

Title: An application of CNN to classify barchan dunes into asymmetry classes

Authors:

Barend van der Merwe<sup>a,\*</sup>

<sup>a</sup>Department of Geography, Geoinformatics and Meteorology

University of Pretoria

Cnr Lynwood Road and Roper Street

Hatfield

Pretoria

South Africa

\*Corresponding Author

Email: barend.vandermerwe@up.ac.za

Nelishia Pillay<sup>b</sup>

<sup>b</sup>Department of Computer Science

University of Pretoria

Cnr Lynwood Road and Roper Street

Hatfield

Pretoria

South Africa

Email: npillay@cs.up.ac.za

Serena Coetzee<sup>a</sup>

<sup>a</sup>Department of Geography, Geoinformatics and Meteorology

University of Pretoria

Cnr Lynwood Road and Roper Street

Hatfield

Pretoria

South Africa

Email: serena.coetzee@up.ac.za

# An application of CNN to classify barchan dunes into asymmetry classes

---

## Abstract

Barchan morphometric data have been used as proxies of meteorological and topographical data in environments where this data is lacking (such as other planetary bodies), gaining insights into barchan dune field dynamics such as barchan collision and sediment dynamics, and estimating migration speeds. However, manual extraction of this data is time-consuming which can impose limits on the spatial extent and temporal frequencies of observations. Combining remotely sensed big data with automated processing techniques such as convolutional neural networks (CNNs) can therefore increase the amount of data on barchan morphology. However, such techniques have not yet been applied to barchans and their efficacy remains unknown. This study addresses this issue by evaluating the classification performance (using the ACC, F 1 -score and MCC metrics) of CNNs on several different morphometric tasks: the side of horn elongation, the magnitude of elongation, the barchans a/c ratio, and a new metric, bilateral asymmetry, which takes a more holistic view of barchan asymmetry. Specifically, bilateral asymmetry offers a means by which the total points of variation on a barchan that is used in describing barchan morphology, can be expressed with a single measure. Twelve different CNN architectures, each with different hyperparameters, are trained and tested on a sample of 90 barchan dunes. Additionally, the potential of transfer learning is assessed using the VGG16 and ResNet50 architectures. The results show that the accuracy of the CNNs can exceed 80% in some cases and that "from scratch" CNNs can match the performance obtained using transfer learning approaches.

*Keywords:* Barchan Asymmetry, Convolutional Neural Networks, VGG16, ResNet50, Outline Classification

---

## 1. Introduction

Barchan dunes are crescent-shaped mobile free dunes (Bourke and Goudie, 2009; Worman *et al.*, 2013; Engel *et al.*, 2018; Parteli *et al.*, 2014) that form in regions of limited sand supply and unidirectional (or narrowly bimodal) wind regimes (Barnes, 2001; Bourke and Goudie, 2009; Bourke, 2010; Elbelrhiti *et al.*, 2008). Their relatively simple morphology (Fig 1), which can change significantly over time (Pike, 2000), has applications in several different fields of study. Barchan shape is often described as an  $\frac{a}{c}$ -ratio which relates the length

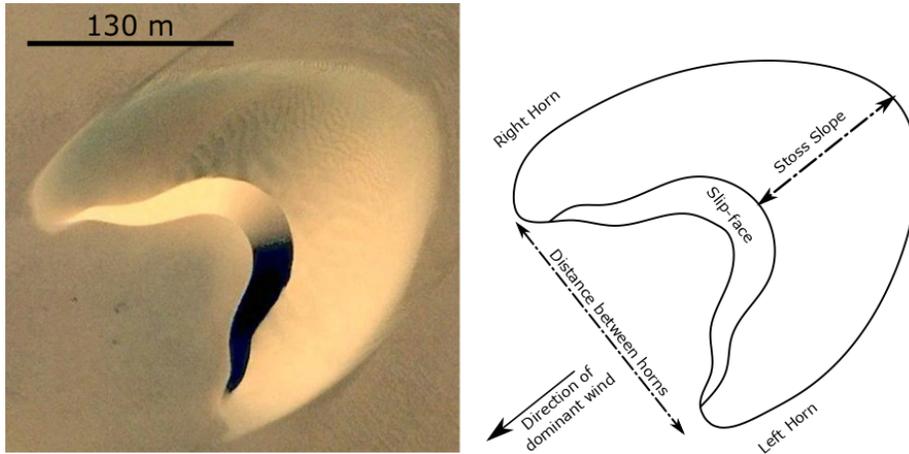


Figure 1: A barchan dune from Sudan along with a description of the main morphological components. Image source: Google Earth™.

of the dune’s stoss slope ( $a$ ) to the distance between the dune’s horns ( $c$ ) (Long and Sharp, 1964; Norris, 1966; Hamdan *et al.*, 2016; Bourke and Goudie, 2009; Barnes, 2001; Boulghobra and Dridi, 2016; Lorenz *et al.*, 2013; Chojnacki *et al.*, 2011). The  $\frac{a}{c}$  ratio can be used to describe the effects of changing sediment flux on barchans (Parteli *et al.*, 2014) and has been linked to variations in the migration speed of barchans (Hamdan *et al.*, 2016). This latter aspect is important for settlements in arid areas given the risk that migrating barchans pose to infrastructure (Pike, 2000; Aydda *et al.*, 2020; Moosavi *et al.*, 2014; Abdu, 1976).

The asymmetry of barchans, which is considered to be present when one of the horns is longer than the other (Bourke, 2010; Barnes, 2001; Tsoar and Parteli, 2016; Lv *et al.*, 2016; ?; Parteli *et al.*, 2014), has also been used in several different research contexts. The simplest expression of asymmetry, namely the side on which the elongated horn occurs, can be used to discern between bimodal wind regimes and dune collisions as agents that modify dune morphology on a large scale (Bourke, 2010; Tsoar and Parteli, 2016; Boulghobra, 2016; Zhang *et al.*, 2018). It can also be used to infer modifications to upwind airflow (Barnes, 2001), barchan sediment dynamics (Franklin and Charru, 2011; Scheidt and Lancaster, 2013) and, for regions where the data is absent, the local topography of a region (Parteli *et al.*, 2014; Bourke, 2010). However, it must be borne in mind that such inferences cannot be made from a single observation of the barchans. To be truly effective, multiple observations need to be made at different periods to fully understand the dynamics that are present within the dune field.

The magnitude of difference in length between the horns is also an important aspect of barchan morphology. Both Parteli *et al.* (2014) and Bourke (2010) show that an increase in the magnitude of difference can be related to

Table 1: Example studies where the different methods depicted in Fig 2 have been used.

Example		Example	
1	Norris and Norris (1961)	14	Franklin and Charru (2011)
2	Bourke <i>et al.</i> (2006)	15	El belrhiti and Douady (2011)
3	Gay (1999)	16	Douglass (1909)
4	Durán <i>et al.</i> (2010)	17	Andreotti <i>et al.</i> (2002)
5	Lonsdale and Malfait (1974)	18	Elbelrhiti <i>et al.</i> (2008)
6	Barnes (2001)	19	Wang <i>et al.</i> (2007)
7	Courrech du Pont (2015)	20	Burrough <i>et al.</i> (2012)
8	Hesse (2009)	21	Finkel (1959)
9	Silvestro <i>et al.</i> (2010)	22	Bailey (1906)
10	Long and Sharp (1964)	23	Sauermann <i>et al.</i> (2000)
11	Al-Harhi (2002)	24	Ma <i>et al.</i> (2014)
12	Parteli <i>et al.</i> (2014)	25	Todd (2005)
13	Rempel (1936)	26	Maghsoudi <i>et al.</i> (2017)

the duration of secondary oblique winds. The migration speed of barchans is dependent on the size of the dune with smaller dunes migrating faster than larger dunes (Sauermann *et al.*, 2000; Wiggs, 2013; Worman *et al.*, 2013; Gay, 1999). Therefore, collisions occur between slower moving barchans and faster moving barchans approaching upwind. The elongation of one horn can be used to identify cases where such a collision between two barchans can potentially occur (Bourke, 2010; Wang and Anderson, 2018) or has already occurred (Zhang *et al.*, 2018).

A considerable amount of information can, therefore, be obtained from collecting data on barchan morphology and several techniques have been developed to extract this data (Fig 2). However, these methods still require the manual extraction of data either in the field or from imagery. Earlier studies, such as Norris (1966), made use of repeated field surveys to monitor changes in barchan morphology over time. Such field surveys are expensive and time-consuming, especially when large numbers of barchans are being studied (Azzaoui *et al.*, 2020; Parteli *et al.*, 2014; Moosavi *et al.*, 2014). Recent advances in technology have led to an increase in the quantity of remotely sensed data (de Lima and Marfurt, 2020; Kattenborn *et al.*, 2021; Boulila *et al.*, 2021). This makes remotely sensed data a valuable source of data for barchan research (Bourke and Goudie, 2009; Boulghobra, 2016; Lorenz *et al.*, 2013; Vermeesch and Drake, 2008).

However, the effective use of remotely sensed data is problematic. The extraction of data from these images is still dominated by manual interpretation (Maxwell *et al.*, 2020). This leads to different interpretations associated with the same image (Pengra *et al.*, 2020; Van Coillie *et al.*, 2014), a reliance on the skill and experience of the interpreter (Hölbling *et al.*, 2017) and, constraints on the areal extent and temporal scale of the research due to the associated time constraints (Huang *et al.*, 2018; Witharana *et al.*, 2020; Gafurov and Yermolayev,

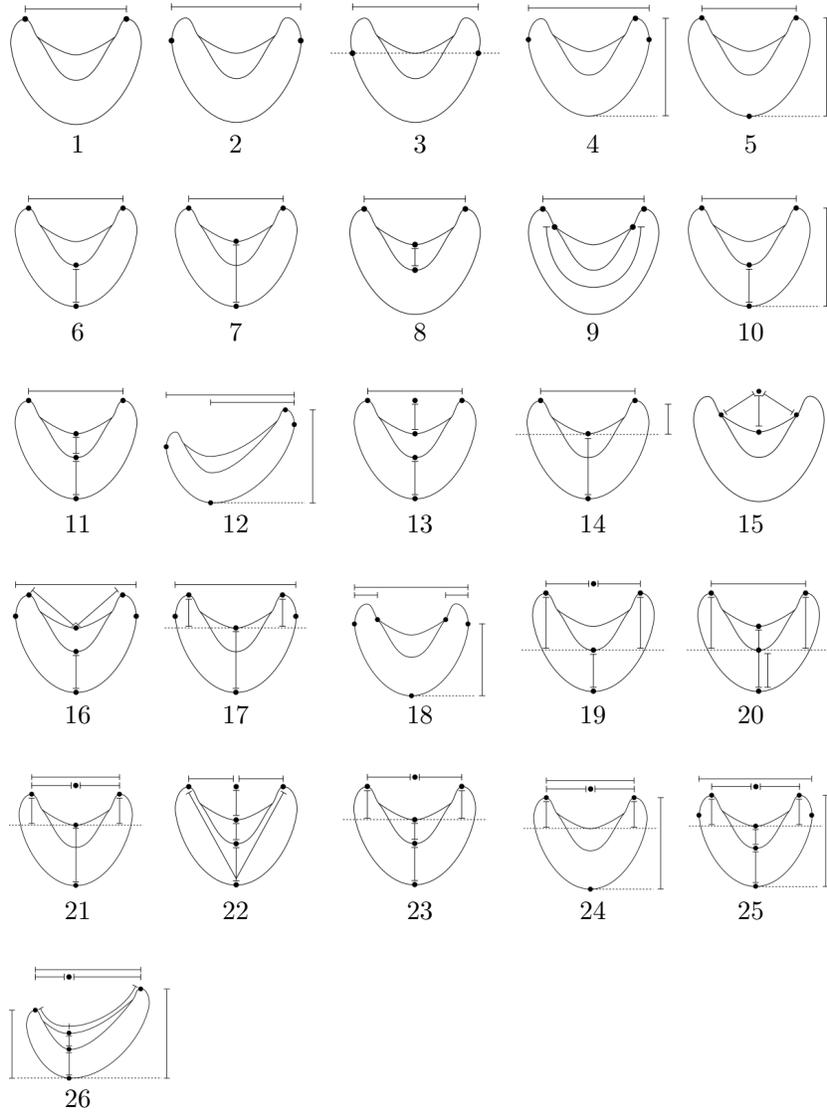


Figure 2: Different types of morphometric data that can be extracted from barchan dunes (cf Table 1). Filled circles ( $\bullet$ ) represent the points on the barchan between which measurements are taken.

2020; Baumhoer *et al.*, 2019; Liu *et al.*, 2019). This last point, which can be considered a scalability problem, can be potentially resolved using convolutional neural networks (CNNs) (Bhuiyan *et al.*, 2020).

CNNs are a class of artificial neural network that is characterised by the presence of convolutional layers (Traore *et al.*, 2018; Lecun *et al.*, 2015). These layers make use of convolutions to transform the image in such a way as to highlight particular patterns (Cady, 2017; Agarwal and Mittal, 2019). These patterns include vertical lines, horizontal lines, curved lines and many more complicated patterns. CNNs are, therefore, designed to learn the spatial features that best describe an object (Kattenborn *et al.*, 2021; Kubilius *et al.*, 2016) and have shown to be very accurate on classification tasks (De Cesare *et al.*, 2021). As such, they have a growing presence within geomorphology (Li and Hsu, 2020; Abolt and Young, 2020; Carbonneau *et al.*, 2020; Dang *et al.*, 2020; Huang *et al.*, 2018; Ji *et al.*, 2020; Palafox *et al.*, 2017; Chen *et al.*, 2018; Azzaoui *et al.*, 2020; Baumhoer *et al.*, 2019; Witharana *et al.*, 2020; Liu *et al.*, 2019; Gao *et al.*, 2021; Gafurov and Yermolayev, 2020). The application of the classification capabilities of CNNs to remotely sensed imagery also has the potential to assist in the interpretation of geomorphic processes (Buscombe and Ritchie, 2018).

