids used are U, Th and K radiometrics as well as specific ratios commonly used in exploration.

The labels used for training and validation are shown in Figure 2 and represent the dominant rock type from the GIS shapefile obtained from the Geological survey (Queensland Department of Mines and Energy 2000).

3.2 Methods

3.2.1 Data pre-processing

All data was spatially resampled to 10m resolution, the best spatial resolution of the Sentinel-2 bands, and projected to a common coordinate system (WGS 1984, UTM Zone 54S).

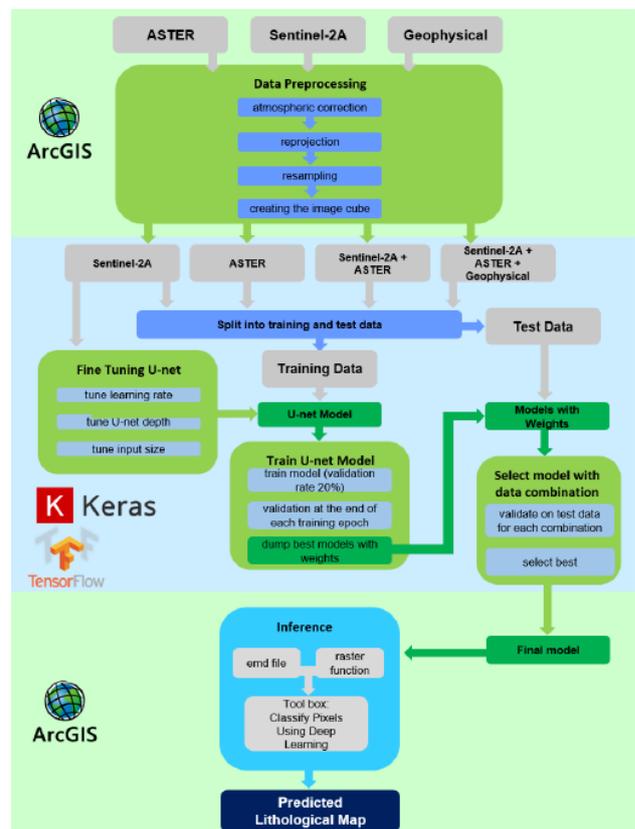


Figure 3. Flowchart showing the overall workflow, data and software used in this study.

Four different image cubes were then exported as labelled tiles (256 pixels x 256 pixels) as input for the CNN: 1) the Sentinel-2 data, 2) the ASTER data, 3) Sentinel-2 + ASTER and 3) Sentinel-2 + ASTER + geophysical data. This was done to compare the accuracies achieved by each combination (Fig. 3).

The exported tiles were split into training and testing datasets (90% and 10%). The training data is automatically split into training and validation during the training process (80% and 20%). To increase the number of training data, geometric data augmentation (rotation) was used on 10% of the tiles. Furthermore, even though the area is arid, we created a mask to eliminate the effect of the sparse vegetation. This was done by calculating and interactively thresholding the NDVI (threshold of 0.3) based on the Sentinel-2 data.

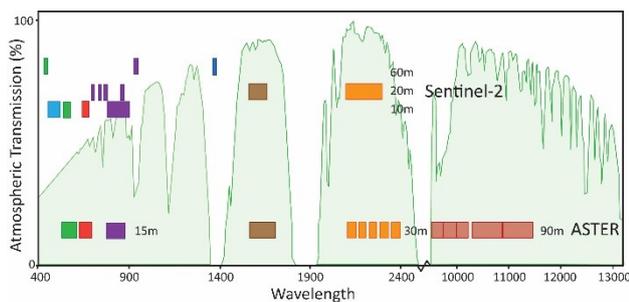


Figure 4. Sentinel-2, ASTER satellite characteristics.

Atmospheric transmittance is plotted on the y-axis (after Van der Meer 2016).

3.2.2 CNN architecture and experiments

Deep learning has been widely used for object detection and classification and there exist many highly performant architectures such as InceptionV3, VGG16, VGG19, ResNet50 and many more (e.g. Mahdianpari et al. 2018). As we are using multispectral data and aim at classifying rocks that are very different from other objects, we cannot use networks that are pretrained on large computer vision datasets but need to train a new network that is well-suited for smaller datasets. Thus, we use a U-Net architecture which was originally invented for biomedical image segmentation (Ronneberger et al. 2015). It consists of an encoding and a decoding path with the encoding path being a typical CNN. During the contraction, the spatial information is reduced while feature information is increased. The decoding path combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the encoding path (Figure 5).

