A Deep Learning Model for Detecting Dust in Earth's Atmosphere from Satellite Remote Sensing Data

Ping Hou
School for Environment and Sustainability
University of Michigan
Pei Guo
Department of Information Systems
University of Maryland Baltimore County
Aryya Gangopadhyay
Department of Information Systems
University of Maryland Baltimore County
Peng Wu
Department of Hydrology and Atomospheric Sciences
University of Arizona
Jianwu Wang
Department of Information Systems
University of Maryland Baltimore County
Zhibo Zhang
Department of Physics
University of Maryland Baltimore County

***Abstract*—In this paper we develop a deep learning model to distinguish dust from cloud and surface using satellite remote sensing image data. The occurrence of dust storms is increasing along with global climate change, especially in the arid and semi-arid regions. Originated from the soil, dust acts as a type of aerosol that causes significant impacts on the environment and human health. The dust and cloud data labels used in this paper are from CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) satellite. The radiometric channels and geometric parameters from VIIRS (Visible Infrared Imaging Radiometer Suite) satellite sensor serve as features for our model.**

We trained and tested our deep learning model using 10,000 samples in March 2012. The developed model has five hidden layers and 512 neurons in each layer. The classification accuracy on the test set is 71.1%. In addition, we performed a shuffling procedure to identify the importance of features, which is calculated as the increase in the prediction error after we permute the feature's values. We also developed a method based on genetic algorithm to find the best subset of features for dust detection. The results show that the genetic algorithm can select a subset of features that have comparable performance as that of a model with all features. The shuffling procedure and the genetic algorithm both identify geometric information as important features for detecting mineral dust. The chosen subset will improve computational efficiency for dust detection and improve physical based methods.

***Index Terms*—Climate pollution; Deep learning; Feature selection; Classification; Genetic algorithm**

Supported by NSF grants IIS-1923982, CNS-0821258, CNS-1228778, OAC-1730250, OAC-1726023, OAC-1942714, and DMS–0821311.

I. Introduction

Dust storm affects many domains related to smart computing, including transportation, environmental protection and healthcare. Dust storm occurrence is increasing under the background of global climate change, especially in the arid and semi-arid regions. Dust, originated from the soil, acts as a type of aerosol that causes significant impacts on the environment and human health. Dust storms reduce visibility and cause dangers for high way traffic. For people with respiratory conditions like asthma, chronic obstructive airways disease (COAD) or emphysema, even small increases of dust concentration can make their symptoms worse.

Dust is also the most abundant aerosol component in terms of dry mass [1], [2]. Dust aerosols can interact with both solar and thermal infrared radiation, which gives them an important role in regulating the radiative energy balance of Earth-Atmosphere system. After lifted and transported by wind, dust aerosols can absorb and scatter solar radiation and warm the surrounding air. It reduces the sun's radiation that reaches the surface, imposing a shortwave cooling effect [3], [4]. On the other hand, dust absorbs longwave radiation and re-emits to the surface, imposing a warming effect on the surface [5], [6]. Dust particles can also act as cloud condensation nuclei (CCN) or ice nuclei (IN) in cloud formation processes and alter cloud lifetime and radiative effect by changing cloud droplet number concentration and size [7], [8].The radiative effects dust depend on a variety of factors including dust loading, dust particle size, dust refractive index and dust vertical distribution. Currently

and in the near future satellite observation is the only means to monitor the occurrence of dust storm and the properties of dust aerosols on a regional to global scale.

With the rapid development of satellite remote sensing, various methods have been proposed to utilize multi-channel observations to detect and retrieve dust information [9]. The Moderate Resolution Imaging Spectroradiometer (MODIS) is a widely used passive sensor with 36 channels in dust detection. Starting from 2011, as a replacement to MODIS, the Visible Infrared Imaging Radiometer Suite (VIIRS) on board Suomi National Polar-orbiting Partnership (NPP) spacecraft and Aqua satellite was launched. The VIIRS sensor has 16 moderate resolution (750 m) channels. Physical-based retrieval methods, such as [10] and [11], have been proposed to identify dust aerosols from the VIIRS moderate resolution channels and some of them were adopted from the MODIS channels.

The physical-based methods, however, highly depend on empirical thresholds to differentiate dust and dust-free pixels. Also, the detection accuracy is 40-50% when compared with collocated active remote sensing (e.g., the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation, CALIPSO) dust index [12]. Because of the ample information from multispectral satellite observations, machine learning and deep learning based approaches have been proposed to automate the dust detection process. In this paper, we propose a deep learning model to detect mineral dust using VIIRS data.