However, there are some aspects of CNNs and their application to geomorphology that need further research. An image contains a lot of sources of information, such as shapes textures etc., and it is difficult to predict how a CNN will use this information when making predictions (Baker *et al.*, 2020). The performance of a CNN is also dependent on its hyperparameters (Shakya *et al.*, 2021). These are aspects such as the number of layers that are present within the network, whether regularizations are present or not, the types of regularizations etc. Their selection requires both skill and experience (Gu *et al.*, 2018). Since CNNs are currently not as widely used as more conventional analytical tools, a lot of users lack both skill and experience and therefore either use predefined hyperparameters, or find more suitable hyperparameters in a trial-and-error fashion (Ghorbanzadeh *et al.*, 2019).

CNNs are also strongly biased toward using the texture of the image as a basis for classification (Baker *et al.*, 2020; Geirhos *et al.*, 2019) even to the extent that changes that are not perceptible to humans can lead to misclassifications (Nguyen *et al.*, 2015; Carrara *et al.*, 2018). Although the extraction of edges from an image does limit the amount of texture information within the image (such as the work carried out by Azzaoui *et al.* (2020)), the impact that this will have on the classification accuracy of landforms is unclear.

Another aspect to consider is the subject within the image that is to be classified. A challenge for CNNs is the relationship between within-class diversity and between-class similarity (Cheng *et al.*, 2018). In other words, the diversity of shapes that the subject can take within each class and how similar objects that should be placed into different classes are to each other. Within geomorphology the ability to detect subtle differences in terrain features is considered essential (Li and Hsu, 2020), yet this poses a challenge for automated systems (Wilhelm *et al.*, 2020). This is of particular concern in barchan research given the variety of shapes that individual barchans can assume (Fig 3. This greatly

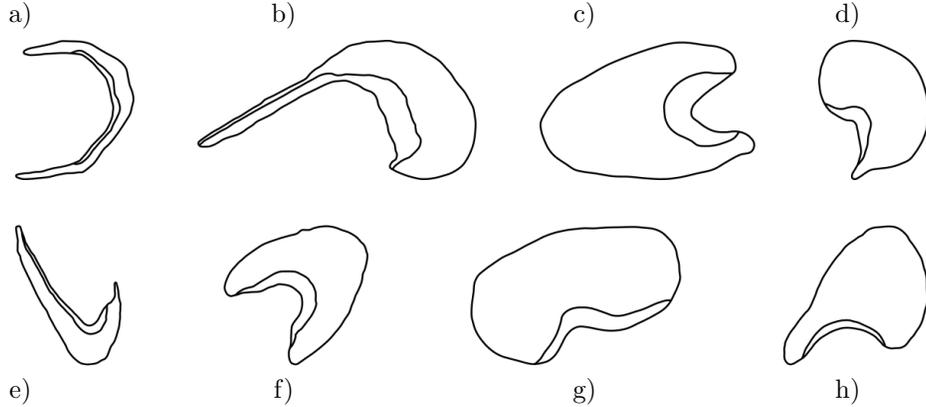


Figure 3: Examples of barchan shape varieties from Mars (top) and Earth (bottom). a-d) Mars (MGD image E0201255, HiRISE images ESP 023536 2560, ESP 020018 1385 and ESP 029660 1350 respectively). e-h) Google Earth<sup>TM</sup> images of barchans on Earth (Namibia, Sudan, USA and Saudi Arabia respectively). Images not to scale.

increases the level of within class diversity.

This study aims to address some of these questions. It aims to evaluate the suitability of different CNN configurations for extracting morphological information (i.e. shape and asymmetry) from barchan outlines. Two use cases are considered (following [Mahdianpari et al. \(2018\)](#)): developing a CNN that is trained on only the landform that it is intended to classify (i.e. creating a full-training network where the model architecture and training are controlled), and using a CNN that was pre-trained on a much larger and diverse data set and applying it to landform classification (i.e making use of the transfer learning paradigm). These two approaches have advantages and disadvantages associated with them. In the first use case, advantages include the ability to determine all of the hyperparameters that form part of the network. Further, the weight of the connections between the nodes may be more optimal for the use case. However, a big disadvantage is the need for large data sets to sufficiently train the network. The advantage of the second use case is that the network has already been trained on hundreds of thousands of images using hardware resources that the majority of users do not have access to. This saves the user time since they do not have to train the network to extract meaningful information from imagery and only focus on training the network to classify the image. A disadvantage, however, is that the user is not able to adjust the hyperparameters that were used during the initial training of the network.

By definition, all barchans are crescent-shaped and therefore a CNN must be able to use subtle local differences in shape to make a correct classification. This addresses the need put forward by [Li and Hsu \(2020\)](#). The use of outlines addresses the potential bias caused by image texture by removing most of the texture information within an image. And since it considers several different

CNN configurations, it addresses the concern raised about the proper selection of hyperparameters by providing a case study of model performance. Lastly, to address the needs of the research community, several different classification tasks are evaluated to approximate the research needs mentioned earlier.

The rest of the paper is laid out as follows. Section 2 contains background information on the origin of barchan shape and the processes that act on barchans to cause asymmetry. It also provides some background on the hyperparameters that need to be considered when constructing a CNN and the characteristics of two well-known CNN models, VGG16 and ResNet50, which will be used to evaluate the usefulness of transfer learning. In section 3 the characteristics of the study site and its barchans are discussed along with the details of the construction, training and evaluation of the different CNN models used. The results of the study are reported in section 4 and then contextualised in section 5. The paper concludes with a synopsis of the findings and suggestions for future research.

## 2. Barchan shape and asymmetry

Barchan dunes originate when an initial accumulation of sand obstructs the flow of wind at the surface (Lv *et al.*, 2016; Elbelrhiti, 2012). When wind encounters this sand accumulation, an acceleration of airflow occurs along its windward side. This results in sediment being eroded from this region (Elbelrhiti, 2012; Zhang *et al.*, 2014). Sediment eroded on the windward side accumulates at the summit of the sand pile where it collapses due to over-steepening resulting in the formation of a slip-face (Lv *et al.*, 2016; Zhang *et al.*, 2014; Wiggs, 2013). Once the slip-face has formed, the sand pile is considered to be a barchan dune (Elbelrhiti, 2012). The horns of the barchan form due to the lower height of the sand pile along its edges. These regions have a lower reconstitution time leading to a more rapid movement leading to downwind elongation and horn development (Wiggs, 2013).

Barchans often occur in an asymmetric shape (Parteli *et al.*, 2014; Lv *et al.*, 2016; Tsoar and Parteli, 2016) and several causal factors have been identified: bimodal winds, barchan collisions, topography, and variations in sediment supply to the dune. Bimodal winds (i.e. where the distribution of wind directions within a region has two distinct modes) result in the preferential elongation of one barchan horn. Three models have been proposed to explain this mechanism. In the model proposed by Bagnold (1954), barchans are aligned along a gentle but steady wind and it is the presence of intermittent strong oblique winds that cause the barchan's shape to deform. This deformation manifests as an extension of the horn closest to the oncoming oblique wind in front of and across the main body of the dune. In contrast, the model proposed by Tsoar (1984) has the barchan aligned along the stronger wind and it is the gentler wind that causes the extension of the horn opposite the oncoming oblique wind. This extension is away from the main body of the dune. It should be noted, however, that these two models were intended to describe the transition of a barchan dune into a seif dune (Tsoar and Parteli, 2016) and not as a cause

of barchan asymmetry. However, due to the association between the arise of barchan asymmetry and the development of a seif dune, these models are still applicable. The last model, proposed by Melton (1940), has one horn aligned with one wind direction and the other with a secondary wind direction. The barchan appears to be orientated at an angle that bisects these two wind directions and the horn that experiences the most elongation is associated with the strongest wind direction. This model has, however, not received widespread acceptance within the research community.

Since the movement speed of a barchan is a function of its size, with smaller barchans moving faster than larger barchans (Finkel, 1959; Norris, 1966; Worman *et al.*, 2013; Gay, 1999), any given dune field will contain collisions between faster moving upwind barchans and the slower barchans downwind. In cases where there is a lateral offset between the longitudinal axes of the two dunes, Such a collision can lead to a preferential elongation of one of the horns that are involved in the collision process (Parteli *et al.*, 2014). Finkel (1959) proposed that the slope of the surface on which the barchan migrates contributes to the emergence of asymmetry. Supporting evidence was provided by Parteli *et al.* (2014) who showed that there is an additional gravitational component that causes a horn to extend in the downslope direction away from the main body of the dune. Lastly, sediment supply variability can cause an alteration to the shape of the dune by influencing the rate of elongation of the horn receiving the additional influx of sediment. Two contrasting views have been proposed. One view, proposed by Parteli *et al.* (2014) holds that the downwind extension of the horn experiencing the additional sediment supply is slowed causing the opposite horn to appear elongated. In contrast, Lv *et al.* (2016) propose that the horn receiving the additional influx is elongated.

### 3. CNNs

#### 3.1. Functioning

A full and comprehensive discussion regarding the operation of CNNs is beyond the scope of this work. Instead, emphasis will be placed on some key aspects of the model’s functioning to show why it is capable of classifying barchan outlines. In simplistic terms, a CNN consists of three parts: a layer that receives the input image, a block of layers that act as a feature extractor, and a final block of layers which classifies the input image (Fig 4). The input image is a 3D collection of pixels with the width and height representing the dimensions of the image and the depth representing the number of channels in the image (3 in the case of an RGB image and 1 for grayscale images) (Chollet and Allaire, 2018). This input image is then transferred to a feature extraction block. These consist of convolutional layers and max-pooling layers (Boulila *et al.*, 2021). The role of the convolution layer is to extract the features within the input (Rocco *et al.*, 2017; Palafox *et al.*, 2017; Boulila *et al.*, 2021; Raghu *et al.*, 2020) while the max-pooling layers downsamples the feature map to reduce the computation demands of the network (Hao *et al.*, 2019; Wang *et al.*, 2019; Ghorbanzadeh

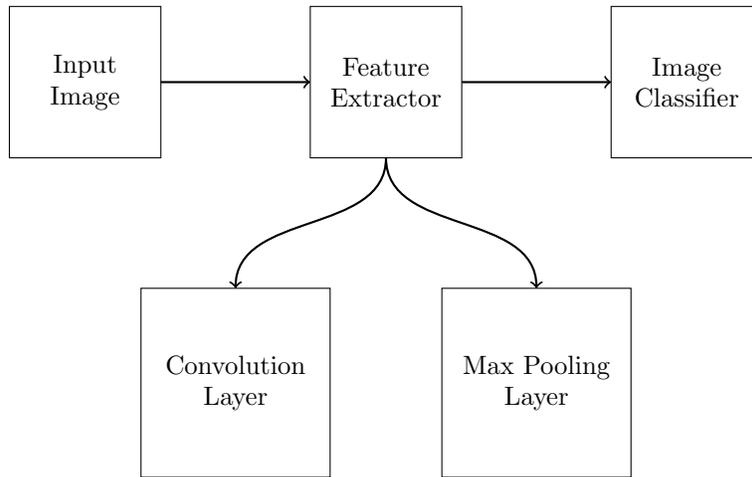


Figure 4: A simplified conceptual representation of a CNN.

*et al.*, 2019). It is important to note that any given CNN can contain one or more of these feature extraction blocks. As an output, the feature extraction process produces a feature map which is then passed on to a classifier block. The classifier block resembles a multilayer-perceptron, also referred to as a fully-connected layer, (Chevtchenko *et al.*, 2018; Hao *et al.*, 2019) and consists of an  $x$  amount of layers followed by a final output layer used to classify the original input image.

Regularizations can be placed in either the feature extraction block or the image classifier block and three types are generally used: L1, L2, and Dropout. They serve the purpose of making the CNN less complex (Chollet and Allaire, 2018; Thakkar and Lohiya, 2021) so that the network becomes more "generalized". This generalization is needed for the network to be able to classify imagery that it has never encountered before. The L1 and L2 regularizations are calculated values that get added to the loss function of the network (Thakkar and Lohiya, 2021). The loss function, in turn, is a function that is used to evaluate the performance of the CNN by computing the difference between the predicted outputs and the expected outputs (Chollet and Allaire, 2018). High values in the loss function translate to poor performance. The difference between L1 and L2 is that in L1, the added value is proportional to the absolute value of the weight coefficients, while in L2 it is proportional to the square of the coefficients (Chollet and Allaire, 2018). In both instances, high weight values will increase the value of the loss function and, in effect, force the network to use lower weight values during its training process. The dropout regularization differs from the L1 and L2 regularization since it does not modify the loss function of the network. Instead, it randomly assigns a value of zero to connections received from previous layers (Agarwal and Mittal, 2019) instead of the value that the activation function would normally produce. Where the L1 and L2 regularizations adjust

the weights in the network indirectly through the loss function, the dropout regularization adjusts the weights in the network by assigning zero to randomly selected connections thereby "forcing" the network to update the weights on the unaffected connections. This is because a low output of the loss function indicates better agreement between the predictions of the model and the real world examples used.