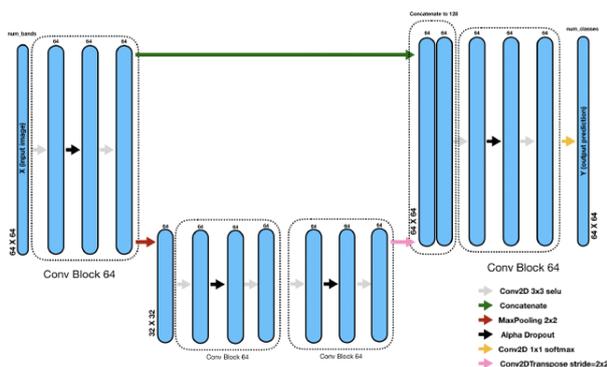


Figure 5. Final U-Net architecture used for lithologic classification.

To find the optimal hyperparameters for our problem, some fine-tuning experiments were conducted. We tested different model depth, input image sizes and learning rates to achieve the best overall accuracy on the validation dataset. The overall validation accuracy (average of the last 20 training epochs) was used as criteria to choose the best parameters. In addition, we monitored the training curves of validation accuracy and loss. Based on these experiments, the learning rate for the final network was set to 0.001 and the model depth to two convolutional blocks of 64 convolutions on each path (Figure 5). The input image size was 64x64 pixels x number of bands. The Adam optimizer (Kingma and Ba 2015) was used for all experiments and loss is calculated by the cross-entropy loss function.

In order to compare these results to traditionally used machine-learning algorithms, we also ran a Random Forest Classifier and a Support Vector Machine on the same dataset. The kernel function from the SVM classifier was the RBF kernel and the error tolerance used was 0.001. The RF classifier had a tree depth of 6 and 100 trees.

For validation we used our test dataset and calculated different metrics to compare our results: overall accuracy, confusion matrices, F1 scores and Intersection over union (IoU).

3.2.3 U-Net integration in ArcGIS Desktop

As our goal was to provide a workflow that is scalable and can be applied easily to future datasets, the final U-Net was imported into a toolbox in ArcGIS Pro 2.3. This was done by using the tool “Classify Pixels Using Deep Learning” that requires the trained Network as well as a .json file and a raster function (a detailed description of this workflow can be found here: <https://github.com/Esri/raster-deep-learning>). Using the same input data, future scenes can be classified rapidly, and results can be synchronized with the Esri cloud or an ArcGIS server to be accessed by mobile mapping devices in the field. This allows for a streamlined workflow for fieldwork and supports geological map production.

4. RESULTS

Results for all data combinations are summarized in Table 1. The best overall accuracy of 75% was achieved by fusing ASTER and Sentinel-2 data. Using each sensor alone, accuracies are slightly lower. However, comparing the results for each class (Figure 6), we observe that several lithologies are classified well using only one sensor, others cannot be captured at all (e.g. Carbonate sediment or Gneiss). Thus, by gaining a broader spectral range when fusing ASTER and Sentinel-2, we improve the classification significantly for several classes (especially for Dolerite, Carbonate sediments or Gneiss). This is not reflected in the overall accuracy metric as classes are not distributed equally in the area.

Dataset	Overall Accuracy	Confidence Rate
ASTER	0.7007	0.6165
Sentinel-2	0.7329	0.7790
ASTER + Sentinel-2	0.7482	0.8340
ASTER+Sentinel-2+Geophysical	0.6064	0.3672

Table 1. Final result for each dataset. Best results are achieved by fusing Sentinel-2 and ASTER data.

Surprisingly, adding the geophysical data does not help to improve classification results but reduces the overall accuracies. One reason for this might be the different spatial resolution of the original data that is an order of magnitude different from the other sensors. As the U-Net extracts features at different levels, spatial resolution might be an important issue and should be investigated further.

In comparison to the U-Net results, the SVM and RF classifier only achieve accuracies of 36% and 46% respectively when using the same input data.

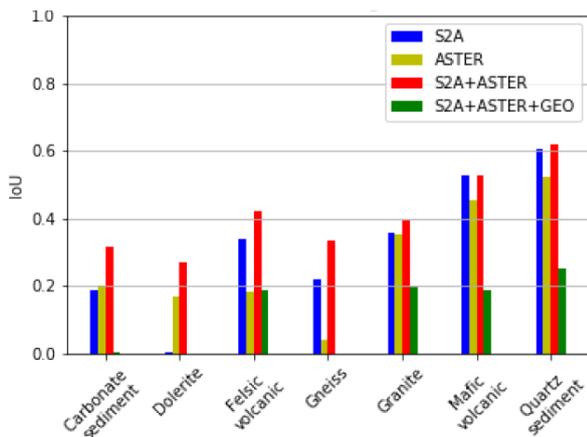


Figure 6. Intersection over Union for each lithological class. Some classes (Carbonate sediment, Dolerite, Gneiss) only achieve good results by combining ASTER and Sentinel-2 data.