Two popular networks that are frequently used in image classification tasks are VGG16 and ResNet50. VGG16 is a type of deep neural network consisting of five 'blocks' with each containing between two and three convolution layers (Pradhan *et al.*, 2020; Mahdianpari *et al.*, 2018). This model was originally trained on the ImageNet dataset (Chevtchenko *et al.*, 2018). ResNet50 has a much deeper (i.e. more layers) architecture than VGG16. It is a residual network (Bendjillali *et al.*, 2020; Pradhan *et al.*, 2020) that makes use of residual learning, in the form of "skip-connections" (Khan *et al.*, 2020), to overcome the vanishing gradient problem (Mahdianpari *et al.*, 2018; Nagle-Mcnaughton *et al.*, 2020). Both VGG16 and ResNet50 require the input dimensions to be  $224 \times 224$  (Bendjillali *et al.*, 2020).

Performance-wise ResNet50 is considered easier to train, from a system resources perspective, than VGG16 (Oo and Oo, 2019). However, studies regarding the predictive accuracies between the models are divided. In some studies (e.g. Agarwal and Mittal, 2019; Sun *et al.*, 2019) VGG16 outperforms ResNet50 while a larger number of studies found that ResNet50 performs better (Mahdianpari *et al.*, 2018; Ji *et al.*, 2020; Wilhelm *et al.*, 2020).

### 3.2. Performance Determinants

There are three broad determinants of CNN model performance. First, are the learnable parameters which include the weights, biases, and the kernels used in the CNN. These are the parameters that are updated during the training process of the model. The weights within a model are analogous to the strength of synaptic connection in biological neural networks (Rodvold *et al.*, 2001) and can be viewed as storing the processing capacity of the model (Gurney, 2010). When the weight is associated with the transfer of information between different layers of the model, it serves as a means by which the contribution of the information to the classification is expressed. High weights mean that the information coming from that neuron contributes greatly to the ultimate classification while low weights indicate the opposite. When associated with kernels, the weights determine the "pattern" that the filter is testing for during the convolution operation. For CNNs, this latter aspect is instrumental in the CNN "learning" which features to extract to more effectively classify the input image. Although these learnable parameters are ultimately the deciding factor in determining the ability of the model to classify images, the user has no direct control over them. They are updated and modified during the training process as part of back-propagation.

However, the user can indirectly influence these learnable parameters through the modification of the model's hyperparameters. The hyperparameters are the

second determinant of model performance. The following hyperparameters influence of the performance of the model. This is by no means an exhaustive list but represents the hyperparameters considered in this study. The performance of a CNN is influenced by the depth of the model (Ghorbanzadeh *et al.*, 2019) with increasing depth associated with increased accuracy (Shakya *et al.*, 2021; Simonyan and Zisserman, 2015; Wan *et al.*, 2018; Ciresan *et al.*, 2011). Increasing the depth of the model, through the addition of more convolutional layers, allows the model to extract more abstract features (Kattenborn *et al.*, 2021). Here, abstract features refer to features that are derived from pre-existing feature maps. For example, filters located in the first convolution layer may extract features such as horizontal lines ( $-$ ), diagonally upwards lines ( $/$ ), and diagonally downwards lines ( $\backslash$ ) from an input image. The resultant feature map would therefore indicate the extent to which these features are present in the input image. The next convolution will take the feature map as an input and may recognise the combined presence of these three features as a triangle ( $\Delta$ ). However, it is also possible for an increase in model depth to lead to poorer performance (He and Sun, 2015; He *et al.*, 2016; Du *et al.*, 2021). It can therefore not be assumed that the addition of more layers will guarantee a performance improvement.

Another contributor to improved model performance is the use of regularizations and dropout layers. A regularization is any modification made to a learning program that attempts to reduce the generalization error (Goodfellow *et al.*, 2016). Here, generalization refers to the process whereby the model finds rules that are consistent with the available data (i.e. the data on which it is trained) that apply to instances that the model has not yet encountered (Zhang *et al.*, 2017). Essentially, regularizations are used to reduce overfitting (Thakkar and Lohiya, 2021; Tombe and Viriri, 2020; Lancashire *et al.*, 2009). A model that is suffering from overfitting cannot, by definition, generalize well. Therefore, by countering the tendency of the model to overfit during training, it is more likely that the model will be able to generalize to new unseen cases. Two types of regularization are commonly used:  $L_1$  regularization, and  $L_2$  regularization. Both involve mathematical terms that are added to the loss function during the training process. In the case of  $L_1$ , the additional term is proportional to the absolute value of the weight coefficients while  $L_2$  is proportional to the square of the weight coefficients (Chollet and Allaire, 2018). This means that the additional penalty term for  $L_1$  regularization is larger than for  $L_2$  regularization. Additionally, the weight of the regularization, i.e. the contribution it makes as part of the loss function, can also be adjusted by setting the regularization parameter (Thakkar and Lohiya, 2021; Tombe and Viriri, 2020; Lancashire *et al.*, 2009) and to reduce model complexity (Rubanenko *et al.*, 2021; Chollet and Allaire, 2018; Zhang *et al.*, 2017).

Dropout layers assign a value of 0 to the output of a set of activation functions in the previous layer of a model (Agarwal and Mittal, 2019). This can also be used as a means to counter model overfitting (Shakya *et al.*, 2021; Chevtchenko *et al.*, 2018) thereby increasing generalization. Aside from the addition of these dropout layers to the network, an additional hyperparame-

ter that can be set is the dropout value which determines the fraction of the neurons, from the previous layer, that would have an output of 0.

The last determinant of model performance is the data on which the model is trained. One of the main constraints of developing CNNs "from scratch" is their performance is dependent on the number of images on which they are trained. When limited data is available, the CNN is susceptible to overfitting (Mahdianpari *et al.*, 2018) which will negatively influence its performance. It is, however, possible to overcome this problem to some extent through the use of data augmentation (e.g. Gu *et al.*, 2018; Huang *et al.*, 2018; Carbonneau *et al.*, 2020). During data augmentation, the data is transformed without altering the nature of the image (Gu *et al.*, 2018). This is achieved by applying geometric transformations such as horizontal and vertical flips, rotations, translations, and scaling (Ji *et al.*, 2020; Nagle-Mcnaughton *et al.*, 2020). Alternatively, in cases where training data is limited, it is also possible to use transfer learning strategies (Azzaoui *et al.*, 2019; Bhuiyan *et al.*, 2020). In transfer learning, the learnable parameters of the model are derived through exposure to a primary task (which can include a completely different image data set). Following successful training, the model is then applied to a secondary task (i.e. the one where the data is limited).

## 4. Methods

### 4.1. Study site sample selection

Ninety barchan dunes were sampled from the Kunene region in northern Namibia using Google Earth<sup>TM</sup> imagery (taken 12 November 2012). The selection of the imagery date has no impact on the analysis and was merely selected due to the high resolution of the imagery. Separate images were collected for each barchan within the dune field due to the varying size of the dunes. This ensured that the highest possible resolution was used for each barchan thereby minimizing interpretation errors in outline delineation. The region experiences southerly winds with a resultant drift potential of 629.2, a resultant drift direction of 352°, and the terrain slopes gently from the interior of the region towards the coast (van der Merwe, 2021). Since boundary delineation of merged barchans can be highly subjective (Hugenholtz *et al.*, 2012), only isolated barchans were selected.

Google Earth imagery is frequently used in dune research (Bourke and Goudie, 2009; Zhang *et al.*, 2018) due to it being freely available. It is also used in research on the application of CNNs to remote sensing imagery (Wang *et al.*, 2019; Gao *et al.*, 2021; Abolt and Young, 2020). However, there are some limitations to using this platform. High-resolution imagery of the globe is not available at all locations (Tooth, 2013). This imposes restrictions on the types of landforms that can be studied. However, given the wide global distribution of barchans (Goudie, 2020), imagery of barchans are available. The positional accuracy of Google Earth imagery is also quite variable between different regions (Benker *et al.*, 2011). Since this study is concerned with the classification

of individual dunes, the exact position of the sampled dunes on the globe is not relevant and the imagery is considered suitable for use.

#### 4.2. Description of Barchans

The barchans in the Kunene region have an average width of  $46.6m \pm 28.1$  and an average stoss length of  $38.8m \pm 20.04$ . This makes the dunes significantly smaller ( $p < 0.01$ ) than those measured by Barnes (2001) in Walvis Bay further to the south. These barchans are also small compared to other barchans measured in other localities (Al-Harhi, 2002; Bailey, 1906; Barnes, 2001; Dong *et al.*, 2000; Douglass, 1909; Embabi, 1982; Finkel, 1959; Hamdan *et al.*, 2016; Jimenez *et al.*, 1999; Long and Sharp, 1964; Norris, 1966; Rempel, 1936; Sagga, 1998; Wang *et al.*, 2007). The size distribution of the barchans within the Kunene dune field is also positively skewed indicating that smaller dunes dominate the data set. Although barchan dunes do show allometry (van der Merwe, 2021), this is not expected to impact the number of dunes in each class since the median value is used in the definition of each class. The majority of dunes ( $n = 76$ ) exhibit an elongation of the left horn while only a few ( $n = 14$ ) have the right horn elongated. This shows a clear preference for left horn elongation within the data set and contrasts with the findings of (Barnes, 2001) who found no preferential elongation further towards the south in Walvis Bay. In terms of the  $\frac{a}{c}$  ratio, for 43 barchans  $\frac{a}{c} \leq 0.88$ , for 34 dunes  $0.63 \leq \frac{a}{c} < 0.88$ , and for 16 dunes  $0.38 \leq \frac{a}{c} < 0.63$ .

#### 4.3. Classification tasks

Four binary classification tasks (i.e. tasks where there were only two classification categories are present) were created that represent incremental increases in difficulty. The difficulty levels are, admittedly, subjective and based on the amount of analysis that needs to be undertaken for a human to complete. Task 1 requires the model to classify the dunes based on which side of the dune, left or right, is elongated. For Task 2, the network has to classify dunes based on the magnitude of difference between the length of the left and right horn. Task 3 requires the model to classify dunes based on their  $\frac{a}{c}$  ratio. Lastly, for Task 4, barchans need to be classified based on the magnitude of their bilateral asymmetry. The requirements of the tasks are graphically illustrated in figure 5.

Tasks 1 and 2 can be completed through visual inspection only and as such is considered to be the simplest of the four. For Task 1 only two "pieces" of information are used, an image of the left side and one of the right. Similarly, the second task compares the one horn with the other but requires the additional task of ranking all images in the set based on the magnitude of difference. Therefore, Task 1 is considered here to be simpler than Task 2. Tasks 3 and 4 cannot be easily solved using only visual inspection and require some calculations to be completed. For Task 3 it is simple division while completion of Task 4 requires matrices and matrix transformations to solve. It uses landmarks that are located along the barchan's boundary (van der Merwe, 2021) to compute

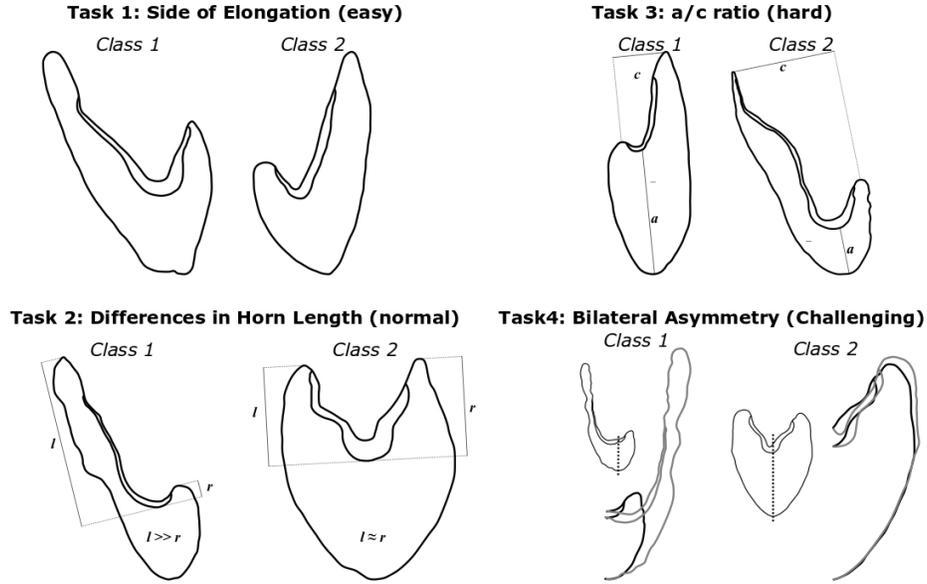


Figure 5: A graphical summary of the classification tasks. The concept of bilateral asymmetry is illustrated by superimposing the right half of the dune (grey) onto the left half (black) along the longitudinal axis of the barchan.

the procrustes distance between the dune and its mirror image and uses this as an indication of asymmetry (Klingenberg and McIntyre, 1998). Therefore, Task 3 is considered to be simpler than Task 4.

#### 4.4. Data preparation

The performance of a CNN is dependent on the training data that is used (Guidotti *et al.*, 2018; Maggiori *et al.*, 2017). More imagery is better since it provides the network with more diverse instances of the subject leading to, potentially, rules that are more generalized. However, it is not always possible to obtain large datasets of images. When limited imagery is available, image augmentation can be used to effectively increase the size of the training dataset (Castelluccio *et al.*, 2015; Jiang *et al.*, 2018; Zhu *et al.*, 2017; Carbonneau *et al.*, 2020; Huang *et al.*, 2018). Image augmentation involves changing the orientation of the input image through rotations, flipping, scaling and translations (Maxwell *et al.*, 2020; Nagle-Mcnaughton *et al.*, 2020) without altering the data within the image (Gu *et al.*, 2018). This effectively increases the diversity of imagery upon which the network is trained (Gao *et al.*, 2021). For the dunes in this data set, only the flipping augmentation, which was carried out along the  $x$ -axis, was used which generated mirror images and doubled the size of the training data.

Because CNNs are strongly biased towards detecting textures in imagery (Geirhos *et al.*, 2019; Nguyen *et al.*, 2015), it was necessary to remove this from the imagery. Earlier work has been successful in extracting the barchan boundary from remote sensing imagery (Rubanenko *et al.*, 2021; Azzaoui *et al.*, 2020;

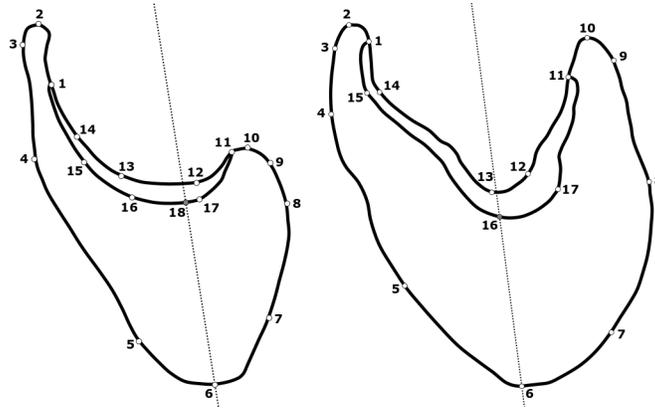


Figure 6: Landmarks used to derive barchan morphologic descriptions. The dotted line represents the longitudinal axis. The dune on the right represents a scenario where 18 landmarks were used while the dune on the left has 17 landmarks.

Vaz *et al.*, 2015) using semantic segmentation. Unfortunately, the brink line is not delineated in these cases as it occurs inside the polygon that defines the entire barchan boundary. Given the importance of the brink line in describing the morphology of the barchan (Fig. 2) this is a significant shortcoming.

Because of this limitation, the barchan outlines were extracted manually from cropped satellite imagery using Inkscape (<https://inkscape.org/>). This involved tracing along the boundaries of the dune in the image. Because Inkscape allows for the creation of scalable vector graphics (SVG) it is possible to scale the outlines to the dimensions required by the neural network without any loss of resolution. Once appropriately scaled, images were transferred to GIMP (GNU Image Manipulation Program - <https://www.gimp.org/>) to create square images of  $224 \times 224$  pixels with the dune being centered. This scaling does not form part of the augmentation process but is required to match the input dimensions required by the CNN. Although the images are black and white, three channels were retained to accommodate the input requirements for the neural networks.

Before training the network, barchan morphology data had to be extracted from the images to create the appropriate training, testing and validation sets. The geometric morphometric technique method used by van der Merwe (2021) was used to assign landmarks to the barchan outline. Seventeen to 18 landmarks were identified for each dune (Fig 6). The 18<sup>th</sup> landmark was only present in some imagery and serves as a reference point for the intersection between the longitudinal axis of the dune and the crest, which is needed for the calculation of the  $\frac{a}{c}$  ratio. In cases where this point coincides with one of the 17 points identified by van der Merwe (2021), only 17 landmarks were used.

Using these landmarks, the relevant morphometric variables can be calculated using coordinate geometry. The total width of the dune was calculated following Sauermann *et al.* (2000) where the total length is the sum of the distance from the horn tip to the longitudinal axis for the left and right horns

respectively. To obtain this value, two triangles were constructed with their corners at the toe, tips of the horns and peak (Fig. 7). Heron’s formula (Eq. 1), where  $s$  is the half-perimeter and  $a$ ,  $b$ , and  $c$  represent the individual sides, can then be used to calculate the surface area of the triangle. This is then used to calculate the height of each triangle ( $\overline{AF}$  and  $\overline{CE}$ ), where  $A$  represents the area of the triangle,  $h$  the height, and  $b$  the base. In this instance, the height of the individual triangles represents the width of each horn and their sum, therefore, represents the total width of the dune. The length of each horn was also calculated using triangles with the brink line as the reference point (following Wang *et al.* (2007)). Since  $\overline{AF}$  and  $\overline{CE}$  were calculated in the previous step, the Pythagorean theorem can be used to obtain  $\overline{DF}$  and  $\overline{DE}$  which represents the lengths of the left and right horn respectively.

$$A = \sqrt{s(s-a)(s-b)(s-c)} \quad (1)$$

$$h = \frac{2A}{b} \quad (2)$$

This data was then used to classify the dunes into the various classes needed for each task (Table 2). The bilateral asymmetry of the barchan cannot be calculated using the distance data obtained in Figure 7. This value was calculated using geometric morphometrics. The mirror image of the dune is used as a reference shape and the procrustes distance is determined between the dune and its mirror image (van der Merwe, 2021). This technique has been used in an earlier study to determine the asymmetry of *tsetse* fly wings (Klingenberg and McIntyre, 1998). Once this was calculated for all dunes, the median was used as the separation criterion (Table 2) and mirror images were created within each class to increase the available training data. To split the data into the appropriate categories the following approach was followed. For the first task, the data was simply split according to which side contained the elongated horn. For the remaining tasks, the median value was used to separate the data into two groups for training and testing. This was done to ensure the maximum amount of images within each category which, in turn, overcomes the problem of class imbalance raised by Nagle-Mcnaughton *et al.* (2020).

Task	Class 1	Class 2
Side of elongation	$\overline{FD} > \overline{ED}$	$\overline{FD} > \overline{ED}$
The difference in horn lengths	$ \overline{FD} - \overline{ED}  > \text{median}$	$ \overline{FD} - \overline{ED}  < \text{median}$
$\frac{a}{c}$ ratio	$\frac{\overline{DB}}{\overline{AF+CE}} > \text{median}$	$\frac{\overline{DB}}{\overline{AF+CE}} < \text{median}$
Bilateral asymmetry	$\rho > \text{median}$	$\rho \leq \text{median}$

Table 2: The criteria that were used to assign individual dunes into different classes. The procrustes distance is represented by  $\rho$ .

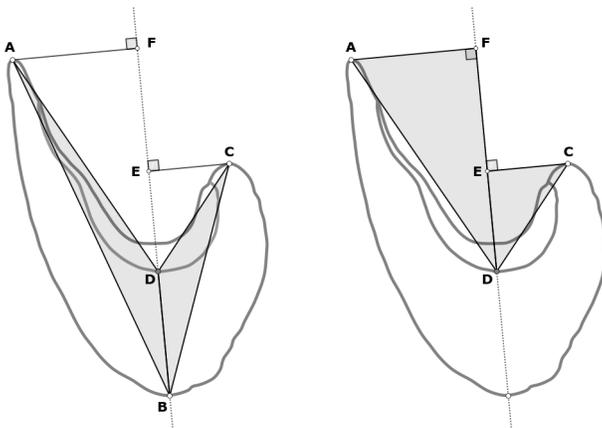


Figure 7: The approach used to calculate barchan width and the lengths of the horns.

#### 4.5. Models

Because the project aims to evaluate the feasibility of using neural networks to classify barchans, two scenarios were considered: a "from scratch" scenario where the model is built and trained by a user; and a transfer learning strategy. In the former, the performance of the model is dependent on both the training data and the hyperparameters. Unfortunately, the number of permutations associated with modifying all of the hyperparameters is too large to be practically feasible. Therefore, it was decided to consider only the following hyperparameters: depth, regularizations, and dropout layers. Because there is no data on which combination of these is best suited to the task, a trial and error approach was followed (Thakkar and Lohiya, 2021; Zeiler and Fergus, 2014; Tajbakhsh *et al.*, 2016). Models were created where the depth ranged from 1 to 3 layers and for each of these models one model had the  $L_1$  regularization, one had the  $L_2$  regularization, and one used dropout layers (Figure 8). This resulted in 12 models for the "from scratch" approach. All architectures were created in Python using the Tensorflow library.

In the second scenario, a prospective user is making use of a pre-trained network. As stated earlier, this approach may be of value in cases where the training data is limited. For this scenario, the VGG16 and ResNet50 models were chosen. This choice is based on the popularity of these two models within the image classification literature (Diaz-Pinto *et al.*, 2019; Oo and Oo, 2019; Ilyas *et al.*, 2019; Bendjillali *et al.*, 2020; Mukti and Biswas, 2019; Sun *et al.*, 2019; Agarwal and Mittal, 2019; LaVeZZi *et al.*, 2020; Raghu *et al.*, 2020; Mahdianpari *et al.*, 2018; Pradhan and Lee, 2010). Because the number of output neurons for each of these models is more than that required for this study, the output layers (i.e. the image classifier) were removed from the original and a new output layer was added using two neurons and the softmax activation function. This serves to retain the feature extractor portion of the network, which was trained on a significantly larger data set while replacing the classifier with one suited

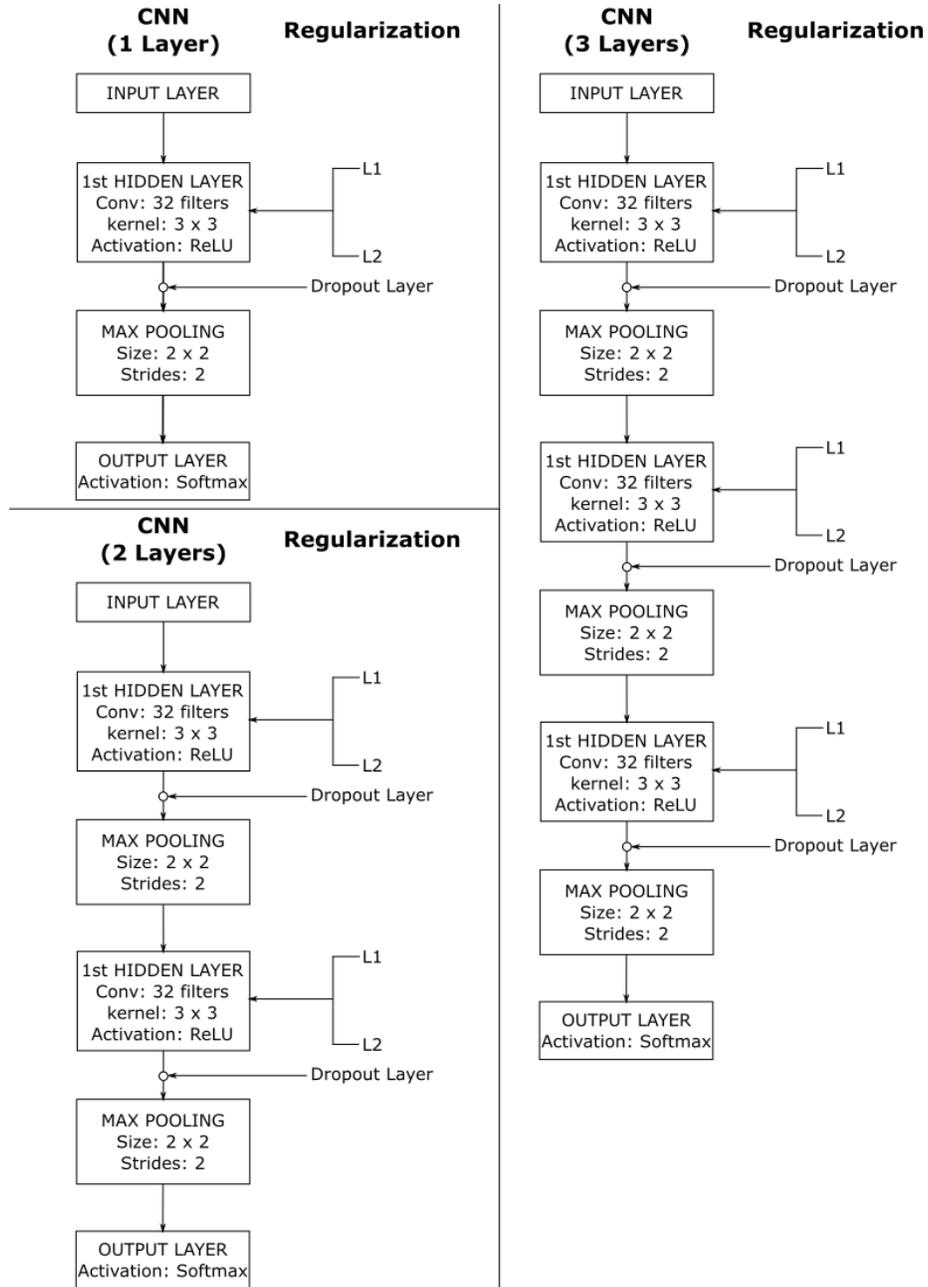


Figure 8: The three CNN architectures, along with the regularizations, that were used in this study.

to the task. A similar approach was followed by [Chevtchenko et al. \(2018\)](#) and [LaVezzi et al. \(2020\)](#).

#### 4.6. Training and testing

All of the training and testing were carried out following the procedure laid out in Pseudocode 1. By including the model construction, compilation, validation and testing within a for loop, it is guaranteed that any weights resulting from a previous training cycle is removed. This results in a more realistic assessment of the architecture’s performance. Both training and prediction data were exported as .csv files for analysis in RStudio.