Figure 7. shows the confusion matrix of the final ASTER+Sentinel-2 result and Figure 8. the classification map. Especially the Gneiss class is still not captured well. Most confusion occurs with Granite which is spectrally very similar and might also be similar in terms of morphology and therefore extracted features.

As the classification is based on the dominant rock unit, results achieved by using a deep-learning approach capable of feature extraction are very good compared to the RF and SVM classifier that only classifies based on the spectral information. Even minor occurrences of other rocks within one label can lead to a substantial reduction of classification accuracy as this affects spectral signatures significantly.

True label \ Predicted label	Carbonate sediment	Dolerite	Felsic volcanic	Gneiss	Granite	Mafic volcanic	Quartz sediment
Carbonate sediment	36,618	12	74	50	467	11,385	12,719
Dolerite	13	2,243	253	1,877	218	348	810
Felsic volcanic	2,294	0	120,369	3,692	13,848	3,088	13,265
Gneiss	820	315	4,177	48,147	18,818	8,571	230
Granite	7,032	8	44,105	32,311	215,790	26,605	12,975
Mafic volcanic	16,464	599	3,000	1,824	19,365	299,667	80,269
Quartz sediment	9,574	0	15,206	581	8,994	28,915	481,071

Figure 7. Confusion matrix of the final classification

For the final classification in ArcGIS Pro we also added another class for pixels that have scores of less than 0.5 for each class and are therefore considered as not confidential. Note that vegetation also falls into this class. These pixels are named “uncertainty class” and might be checked in the field using mobile mapping. This can be achieved easily by synchronizing the results with the cloud or server and accessing them with Collector for ArcGIS for mobile mapping and sample collection.

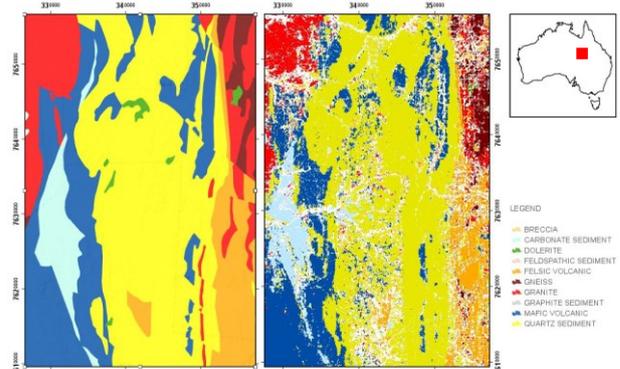


Figure 8. Results of the final U-Net. Right: Labels; Left: Results. Note the white pixels. This is where results are below 0.5 for each class and we cannot predict accurately.

5. DISCUSSION & SUMMARY

This study evaluated several aspects that need separate discussion: (1) we present a preliminary approach to use deep-learning for lithological mapping and compare this approach to purely spectral classifiers; (2) we evaluate Sentinel-2 data with respect to its capacity to capture spectral characteristics in comparison to ASTER data and image cubes with both datasets as well as radiometric data; (3) we provide a complete integration into ArcGIS and, thus, a workflow that can be applied on other scenes by non-data scientists, is scalable and allows smooth data access on mobile devices.

With respect to the performance of deep learning we could show that U-Net is a great alternative to traditionally used machine-learning algorithms for multispectral classification tasks. This is especially true if training and testing data covers large areas and no spectrally pure regions are selected for training. In the latter scenario SVM and RF classifiers perform much better than in the setting of this study (Brandmeier et al. 2013; Rowan 2003). The capability of feature extraction and therefore also a spatial component obviously helps to correctly classify the different lithologies. This agrees with findings of other studies on landcover classification using multispectral data and deep learning (Nijhawan et al. 2019; Zhang et al. 2019). The biggest issue in our study was that labels were not very reliable and only referred to the dominant rock type. This is why the purely spectral approach of RF and SVM classifiers only achieved very low overall accuracies. The U-Net still performs at a high overall accuracy due to the feature extraction. Future studies are necessary to investigate this in more detail.

With respect to Sentinel-2 we found that the data on its own classifies some types of rocks even better than ASTER in this deep-learning approach, namely mafic volcanics, felsic volcanics and quartz sediments. This might be due to the higher spatial resolution of 10m compared to the 30m of ASTER and to the higher spectral resolution in the VNIR range of the electromagnetic spectrum that is important to capture absorption caused by iron, for example. However, the combination with ASTER is what really allows to improve results on all classes as

the joint use of the two sensors covers both important spectral ranges, the VNIR and SWIR at an adequate spectral resolution. Finally, the integration into ArcGIS allows users not familiar with programming and data science to use the pre-trained network in a GUI. Thus, the approach is also scalable, and data can be accessed with mobile mapping devices during fieldwork.