---

**Pseudocode 1:** Training and testing neural networks ( $n = 30$ )

---

**Result:** Train and test network  $n$  times  
 Set  $n$  equal to the desired number of runs;  
**for**  $i$  in range 1 to  $n$  **do**  
     build model;  
     compile model;  
     train model;  
     store train data in variable  $train.i$ ;  
     write  $train.i$  to disc as  $train.i.csv$ ;  
     assign test batch labels to variable  $labels.i$ ;  
     test model;  
     store predictions in variable  $pred.i$ ;  
     append  $pred.i$  with  $labels.i$ ;  
     write  $pred.i$  to disc as  $pred.i.csv$ ;  
**end**

---

For tasks 1 to 3, the distribution per class for training, validation and testing were 56, 24 and 10 respectively. These samples were randomly selected from the appropriate image banks. Task 4 had 48 images in class one and 64 in class two. The reason for the discrepancy was the presence of several dunes that had a procrustes distance equal to the median value. Since no justifiable reason can be used to assign dunes of equal magnitudes of bilateral asymmetry to different classes, they were all incorporated into one of the classes. This does, unfortunately, result in a slight imbalance in sizes that are skewed in the favour of symmetrical dunes. During the training process, the training parameters of all the architectures, including those that had regularizations, were kept identical across all of the tasks. Since the number of images was small, a batch size of two was used along with 10 epochs of training.

#### 4.7. Model evaluation

Before the evaluation, the prediction data for each model was converted into a confusion matrix (Eq. 3). A confusion matrix represents the performance of the model as an  $m \times m$  matrix where  $m$  is the number of classes used in the model. It shows the extent of agreement and disagreement between the actual

category of the image, and the assigned category that the model provided. In the case of a binary classification scheme, cases where the model correctly classified the image can be either true positive (TP) in the case of category one or true negative (TN) in the case of category two. When there is a disagreement between the model’s classification and the actual classification, the result is returned as either a false positive (FP) or a false negative (FN) depending on the nature of the discrepancy.

$$\begin{matrix} & \textit{Model}_{C1} & \textit{Model}_{C2} \\ \textit{Actual}_{C1} & \left[ \begin{matrix} TP & FP \end{matrix} \right] \\ \textit{Actual}_{C2} & \left[ \begin{matrix} FN & TN \end{matrix} \right] \end{matrix} \quad (3)$$

Three parameters were chosen for model evaluation: accuracy (ACC),  $F_1$ -score and Matthew’s Correlation Coefficient (MCC). Accuracy is a common metric which expresses the number of predictions that correctly identified both positive and negative instances (Brown, 2018). This metric is frequently used in binary classification tasks (Hossin and Sulaiman, 2015). Here, a high accuracy score represents a situation where the number of images that were correctly classified as belonging to either class is high relative to the number of predictions made by the model. The equation for ACC, following (Hossin and Sulaiman, 2015), is given in equation 4.

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

The  $F_1$ -score (Eq. 5) represents the harmonic mean between recall (TPR) and precision (PPV) (Hossin and Sulaiman, 2015). TPR is the proportion between the TP predictions and all the positive predictions made by the model (Bellows *et al.*, 2011). Precision, also known as the positive prediction value (PPV) (Bellows *et al.*, 2011), is the number of positive predictions made by the model that are true positives (Hossin and Sulaiman, 2015). The equation for the  $F_1$  score, following (Brown, 2018), is given in equation 5. In the surface plot of the potential  $F_1$ -scores (Fig 9), both the TPR and the PPV scores of the network need to be high. This means that a high  $F_1$ -score suggests that the model had a large portion of true positive predictions and low amounts of false positives and false negatives. Unfortunately, the surface also suggests that it is not possible to determine from the  $F_1$ -score alone where shortcomings occur. In other words, it can either be from a high false positive rate, a high false negative rate or a combination of both. Notwithstanding this, the  $F_1$ -score is still widely used in machine learning (Chicco and Jurman, 2020).

$$F1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} \quad (5)$$

Both the ACC and the MCC (Eq. 6) use the same variables (Brown, 2018; Hossin and Sulaiman, 2015), but the MCC metric is considered to be more reliable (Chicco and Jurman, 2020). This is because it penalises Type-I (false positive) and Type-II (false negative) errors more than the ACC metric does

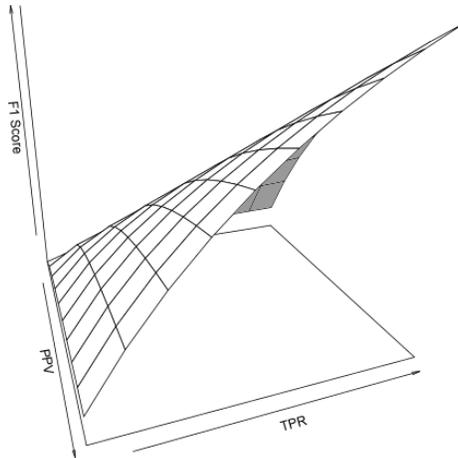


Figure 9: The value surface of the F1 metric. In this example, PPV values are plotted along the  $x$ -axis, TPR values along the  $y$ -axis and the resultant F1 score along the  $z$ -axis.

(Brown, 2018). Essentially, it serves as an indication of how well the actual data correlates with the predicted data with values of +1, zero and -1 indicating perfect agreement, random predictions and total disagreement respectively (Tharwat, 2018). The equation for the MCC metric is given in equation 6 (Brown, 2018). A problem with the MCC metric is that, under certain conditions, it cannot be defined due to imbalances in the confusion matrix (Chicco and Jurman, 2020). This happens in cases where both TP and FN are equal to zero causing the denominator to become zero.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (6)$$

## 5. Results

The baseline model (CNN-1 [-]) had an overall ACC value of more than 70% on all the tasks. The mean ACC value decreased as the task difficulty increased. Significant differences in ACC performance were found between the following task pairs: Task 1 and Task 3 ( $p = 0.002$ ); Task 1 and Task 4 ( $p = 0.002$ ); Task 2 and Task 3 ( $p = 0.02$ ); and between Tasks 2 and 4 ( $p = 0.016$ ). Similar to the ACC values, the  $F_1$ -score decreased as the as the task difficulty increased. However, significant differences were only present between Tasks 2 and 3 ( $p = 0.008$ ), and between Tasks 2 and 4 ( $p = 0.006$ ). The MCC values also decreased as the task difficulty increased and a significant difference ( $p = 0.032$ ) was only found between Tasks 1 and 4.

For the majority of cases, there was no significant difference in the metrics as the depth of the models were increased. When the depth of the model is

Metric	Task 1	Task 2	Task 3	Task 4
ACC	78.667	77.333	73.00	72.167
$F_1$ -Score	75.925	75.179	72.52	68.733
MCC	0.653	0.63	0.604	0.575

Table 3: Mean performance on the different metrics for the CNN-1 [-] model.

Metric	Depth	Task 1	Task 2	Task 3	Task 4
ACC	2 Layers	0.333	4.500	-1.333	2.000
	3 Layers	-2.167	4.333	0	1.167
$F_1$ -Score	2 Layers	0.985	6.253	-2.046*	4.780
	3 Layers	-0.142	6.442	-0.423	5.692*
MCC	2 Layers	0.014	0.073	-0.017	0.035
	3 layers	-0.017	0.073	-0.009	0.038

Table 4: Influence of model depth (i.e. CNN-2 [-] and CNN-3 [-]) on performance as deviations from the baseline. Values marked with an asterisk (\*) indicate significant differences compared with the baseline performance.

increased to two layers, the only significant difference ( $p = 0.05$ ) that was found involved a decrease in the  $F_1$ -score for Task 3. Increasing the depth to three layers resulted in a significant improvement ( $p = 0.05$ ) of the  $F_1$ -score for task 4.

The addition of regularizations (Table 5) did not significantly influence the performance of the model on Task 1. Also, the dropout regularization had no significant influence on any of the metrics for any of the four tasks. The L1 regularization improved ( $p = 0.05$ ) the performance of the  $F_1$ -score for Task 4. Most changes in model performance were associated with the incorporation of the L2 regularization. This led to significant improvements in the  $F_1$ -score for Tasks 2 and 4 ( $p = 0.004$  and  $p = 0.02$  respectively). It also led to a decrease in the ACC and MCC values for Task 2 ( $p = 0.01$  and  $p = 0.009$  respectively).

Metric	Reg	Task 1	Task 2	Task 3	Task 4
ACC	L1	2.667	-1.667	-2.667	2.333
	L2	-1.167	-0.667*	0.167	2.500
	Dropout	-0.167	1.000	0.833	0.500
$F_1$ -Score	L1	3.060	-0.642	-1.522	3.674*
	L2	-2.277	2.696*	1.582	7.934*
	Dropout	0.329	1.703	0.078	2.364
MCC	L1	0.037	-0.022	-0.025	0.033
	L2	-0.013	-0.002	0.006	0.068
	Dropout	0.014	0.021	0.003	0.002

Table 5: Changes in model performance from baseline conditions by adding regularizations (i.e models CNN-1 [L1], [L2] and [D]). Values marked with an asterisk (\*) indicate significant differences compared with the baseline performance.

Metric	Depth	Reg.	Task 1	Task 2	Task 3	Task 4
ACC	2	L1	1.333	5.833*	-0.667	3.000
		L2	1.167	4.333	0	2.333
		D	3.333	3.833	-0.833	3.667*
	3	L1	2.000	2.333	2.833*	1.000
		L2	-0.833	5.000*	-1.167	0.167
		D	-3.000	6.667**	0.833	-1.167
$F_1$ -Score	2	L1	2.930	8.582**	-1.090	6.560*
		L2	2.416	6.390	-0.407	3.158
		D	4.639	5.418	-1.633	4.360*
	3	L1	4.105	2.562	1.256	2.723
		L2	0.629	6.886*	-1.792	0.441
		D	-1.805	8.569**	-0.784	1.300
MCC	2	L1	0.022	0.090**	-0.008	0.046
		L2	0.025	0.068	0.003	0.031
		D	0.054	0.061	-0.008	0.050**
	3	L1	0.049	0.046	0.024*	0.029
		L2	0.002	0.084*	-0.030	0.010
		D	-0.033	0.102**	0.001	0.007

Table 6: The influence of a combination of depth and regularization on model performance. For brevity, only combinations that involve significant differences are shown. The magnitude of the significant difference is indicated by asterisks (\* =  $p \leq 0.05$ , \*\* =  $p \leq 0.001$ ).

As to be expected, more significant differences emerged when the model depth is combined with regularizations (Table 6). These differences were only present for Tasks 2 to 4 and no significant differences were found for Task 1. In contrast, the majority of significant changes occurred for Task 2. Combining a three-layer model with either the L2 or dropout regularization led to significant improvements to the ACC value for Task 2. When a two-layer model is combined with the L1 regularization a significant improvement in ACC values is also found. These configurations were also associated with improvements to the  $F_1$ -score and MCC values for this task.

The performance of Task 3 was only improved by combining an L1 regularization with a three-layer network and this only improved the ACC and MCC values. By combining a two-layer network with the dropout regularizer, the ACC,  $F_1$ -score and MCC values for Task 4 were significantly improved. Also, the use of a two-layer network with L1 regularization improved only the  $F_1$ -score of this task.

Using transfer learning resulted in significant improvements for all metrics on all tasks except for Task 2 (Table 7) where the pre-trained models did not perform significantly better than the baseline model. The magnitude of improvement was also greater for the complex tasks (Tasks 3 and 4) than for the simpler tasks (Task 1 and 2).

Metric	Model	Task 1	Task 2	Task 3	Task 4
ACC	VGG16	3.500*	2.333	11.167***	11.000***
	ResNet50	4.667*	4.833	14.833***	15.500***
$F_1$ -Score	VGG16	6.347*	3.129	10.410***	12.932***
	ResNet50	7.843**	5.240	15.151***	17.268***
MCC	VGG16	0.057*	0.048	0.131***	0.145***
	ResNet50	0.069*	0.076	0.183***	0.207***

Table 7: Changes in performance when the pre-trained networks are used. The magnitude of the significant difference is indicated by asterisks (\* =  $p \leq 0.05$ , \*\* =  $p \leq 0.001$ , \*\*\* =  $p \leq 0.0001$ ).

## 6. Discussion

Determining whether or not a model performs satisfactorily involves a subjective decision based on the performance indicators. In the case of human interpreters, several studies have commented on the level of variability that exists between individuals when carrying out image classification tasks (Van Coillie *et al.*, 2014; Pengra *et al.*, 2020; Lloyd *et al.*, 2002). The study by Lloyd *et al.* (2002) found that the accuracy of human operators on a classification task can range from 69.6% (for specific within category distinctions) to 90.5% for more general category classifications. Although these tasks are different from the classification tasks carried out here, it does indicate that even when trained on a small data set, CNNs have accuracies that are comparable to humans which agrees with De Cesarei *et al.* (2021).

Out of the 14 model configurations tested (including the baseline), 78.5% of the CNNs achieve high accuracy levels (i.e an ACC and  $F_1$ -score greater than 80% and an MCC score greater than or equal to 0.7). This in itself answers the main aim of this research and it is clear that CNNs have the potential to serve as useful tools to accurately classify barchan dunes. This is in agreement with De Cesarei *et al.* (2021) who view neural networks as being well suited to visual tasks. Even the baseline model, which does not match the high accuracy of some of the other models, still performs within the classification range of participants in the study by (Lloyd *et al.*, 2002).

However, the question regarding how the baseline performance can be improved is not as clear from the data generated. It is tempting to assume that increasing the depth of the network will improve performance. Deeper layers within a network can extract more features from the image (Kattenborn *et al.*, 2021; de Lima and Marfurt, 2020; Wang *et al.*, 2019) and are associated with performance improvements (Shakya *et al.*, 2021). The data in this study does not support such a generalization. An increase in the number of layers is only beneficial within the context of the task and may, in some cases, result in decreased performance. The former is to be expected. The output of the feature extractor is a feature map (Ghorbanzadeh *et al.*, 2019; Gu *et al.*, 2018; Traore *et al.*, 2018; Boulila *et al.*, 2021). These feature maps, in turn, include patterns such as corners, edges etc (Kattenborn *et al.*, 2021). This "collection" of fea-

tures is then passed to the fully-connected layer (Palafox *et al.*, 2017) where the input image is then classified into one of the provided categories. But, it may not always be prudent or necessary to extract more complex features from an image to make a classification. This may be the case for barchan dune outlines which consist of very simple feature elements.

The worsening performance of the increased depth, which in this case resulted in a decreased  $F_1$ -score, may be attributable to the removal of image texture elements. The convolution operation is, essentially, the dot product between the kernel (which forms part of the convolution layer) and the image (Agarwal *et al.*, 2010). In cases where the majority of the pixel values are 0, as is the case when outlines are used, the more complex transformations within the deeper layers may reach a level such that the similarity between images from different classes increases. This would result in misclassifications that, in turn, will negatively affect the  $F_1$ -score.

Regularizations attempt to reduce model complexity and overfitting (Chollet and Allaire, 2018; Thakkar and Lohiya, 2021). This should, in principle, increase the generalization capability of the network. This generalization should allow the network to make better predictions on data that it has not seen before (Zhang *et al.*, 2017). Within the context of this study, it would therefore be expected that the addition of regularizations should improve model performance. However, this was not universally the case. The L1 and L2 regularizations are added to the loss function of the network (Thakkar and Lohiya, 2021) and are proportional to the weights of the connection (Chollet and Allaire, 2018). It may be that in comparatively shallow networks such as the ones used in this study, this addition to the loss function may not contribute meaningfully to the performance of the network. This would account for the low portion of significant changes that were observed. Also, the barchans may be similar enough as to not warrant the need for further generalization within the network.

Dropout layers operate differently within a network than the L1 and L2 regularizations. Dropout layers do not act on the loss function. Instead, it randomly assigns a value of zero to connections between layers (Agarwal and Mittal, 2019). This affects the other connections between layers and in effect serves the role of preventing the network from becoming over-reliant on the values of a few select nodes. Stated differently, it prevents the network from depending on only a few features from the feature map in making the classification. However, these benefits did not materialize when applied to barchan classification and no significant differences were observed. This could, potentially, indicate that the network’s weights are already configured in a generalized fashion so that any additional incentive to generalize will be negligible.

Although the isolated adjustments of network depth and regularization do not contribute to large scale improvements, the combined effect does lead to a larger portion of significant changes. These results indicate that the modification of a single hyperparameter may not be sufficient to improve the performance of a CNN. But by combining hyperparameters, more and exclusively positive changes are observed. Significant changes are still not always present which means that the influence of stochasticity on the part of the network initialization cannot

be excluded. Nevertheless, it appears that by combining hyperparameters, i.e. changing more than one hyperparameter, better performance can be obtained from the network. Unfortunately, because more hyperparameters need to be considered, it also means that there are more permutations to work through to boost performance. Fortunately, it is possible to automate this process using hyperparameter optimization techniques, such as grid search, to find the best configuration.

Perhaps unsurprisingly, the use of transfer learning led to the most improvement in the classification of barchans. It also conflicts to some extent with the interpretations made earlier. ResNet50 consists of 152 layers (Pradhan *et al.*, 2020) while VGG16 contains 13 convolution layers (Mahdianpari *et al.*, 2018). This makes them substantially larger than the custom fully trained networks created for this study. The ACC scores of both VGG16 and ResNet50 are higher in all cases than those of the custom networks, but it is interesting to note that for the second task the improved performance is not significant. Why this is the case is not known, but it does indicate that the use of transfer learning cannot be viewed as being the better option in all use cases. Between the two models, ResNet50 outperformed VGG16 on all metrics. This agrees with the findings of some authors (e.g. Wilhelm *et al.*, 2020; Ji *et al.*, 2020) and differs from those of others (e.g. Sun *et al.*, 2019; Agarwal and Mittal, 2019). This lack of universal agreement reaffirms the task-specific performance of CNNs.

Earlier work on CNNs concluded that they are biased towards the presence of image texture (Geirhos *et al.*, 2019; Baker *et al.*, 2020). In this study, the majority of textural information was removed from the input imagery. Given that image texture is used during convolution operations to generate feature maps, it should follow that limiting the textural information should increase the difficulty in classification for the model. However, the results indicate that such a view might be oversimplified since the models performed well with limited training data. Although the models developed here are far from being ready to be generally used within the aeolian geomorphology community, these initial results do suggest that this line of research has some practical benefits.

## 7. Conclusion

It can be concluded that CNNs do hold promise as a means to automatically extract morphological data from images of barchan outlines although it is still not possible to conclude absolute strategies for improving model performance. This work has shown that CNNs can potentially be used in place of manual methods to classify barchan dunes into morphometric classes. This, by implication, means that it is possible to extract morphometric data from barchan images using CNNs. But more research is still needed before this tool can be fine-tuned to the needs of the intended user base. Specifically, work is needed on the following.

First, the lack of improvement in performance with the addition of a dropout layer, which contradicts the findings of other researchers working on CNNs, may be the result of its position within the network. In the present configuration,

these layers are restricted to the feature extractor. This, in affect, alters the type of features present within the feature map upon which a decision is being made. The question then arises regarding the effects of shifting the position of the dropout layer to the image classifier. This will effect the contribution of the features within the feature map without removing them from the network.

Another area of research deals directly with the issues raised on training data. The effect of data set sizes is taken here as self-evident. However, a more useful approach would be to look at the amount of variability in images that make up the training data. Specifically, are there optimal values between class differences and within class similarities for network performance?

Lastly, an investigation of the activity maps of CNNs involved in barchan classification can be beneficial. First, it will assist in determining whether the CNN is making use of the barchan in making its decision instead of some other artefact within the image. And, second, it will help to identify the regions that the barchan uses to make its classification. This latter aspect will be particularly important in the context of the earlier discussion regarding the number of features needed for optimal network performance and may, hopefully, address how a network can have significant improvements in the  $F_1$ -score which does not translate to significant improvements in ACC and MCC values.

## **8. Conflict of interest**

The authors have no conflicts of interest that need to be declared.

## References

- Abdu, A. (1976) The crescentic dunes of northern Sudan. *East African Geographical Review*, 14(June), 61–71.
- Abolt, C.J. and Young, M.H. (2020) High-resolution mapping of spatial heterogeneity in ice wedge polygon geomorphology near Prudhoe Bay, Alaska. *Scientific Data*, 7(1). doi:10.1038/s41597-020-0423-9.
- Agarwal, M., Agrawal, H., Jain, N. and Kumar, M. (2010) Face recognition using principle component analysis, eigenface and neural network. *2010 International Conference on Signal Acquisition and Processing*, 310–314. doi:10.1109/ICSAP.2010.51.
- Agarwal, T. and Mittal, H. (2019) Performance Comparison of Deep Neural Networks on Image Datasets. In: *2019 Twelfth International Conference on Contemporary Computing (IC3)*. IEEE, pp. 1–6. doi:10.1109/IC3.2019.8844924.
- Al-Harhi, A.A. (2002) Geohazard assessment of sand dunes between Jeddah and Al-Lith, western Saudi Arabia. *Environmental Geology*, 42(4), 360–369. doi:10.1007/s00254-001-0501-z.
- Andreotti, B., Claudin, P. and Douady, S. (2002) Selection of dune shapes and velocities part 1: Dynamics of sand, wind and barchans. *European Physical Journal B*, 28(3), 321–339. doi:10.1140/epjb/e2002-00236-4.
- Aydda, A., Althuwaynee, O.F. and Pokharel, B. (2020) An easy method for barchan dunes automatic extraction from multispectral satellite data. *IOP Conference Series: Earth and Environmental Science*, 419(1). doi:10.1088/1755-1315/419/1/012015.
- Azzaoui, M.A., Adnani, M., El Belrhiti, H., Chaouki, I.E. and Masmoudi, L. (2019) Detection of crescent sand dunes contours in satellite images using an active shape model with a cascade classifier. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(4/W12), 17–24. doi:10.5194/isprs-archives-XLII-4-W12-17-2019.
- Azzaoui, M.A., Masmoudi, L., El Belrhiti, H. and Chaouki, I.E. (2020) Barchan sand dunes collisions detection in high resolution satellite images based on image clustering and transfer learning. *International Journal of Advanced Computer Science and Applications*, 11(1), 376–383. doi:10.14569/ijacsa.2020.0110147.
- Bagnold, R.A. (1954) *The physics of blown sand and desert dunes*. Mineola: Dover Publications.

- Bailey, S.I. (1906) The sand dunes of the Desert of Islay. In: *Annals of the Astronomical Observatory of Harvard College*, volume XXXIX. Observatory of Harvard College.
- Baker, N., Lu, H., Erlikhman, G. and Kellman, P.J. (2020) Local features and global shape information in object classification by deep convolutional neural networks. *Vision Research*, 172(March), 46–61. doi:10.1016/j.visres.2020.04.003.
- Barnes, J. (2001) Barchan dunes on the Kuiseb River Delta, Namibia. *South African Geographical Journal*, 83(3), 283–292. doi:10.1080/03736245.2001.9713747.
- Baumhoer, C.A., Dietz, A.J., Kneisel, C. and Kuenzer, C. (2019) Automated extraction of antarctic glacier and ice shelf fronts from Sentinel-1 imagery using deep learning. *Remote Sensing*, 11(21), 1–22. doi:10.3390/rs11212529.
- Bellows, B., Bhandari, A., Ibrahim, M. and Sandhu, J.S. (2011) Peering into the Black Box. *Information Communication Technologies*, (January), 38–43. doi:10.4018/9781599049496.ch067.
- Bendjillali, R.I., Beladgham, M., Merit, K. and Taleb-Ahmed, A. (2020) Illumination-robust face recognition based on deep convolutional neural networks architectures. *Indonesian Journal of Electrical Engineering and Computer Science*, 18(2), 1015–1027. doi:10.11591/ijeecs.v18.i2.pp1015-1027.
- Benker, S.C., Langford, R.P. and Pavlis, T.L. (2011) Positional accuracy of the Google Earth terrain model derived from stratigraphic unconformities in the Big Bend region, Texas, USA. *Geocarto International*, 26(4), 291–303. doi:10.1080/10106049.2011.568125.
- Bhuiyan, M.A.E., Witharana, C., Liljedahl, A.K., Jones, B.M., Daanen, R., Epstein, H.E., Kent, K., Griffin, C.G. and Agnew, A. (2020) Understanding the effects of optimal combination of spectral bands on deep learning model predictions: A case study based on permafrost tundra landform mapping using high resolution multispectral satellite imagery. *Journal of Imaging*, 6(9). doi:10.3390/JIMAGING6090097.
- Boulghobra, N. (2016) Climatic data and satellite imagery for assessing the aeolian sand deposit and barchan migration, as a major risk sources in the region of In-Salah (Central Algerian Sahara). *Arabian Journal of Geosciences*, 9(6), 450. doi:10.1007/s12517-016-2491-x.
- Boulghobra, N. and Dridi, H. (2016) Fine resolution imagery and gis for investigating the morphological characteristics, and migration rate of barchan dunes in the erg sidi moussa dunefield near in-salah (Algeria). *Geographia Technica*, 11(2), 14–21. doi:10.21163/GT\_2016.112.02.

- Boulila, W., Sellami, M., Driss, M., Al-Sarem, M., Safaei, M. and Ghaleb, F.A. (2021) RS-DCNN: A novel distributed convolutional-neural-networks based-approach for big remote-sensing image classification. *Computers and Electronics in Agriculture*, 182(January), 106014. doi:10.1016/j.compag.2021.106014.
- Bourke, M.C. (2010) Barchan dune asymmetry: Observations from Mars and Earth. *Icarus*, 205(1), 183–197. doi:10.1016/j.icarus.2009.08.023.
- Bourke, M.C., Balme, M., Beyer, R.A., Williams, K.K. and Zimbelman, J. (2006) A comparison of methods used to estimate the height of sand dunes on Mars. *Geomorphology*, 81(3-4), 440–452. doi:10.1016/j.geomorph.2006.04.023.
- Bourke, M.C. and Goudie, A.S. (2009) Varieties of barchan form in the Namib Desert and on Mars. *Aeolian Research*, 1(1-2), 45–54. doi:10.1016/j.aeolia.2009.05.002.
- Brown, J.B. (2018) Classifiers and their Metrics Quantified. *Molecular Informatics*, 37(1), 1–11. doi:10.1002/minf.201700127.
- Burrough, S.L., Thomas, D.S.G., Bailey, R.M. and Davies, L. (2012) From landform to process: Morphology and formation of lake-bed barchan dunes, Makgadikgadi, Botswana. *Geomorphology*, 161-162, 1–14. doi:10.1016/j.geomorph.2012.03.027.
- Buscombe, D. and Ritchie, A. (2018) Landscape classification with deep neural networks. *Geosciences*, 8(7), 244. doi:10.3390/geosciences8070244.
- Cady, F. (2017) *The Data Science Handbook*. Hoboken, New Jersey: John Wiley & Sons, Inc. doi:10.1002/9781119092919.
- Carbonneau, P.E., Belletti, B., Micotti, M., Lastoria, B., Casaioli, M., Mariani, S., Marchetti, G. and Bizzi, S. (2020) UAV-based training for fully fuzzy classification of Sentinel-2 fluvial scenes. *Earth Surface Processes and Landforms*, 45(13), 3120–3140. doi:10.1002/esp.4955.
- Carrara, F., Falchi, F., Caldelli, R., Amato, G. and Becarelli, R. (2018) Adversarial image detection in deep neural networks. *Multimedia Tools and Applications*, 1–21. doi:10.1007/s11042-018-5853-4.
- Castelluccio, M., Poggi, G., Sansone, C. and Verdoliva, L. (2015) Land Use Classification in Remote Sensing Images by Convolutional Neural Networks, 1–11. doi:10.1021/cr8002505.
- Chen, Z., Zhang, Y., Ouyang, C., Zhang, F. and Ma, J. (2018) Automated landslides detection for mountain cities using multi-temporal remote sensing imagery. *Sensors (Switzerland)*, 18(3). doi:10.3390/s18030821.
- Cheng, G., Yang, C., Yao, X., Guo, L. and Han, J. (2018) When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5), 2811–2821. doi:10.1109/TGRS.2017.2783902.