6. CONCLUSIONS

In conclusion, deep-learning is very promising for classification tasks with multispectral data and can be seamlessly integrated into a GIS environment to support mapping and inspection tasks. Sentinel-2 data proved to be valuable for geological remote sensing, especially when combined with ASTER data. The joint use of radiometric data was problematic, possibly due to the different spatial resolution and needs further investigation.

ACKNOWLEDGEMENTS (OPTIONAL)

We want to thank Esri Germany for funding the thesis that led to this manuscript. In addition, many thanks to Dr. Arianne Ford for providing the geological data and data references.

REFERENCES

- Brandmeier, M., Erasmi, S., Hansen, C., Höweling, A., Nitzsche, K., Ohlendorf, T., Mamani, M., & Wörner, G., 2013. Mapping patterns of mineral alteration in volcanic terrains using ASTER data and field spectrometry in Southern Peru. *Journal of South American Earth Sciences*, 48, 296-314.
- Hewson, R.D.C., Mizuhiko, S., Ueda, K., Mauger, A.J., 2005. Seamless geological map generation using ASTER in the Broken Hill-Curnamona province of Australis. *Remote Sensing of Environment*, 159-172.
- Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90.
- Kingma, D.P., Ba, L.J., 2015. Adam: A Method for Stochastic Optimization. In, *International Conference on Learning Representations (ICLR)*: arXiv.org
- Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F., Zhang, Y., 2018. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. *Remote Sensing*, 10, 1119.
- MIT (2013). MIT Technology Review. (<https://www.technologyreview.com/lists/technologies/2013/>, accessed 26.06.2019)
- Murphy, F.C., Hutton, L.J., Walshe, J.L., Cleverley, J.S., Kendrick, M.A., McLellan, J., Rubenach, M.J., Oliver, N.H.S., Gessner, K., Bierlein, F.P., Jupp, B., Aillères, L., Laukamp, C., Roy, I.G., Miller, J.M., Keys, D., Nortje, G.S., 2011. Mineral system analysis of the Mt Isa–McArthur River region, Northern Australia. *Australian Journal of Earth Sciences*, 58, 849-873.
- Nijhawan, R., Joshi, D., Narang, N., Mittal, A., Mittal, A., 2019. A Futuristic Deep Learning Framework Approach for Land Use-Land Cover Classification Using Remote Sensing Imagery. In J.K. Mandal, D. Bhattacharyya, & N. Auluck (Eds.), *Advanced Computing and Communication Technologies*. 87-96.
- Ninomiya, Y.F., B.; Cudahy, T., 2005. Detecting lithology with Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) multispectral thermal infrared "radiance at sensor" data. *Remote Sensing of Environment*, 99, 127-139.
- Queensland Department of Mines and Energy, Taylor Wall and Associates, SRK Consulting Pty. Ltd., ESRI Australia, 2000. North-West Queensland Mineral Province Report. Queensland Department of Mines and Energy. Brisbane.
- Rogers, K.A., 2009. Combined annual technical report for exploration licenses EL 26302, 26303, 26304, 26305, 26307, 26308, 26309, 26310, 26311, 26312, 26314 and EL 26701, 26702, 26703, Georgina Basin, Northern Territory for year ending 7 April 2009. Australis Exploration Pty Ltd, Dept. of Regional Development, Primary Industry, Fisheries and Resources.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: N. Navab, J. Hornegger, W.M. Wells, & A.F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 234-241. Cham: Springer International Publishing.
- Rowan, L.C.M., J.C., 2003. Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. *Remote Sensing of Environment*, 350-366.
- Rowan, L.C.M., J.C.; Simpson, C.J., 2005. Lithologic mapping of the Mordor, NT, Australia ultramafic complex by using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *Remote Sensing of Environment*, 105-126.
- Van der Werff, H., Van der Meer, F., 2015. Sentinel-2 for Mapping Iron Absorption Feature Parameters. *Remote Sensing*, 7, 12635-12653.
- Van der Werff, H., Van der Meer, F., 2016. Sentinel-2A MSI and Landsat 8 OLI Provide Data Continuity for Geological Remote Sensing. *Remote Sensing*, 8, 883.
- Wilde, A.R., 2011. Mount Isa copper orebodies: improving predictive discovery. *Australian Journal of Earth Sciences*, 58, 937-951.
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M., 2019. Joint Deep Learning for land cover and land use classification. *Remote Sensing of Environment*, 221, 173-187.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5, 8-36.