- Chevtchenko, S.F., Vale, R.F., Macario, V. and Cordeiro, F.R. (2018) A convolutional neural network with feature fusion for real-time hand posture recognition. *Applied Soft Computing Journal*, 73, 748–766. doi:10.1016/j.asoc.2018.09.010.
- Chicco, D. and Jurman, G. (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), 1–13. doi:10.1186/s12864-019-6413-7.
- Chojnacki, M., Burr, D.M., Moersch, J.E. and Michaels, T.I. (2011) Orbital observations of contemporary dune activity in Endeavor crater, Meridiani Planum, Mars. *Journal of Geophysical Research E: Planets*, 116(4), 1–20. doi:10.1029/2010JE003675.
- Chollet, F. and Allaire, J. (2018) *Deep learning with R*. Shelter Island: Manning.
- Ciresan, D., Meier, U., Masci, J. and Schmidhuber, J. (2011) A committee of neural networks for traffic sign classification. In: *The 2011 International Joint Conference on Neural Networks*, volume 1. IEEE, pp. 1918–1921. doi:10.1109/IJCNN.2011.6033458.
- Courrech du Pont, S. (2015) Dune morphodynamics. *Comptes Rendus Physique*, 16(1), 118–138. doi:http://dx.doi.org/10.1016/j.crhy.2015.02.002.
- Dang, K.B., Dang, V.B., Bui, Q.T., Nguyen, V.V., Pham, T.P.N. and Ngo, V.L. (2020) A Convolutional Neural Network for Coastal Classification Based on ALOS and NOAA Satellite Data. *IEEE Access*, 8, 11824–11839. doi:10.1109/ACCESS.2020.2965231.
- De Cesarei, A., Cavicchi, S., Cristadoro, G. and Lippi, M. (2021) Do Humans and Deep Convolutional Neural Networks Use Visual Information Similarly for the Categorization of Natural Scenes? *Cognitive Science*, 45(6). doi:10.1111/cogs.13009.
- de Lima, R.P. and Marfurt, K. (2020) Convolutional neural network for remote-sensing scene classification: Transfer learning analysis. *Remote Sensing*, 12(1). doi:10.3390/rs12010086.
- Diaz-Pinto, A., Morales, S., Naranjo, V., Köhler, T., Mossi, J.M. and Navea, A. (2019) CNNs for automatic glaucoma assessment using fundus images: An extensive validation. *BioMedical Engineering Online*, 18(1), 1–19. doi:10.1186/s12938-019-0649-y.
- Dong, Z., Wang, X. and Chen, G. (2000) Monitoring sand dune advance in the Taklimakan Desert. *Geomorphology*, 35(3-4), 219–231. doi:10.1016/S0169-555X(00)00039-8.
- Douglass, A.E. (1909) The crescentic dunes of Peru. *Appalachia*, 12(1), 34 – 45.

- Du, B., Zhao, Z., Hu, X., Wu, G., Han, L., Sun, L. and Gao, Q. (2021) Landslide susceptibility prediction based on image semantic segmentation. *Computers and Geosciences*, 155(November 2020), 104860. doi:10.1016/j.cageo.2021.104860.
- Durán, O., Parteli, E.J.R. and Herrmann, H.J. (2010) A continuous model for sand dunes: Review, new developments and application to barchan dunes and barchan dune fields. *Earth Surface Processes and Landforms*, 35(13), 1591–1600. doi:10.1002/esp.2070.
- El belrhiti, H. and Douady, S. (2011) Equilibrium versus disequilibrium of barchan dunes. *Geomorphology*, 125(4), 558–568. doi:10.1016/j.geomorph.2010.10.025.
- Elbelrhiti, H. (2012) Initiation and early development of barchan dunes: A case study of the Moroccan Atlantic Sahara desert. *Geomorphology*, 138(1), 181–188. doi:10.1016/j.geomorph.2011.08.033.
- Elbelrhiti, H., Andreotti, B. and Claudin, P. (2008) Barchan dune corridors: Field characterization and investigation of control parameters. *Journal of Geophysical Research: Earth Surface*, 113(2), 1–21. doi:10.1029/2007JF000767.
- Embabi, N.S. (1982) Barchans of the Kharga Depression. In: F. El-Baz and T.A. Maxwell (Eds.) *Desert Landforms of Southwest Egypt: A Basis for Comparison with Mars*, chapter 11. Washington, DC: NASA Scientific and Technical Information Branch, p. 372.
- Engel, M., Boesl, F. and Brückner, H. (2018) Migration of barchan dunes in qatar—controls of the shamal, teleconnections, sea-level changes and human impact. *Geosciences (Switzerland)*, 8(7). doi:10.3390/geosciences8070240.
- Finkel, H.J. (1959) The barchans of southern Peru. *The Journal of Geology*, 67(6), 614–647. doi:10.1086/626622.
- Franklin, E.M. and Charru, F. (2011) Subaqueous barchan dunes in turbulent shear flow. Part 1. Dune motion. *Journal of Fluid Mechanics*, 675, 199–222. doi:10.1017/S0022112011000139.
- Gafurov, A.M. and Yermolayev, O.P. (2020) Automatic Gully Detection: Neural Networks and Computer Vision. *Remote Sensing*, 12(11). doi:10.3390/rs12111743.
- Gao, B., Chen, N., Blaschke, T., Wu, C.Q., Chen, J., Xu, Y., Yang, X. and Du, Z. (2021) Automated characterization of yardangs using deep convolutional neural networks. *Remote Sensing*, 13(4), 1–19. doi:10.3390/rs13040733.
- Gay, S.P. (1999) Observations regarding the movement of barchan sand dunes in the Nazca to Tanaca area of southern Peru. *Geomorphology*, 27, 279–293. doi:10.1016/S0169-555X(98)00084-1.

- Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F.A. and Brendel, W. (2019) ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. In: *International Conference on Learning Representations*. pp. 1 – 22.
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D. and Aryal, J. (2019) Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2). doi:10.3390/rs11020196.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*. Cambridge, Massachusetts: MIT Press.
- Goudie, A.S. (2020) Global barchans: A distributional analysis. *Aeolian Research*, 44(February). doi:10.1016/j.aeolia.2020.100591.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. (2018) Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354–377. doi:https://doi.org/10.1016/j.patcog.2017.10.013.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D. and Giannotti, F. (2018) A survey of methods for explaining black box models. *arXiv*, 51(5).
- Gurney, K. (2010) Neural networks for perceptual processing: From simulation tools to theories. In: C. Tosh and G. Ruxton (Eds.) *Modelling Perception with Artificial Neural Networks*, chapter 1. Cambridge: Cambridge University Press, pp. 7–34. doi:10.1017/CBO9780511779145.002.
- Hamdan, M.A., Refaat, A.A. and Abdel Wahed, M. (2016) Morphologic characteristics and migration rate assessment of barchan dunes in the Southeastern Western Desert of Egypt. *Geomorphology*, 257, 57–74. doi:10.1016/j.geomorph.2015.12.026.
- Hao, D., Peng, J., Wang, Y., Liu, J., Zhou, X. and Zheng, D. (2019) Evaluation of convolutional neural network for recognizing uterine contractions with electrohysterogram. *Computers in Biology and Medicine*, 113(August), 103394. doi:10.1016/j.compbiomed.2019.103394.
- He, K. and Sun, J. (2015) Convolutional neural networks at constrained time cost. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 5353–5360. doi:10.1109/CVPR.2015.7299173.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, 770–778. doi:10.1109/CVPR.2016.90.

- Hesse, R. (2009) Do swarms of migrating barchan dunes record paleoenvironmental changes? - A case study spanning the middle to late Holocene in the Pampa de Jaguay, southern Peru. *Geomorphology*, 104(3-4), 185–190. doi:10.1016/j.geomorph.2008.08.006.
- Hölbling, D., Eisank, C., Albrecht, F., Vecchiotti, F., Friedl, B., Weinke, E. and Kociu, A. (2017) Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. *Geosciences (Switzerland)*, 7(2). doi:10.3390/geosciences7020037.
- Hossin, M. and Sulaiman, M.N. (2015) A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 01–11. doi:10.5121/ijdkp.2015.5201.
- Huang, L., Liu, L., Jiang, L. and Zhang, T. (2018) Automatic mapping of thermokarst landforms from remote sensing images using deep learning: A case study in the northeastern Tibetan Plateau. *Remote Sensing*, 10(12). doi:10.3390/rs10122067.
- Hugenholtz, C.H., Levin, N., Barchyn, T.E. and Baddock, M.C. (2012) Remote sensing and spatial analysis of aeolian sand dunes: A review and outlook. *Earth-Science Reviews*, 111(3-4), 319–334. doi:10.1016/j.earscirev.2011.11.006.
- Ilyas, B.R., Mohammed, B., Khaled, M. and Miloud, K. (2019) Enhanced Face Recognition System Based on Deep CNN. *Proceedings - 2019 6th International Conference on Image and Signal Processing and their Applications, ISPA 2019*, 15–20. doi:10.1109/ISPA48434.2019.8966797.
- Ji, S., Yu, D., Shen, C., Li, W. and Xu, Q. (2020) Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. *Landslides*, 17(6), 1337–1352. doi:10.1007/s10346-020-01353-2.
- Jiang, J., Zhang, Z., Dong, Q. and Ni, F. (2018) Characterization and identification of asphalt mixtures based on Convolutional Neural Network methods using X-ray scanning images. *Construction and Building Materials*, 174, 72–80. doi:10.1016/j.conbuildmat.2018.04.083.
- Jimenez, J.A., Maia, L.P., Serra, J. and Morais, J. (1999) Aeolian dune migration along the Ceará coast, north-eastern Brazil. *Sedimentology*, 46(4), 689–701. doi:10.1046/j.1365-3091.1999.00240.x.
- Kattenborn, T., Leitloff, J., Schiefer, F. and Hinz, S. (2021) Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173(January), 24–49. doi:10.1016/j.isprsjprs.2020.12.010.

- Khan, H.A., Jue, W., Mushtaq, M. and Mushtaq, M.U. (2020) Brain tumor classification in MRI image using convolutional neural network. *Mathematical Biosciences and Engineering*, 17(5), 6203–6216. doi:10.3934/MBE.2020328.
- Klingenberg, C.P. and McIntyre, G.S. (1998) Geometric morphometrics of developmental instability: Analyzing patterns of fluctuating asymmetry with procrustes methods. *Evolution*, 52(5), 1363 – 1375. doi:10.2307/2411306.
- Kubilius, J., Bracci, S. and Op de Beeck, H.P. (2016) Deep Neural Networks as a Computational Model for Human Shape Sensitivity. *PLoS Computational Biology*, 12(4), 1–26. doi:10.1371/journal.pcbi.1004896.
- Lancashire, L.J., Lemetre, C. and Ball, G.R. (2009) An introduction to artificial neural networks in bioinformatics - Application to complex microarray and mass spectrometry datasets in cancer studies. *Briefings in Bioinformatics*, 10(3), 315–329. doi:10.1093/bib/bbp012.
- LaVezzi, G., Margapuri, V., Stewart, R. and Wagner, D. (2020) Bombus Species Image Classification. *arXiv*.
- Lecun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. *Nature*, 521(7553), 436–444. doi:10.1038/nature14539.
- Li, W. and Hsu, C.Y. (2020) Automated terrain feature identification from remote sensing imagery: a deep learning approach. *International Journal of Geographical Information Science*, 34(4), 637–660. doi:10.1080/13658816.2018.1542697.
- Liu, X.Y., Jia, R.S., Liu, Q.M., Zhao, C.Y. and Sun, H.M. (2019) Coastline extraction method based on convolutional neural networks-A case study of jiaozhou bay in qingdao, China. *IEEE Access*, 7, 180281–180291. doi:10.1109/ACCESS.2019.2959662.
- Lloyd, R., Hodgson, M.E. and Stokes, A. (2002) Visual categorization with aerial photographs. *Annals of the Association of American Geographers*, 92(2), 241–266. doi:10.1111/1467-8306.00289.
- Long, J.T. and Sharp, R.P. (1964) Barchan-dune movement in Imperial Valley, California. *Geological Society of America Bulletin*, 75(1), 149–156. doi:10.1130/0016-7606(1964)75.
- Lonsdale, P. and Malfait, B. (1974) Abyssal dunes of foraminiferal sand on the Carnegie Ridge. *Bulletin of the Geological Society of America*, 85(11), 1697–1712. doi:10.1130/0016-7606(1974)85<1697:ADOF5O>2.0.CO;2.
- Lorenz, R.D., Gasmi, N., Radebaugh, J., Barnes, J.W. and Ori, G.G. (2013) Dunes on planet Tatooine: Observation of barchan migration at the Star Wars film set in Tunisia. *Geomorphology*, 201, 264–271. doi:10.1016/j.geomorph.2013.06.026.

- Lv, P., Dong, Z., Narteau, C. and Rozier, O. (2016) Morphodynamic mechanisms for the formation of asymmetric barchans: improvement of the Bagnold and Tsoar models. *Environmental Earth Sciences*, 75(3), 259. doi:10.1007/s12665-015-5083-2.
- Ma, X., Yan, J. and Fan, F. (2014) Morphology of submarine barchans and sediment transport in barchans fields off the Dongfang coast in Beibu Gulf. *Geomorphology*, 213, 213–224. doi:10.1016/j.geomorph.2014.01.010.
- Maggiori, E., Tarabalka, Y., Charpiat, G. and Alliez, P. (2017) Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 645–657. doi:10.1109/TGRS.2016.2612821.
- Maghsoudi, M., Hajizadeh, A., Nezammahalleh, M.A. and Sedaghat, Z.B. (2017) New method for measurement of barchans parameters Case study: Lut desert, Iran. *Desert*, 22(1), 11–19. doi:https://dx.doi.org/10.22059/jdesert.2017.62216.
- Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F. and Zhang, Y. (2018) Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing*, 10(7). doi:10.3390/rs10071119.
- Maxwell, A.E., Pourmohammadi, P. and Poyner, J.D. (2020) Mapping the topographic features of mining-related valley fills using mask R-CNN deep learning and digital elevation data. *Remote Sensing*, 12(3). doi:10.3390/rs12030547.
- Melton, F.A. (1940) A tentative classification of sand dunes its application to dune history in the southern high plains. *The Journal of Geology*, 48(2), 113–174. doi:10.1086/624871.
- Moosavi, V., Moradi, H., Shamsi, S.R.F. and Shirmohammadi, B. (2014) Assessment of the planimetric morphology of barchan dunes. *Catena*, 120, 12–19. doi:10.1016/j.catena.2014.03.017.
- Mukti, I.Z. and Biswas, D. (2019) Transfer Learning Based Plant Diseases Detection Using ResNet50. *2019 4th International Conference on Electrical Information and Communication Technology, EICT 2019*, (December), 20–22. doi:10.1109/EICT48899.2019.9068805.
- Nagle-Mcnaughton, T., McClanahan, T. and Scuderi, L. (2020) Planet: A neural network for detecting transverse aeolian ridges on mars. *Remote Sensing*, 12(21), 1–15. doi:10.3390/rs12213607.
- Nguyen, A., Yosinski, J. and Clune, J. (2015) Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June. IEEE, pp. 427–436. doi:10.1109/CVPR.2015.7298640.

- Norris, R.M. (1966) Barchan dunes of Imperial Valley, California. *The Journal of Geology*, 74(3), 292–306. doi:10.1086/627164.
- Norris, R.M. and Norris, K.S. (1961) Algodones dunes of southern California. *Geological Society of America Bulletin*, 7(April), 605–620.
- Oo, S.L.M. and Oo, A.N. (2019) Child Face Recognition with Deep Learning. *2019 International Conference on Advanced Information Technologies, ICAIT 2019*, 155–160. doi:10.1109/AITC.2019.8921152.
- Palafox, L.F., Hamilton, C.W., Scheidt, S.P. and Alvarez, A.M. (2017) Automated detection of geological landforms on Mars using Convolutional Neural Networks. *Computers and Geosciences*, 101(January), 48–56. doi:10.1016/j.cageo.2016.12.015.
- Parteli, E.J.R., Durán, O., Bourke, M.C., Tsoar, H., Pöschel, T. and Herrmann, H. (2014) Origins of barchan dune asymmetry: Insights from numerical simulations. *Aeolian Research*, 12, 121–133. doi:10.1016/j.aeolia.2013.12.002.
- Pengra, B.W., Stehman, S.V., Horton, J.A., Dockter, D.J., Schroeder, T.A., Yang, Z., Cohen, W.B., Healey, S.P. and Loveland, T.R. (2020) Quality control and assessment of interpreter consistency of annual land cover reference data in an operational national monitoring program. *Remote Sensing of Environment*, 238(June 2019), 111261. doi:10.1016/j.rse.2019.111261.
- Pike, R.J. (2000) Geomorphometry - diversity in quantitative surface analysis. *Progress in Physical Geography*, 24(1), 1–20. doi:10.1177/030913330002400101.
- Pradhan, A., Dey, B.K., Nath, R.K. and Sarma, B. (2020) Transfer Learning based Classification of Diabetic Retinopathy Stages. *2020 International Conference on Computational Performance Evaluation, ComPE 2020*, 813–819. doi:10.1109/ComPE49325.2020.9200092.
- Pradhan, B. and Lee, S. (2010) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60(5), 1037–1054. doi:10.1007/s12665-009-0245-8.
- Raghu, S., Sriraam, N., Temel, Y., Rao, S.V. and Kubben, P.L. (2020) EEG based multi-class seizure type classification using convolutional neural network and transfer learning. *Neural Networks*, 124, 202–212. doi:10.1016/j.neunet.2020.01.017.
- Rempel, P.J. (1936) The crescentic Dunes of the Salton Sea and their relation to the vegetation. *Ecology*, 17(3), 347–358. doi:10.2307/1931839.
- Rocco, I., Arandjelovic, R. and Sivic, J. (2017) Convolutional neural network architecture for geometric matching. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua, 39–48. doi:10.1109/CVPR.2017.12.

- Rodvold, D.M., McLeod, D.G., Brandt, J.M., Snow, P.B. and Murphy, G.P. (2001) Introduction to artificial neural networks for physicians: Taking the lid off the black box. *Prostate*, 46(1), 39–44. doi:10.1002/1097-0045(200101)46:1(39::AID-PROS1006)3.0.CO;2-M.
- Rubanenko, L., Perez-Lopez, S., Schull, J. and Lapotre, M.G. (2021) Automatic Detection and Segmentation of Barchan Dunes on Mars and Earth Using a Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 9364–9371. doi:10.1109/JSTARS.2021.3109900.
- Sagga, A. (1998) Barchan dunes of Wadi Khulays, western region of Saudi Arabia: Geomorphology and sedimentology relationships. *Journal of King Abdulaziz University-Earth Sciences*, 10(1), 105–114. doi:10.4197/Ear.10-1.7.
- Sauermann, G., Rognon, P., Poliakov, A. and Herrmann, H.J. (2000) The shape of the barchan dunes of Southern Morocco. *Geomorphology*, 36(1-2), 47–62. doi:10.1016/S0169-555X(00)00047-7.
- Scheidt, S.P. and Lancaster, N. (2013) The application of COSI-Corr to determine dune system dynamics in the southern Namib Desert using ASTER data. *Earth Surface Processes and Landforms*, 38(February), 1004–1019. doi:10.1002/esp.3383.
- Shakya, A., Biswas, M. and Pal, M. (2021) Parametric study of convolutional neural network based remote sensing image classification. *International Journal of Remote Sensing*, 42(7), 2663–2685. doi:10.1080/01431161.2020.1857877.
- Silvestro, S., Di Achille, G. and Ori, G.G. (2010) Dune morphology, sand transport pathways and possible source areas in east Thaumasia Region (Mars). *Geomorphology*, 121(1-2), 84–97. doi:10.1016/j.geomorph.2009.07.019.
- Simonyan, K. and Zisserman, A. (2015) Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.
- Sun, Y., Huang, J., Ao, Z., Lao, D. and Xin, Q. (2019) Deep learning approaches for the mapping of tree species diversity in a tropical wetland using airborne LiDAR and high-spatial-resolution remote sensing images. *Forests*, 10(11). doi:10.3390/F10111047.
- Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B. and Liang, J. (2016) Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312. doi:10.1109/TMI.2016.2535302.
- Thakkar, A. and Lohiya, R. (2021) Analyzing fusion of regularization techniques in the deep learning-based intrusion detection system. *International Journal of Intelligent Systems*, (July), 1–49. doi:10.1002/int.22590.

- Tharwat, A. (2018) Classification assessment methods. *Applied Computing and Informatics*. doi:10.1016/j.aci.2018.08.003.
- Todd, B.J. (2005) Morphology and composition of submarine barchan dunes on the Scotian Shelf, Canadian Atlantic margin. *Geomorphology*, 67(3-4), 487–500. doi:10.1016/j.geomorph.2004.11.016.
- Tombe, R. and Viriri, S. (2020) Effective Processing of Convolutional Neural Networks for Computer Vision: A Tutorial and Survey. *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, 0(0), 1–14. doi:10.1080/02564602.2020.1823252.
- Tooth, S. (2013) Google Earth™ in geomorphology: Re-enchanting, revolutionizing, or just another resource? In: *Treatise on Geomorphology*, volume 14, chapter 5. Elsevier, pp. 53–64. doi:10.1016/B978-0-12-374739-6.00372-9.
- Traore, B.B., Kamsu-Foguem, B. and Tangara, F. (2018) Deep convolution neural network for image recognition. *Ecological Informatics*, 48(September), 257–268. doi:10.1016/j.ecoinf.2018.10.002.
- Tsoar, H. (1984) The formation of seif dunes from barchans - a discussion. *Zeitschrift Fur Geomorphologie*, 28(JANUARY 1984), 99–103.
- Tsoar, H. and Parteli, E.J. (2016) Bidirectional winds, barchan dune asymmetry and formation of seif dunes from barchans: a discussion. *Environmental Earth Sciences*, 75(18), 1–10. doi:10.1007/s12665-016-6040-4.
- Van Coillie, F.M., Gardin, S., Anseel, F., Duyck, W., Verbeke, L.P. and De Wulf, R.R. (2014) Variability of operator performance in remote-sensing image interpretation: The importance of human and external factors. *International Journal of Remote Sensing*, 35(2), 754–778. doi:10.1080/01431161.2013.873152.
- van der Merwe, B.J. (2021) The relationship between barchan size and barchan morphology: a case study from Northern Namibia. *South African Geographical Journal*, 103(1), 119 – 138. doi:10.1080/03736245.2021.1876753.
- Vaz, D.A., Sarmiento, P.T., Barata, M.T., Fenton, L.K. and Michaels, T.I. (2015) Object-based Dune Analysis: Automated dune mapping and pattern characterization for Ganges Chasma and Gale crater, Mars. *Geomorphology*, 250, 128–139. doi:10.1016/j.geomorph.2015.08.021.
- Vermeesch, P. and Drake, N. (2008) Remotely sensed dune celerity and sand flux measurements of the world’s fastest barchans (Bod\{e}\{e}, Chad). *Geophysical Research Letters*, 35(24), L24404. doi:10.1029/2008GL035921.
- Wan, S., Liang, Y. and Zhang, Y. (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Computers and Electrical Engineering*, 72, 274–282. doi:10.1016/j.compeleceng.2018.07.042.

- Wang, C. and Anderson, W. (2018) Large-eddy simulation of turbulent flow over spanwise-offset barchan dunes: Interdune vortex stretching drives asymmetric erosion. *Physical Review E*, 98(3), 1–15. doi:10.1103/PhysRevE.98.033112.
- Wang, Y., Fang, Z. and Hong, H. (2019) Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Science of the Total Environment*, 666, 975–993. doi:10.1016/j.scitotenv.2019.02.263.
- Wang, Z.T., Tao, S.C., Xie, Y.W. and Dong, G.H. (2007) Barchans of Minqin: Morphometry. *Geomorphology*, 89(3-4), 405–411. doi:10.1016/j.geomorph.2006.12.014.
- Wiggs, G. (2013) Dune morphology and dynamics. In: J.F. Shroder (Ed.) *Treatise on Geomorphology*, volume 11, chapter 11. Elsevier, pp. 201–218. doi:10.1016/B978-0-12-374739-6.00304-3.
- Wilhelm, T., Geis, M., Püttschneider, J., Sievernich, T., Weber, T., Wohlfarth, K. and Wöhler, C. (2020) DoMars16k: A diverse dataset for weakly supervised geomorphologic analysis on mars. *Remote Sensing*, 12(23), 1–38. doi:10.3390/rs12233981.
- Witharana, C., Bhuiyan, M.A.E. and Liljedahl, A.K. (2020) Big Imagery and High Performance Computing As Resources To Understand Changing Arctic Polygonal Tundra. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIV-M-2-2(November), 111–116. doi:10.5194/isprs-archives-xxiv-m-2-2020-111-2020.
- Worman, S.L., Murray, A.B., Littlewood, R., Andreotti, B. and Claudin, P. (2013) Modeling emergent large-scale structures of barchan dune fields. *Geology*, 41, 1059–1062. doi:10.1130/G34482.1.
- Zeiler, M.D. and Fergus, R. (2014) Visualizing and Understanding Convolutional Networks. In: *Analytical Chemistry Research, ICLR*. pp. 818–833. doi:10.1007/978-3-319-10590-1\_53.
- Zhang, C., Recht, B., Bengio, S., Hardt, M. and Vinyals, O. (2017) Understanding deep learning requires rethinking generalization. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 107–116. doi:10.1145/3446776.
- Zhang, D., Yang, X., Rozier, O. and Narteau, C. (2014) Mean sediment residence time in barchan dunes. *Journal of Geophysical Research: Earth Surface*, 119(3), 451–463. doi:10.1002/2013JF002833.
- Zhang, Z., Dong, Z., Hu, G. and Parteli, E. (2018) Migration and Morphology of Asymmetric Barchans in the Central Hexi Corridor of Northwest China. *Geosciences*, 8(6), 204. doi:10.3390/geosciences8060204.

Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F. and Fraundorfer, F. (2017) Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8–36. doi:10.1109/MGRS.2017.2762307.