The rest of the paper is organized as follows. Section II connects our study to existing studies in the literature. Section III discusses datasets and the deep learning model used in this paper. Section IV presents the results of cloud and dust identification from the deep learning model, followed by discussions of the results in terms of significance and impacts in Section V. Section VI concludes the paper.

II. Related Work

Some widely used physical-based retrieval methods include (a) Normalized Difference Dust Index (NDDI) [10], [11]; (b) Brightness Temperature Difference (BTD) (11 μm $-$ 12 μm) [13], [14]; (c) BTD (8.6 μm $-$ 11 μm) [15]; (d) Reflective Solar Band (RSB) [16]. NDDI is only appropriate in detecting dust storms when a dust-free image from a nearby time period is available. The BTD methods are simple and efficient in detecting dust. However, any pixel that have BTD exceed the threshold difference will be classified as dust pixel and this can mis-identify land pixels as dust. The RSB method requires significant amount of dust-free pixels in determining the threshold and this method has the same problem of mis-identifying land pixels as dust as in the BTD methods.

Several machine learning based approaches have also been proposed over the last decade. Based on the feature set of spectral bands reported in literature, Murguia et al. developed a Maximum Likelihood (ML) classifier and a Probabilistic Neural Network (PNN) to detect the dust storms [17]. They found PNN provides improved classification performance with reference to the ML classifier and their method allows for real-time processing. Han et al. developed a decision tree classifier based on visually inspected dust events and their associated multispectral MODIS images, and found that the prediction agree in general with surface observations [18]. Mario et al. developed an Artificial Neural Network (ANN) to detect dust storm from multispectral MODIS images [19]. Using selected MODIS channels, Amir and Sanaz proposed a random forests algorithm to detect dust plums over water and land. The results were shown to be better than those physics-based methods [20]. In [12], Shi et al. proposed a hybrid approach combining physical model with traditional data mining models such as Random Forest, which achieved better accuracy than each individual method.

III. Methodology

Existing machine learning studies have used either visually selected dust events or pre-selected channels to train the models. The models may not always perform well in the thin dust layer events, and the models may miss information from spectral channels that were not selected as input features. In addition, the proposed methods above did not perform feature selection after training their models. Selecting a subset of features with similar prediction accuracy can reduce the computational time and make it easier to apply the model to real-time detection. In addition, knowing important features on dust detection could also help sensor design for future satellite missions. In this paper, we develop and test a deep learning model for dust detection and select a best subset that have similar prediction accuracy as using all channels.

Besides simply classifying dust and dust-free pixels, we also add cloud information in the learning procedure. Dust affect climate through cloud by changing cloud lifetime, thus it is imperative to study dust effect on cloud using synthetic dust and cloud observations. The true dust and cloud information is from collocated CALIPSO level-2 data and it is used in training and verifying the deep learning model.

The Visible Infrared Imaging Radiometer Suite (VIIRS) is one of the key instruments on board the Suomi

National Polar-Orbiting Partnership (Suomi NPP) spacecraft, which was successfully launched on October 28, 2011. VIIRS has 22 channels, 16 of which are moderate resolution bands (M-bands) and have a spatial resolution of 750 m at the nadir. The other six channels are made up of five imaging resolution bands (I-bands), which have a spatial resolution of 375 m at the nadir, and one day/night panchromatic band with a spatial resolution of 750 m. The 22 channels cover wavelengths from 0.41 to 12.5 μm and can provide data records for clouds, aerosol, sea surface temperature, snow and ice, vegetation and fire.

The satellite, sun, and target relative positions are important factors affecting the amount of radiation received by the satellite sensor. Though it is hard to quantify the 3D radiative effect, the geometric information cannot be ignored in classifying pixel categories. In this study, in addition to the radiometric channels, four geometric parameters from VIIRS, namely view zenith angels, solar zenith angels, view azimuth angels and solar azimuth angels, are used in training the deep learning model.

The dust and cloud information from CALIPSO satellite was used as labels in the deep learning model. CALIPSO was launched in 2006 as part of A-Train and on board CALIPSO. There is an active remote sensor, lidar, available that can provide reliable cloud and dust aerosol index.

Four categories were classified in this project with (1) dust with no cloud, (2) cloud with no dust, (3) dust with cloud, and (4) others. From global perspective, cloud occurrence is much higher than dust, this will likely to cause imbalanced samples in the training and test datasets. To avoid this, we chose the same number of samples for each category. For test purpose and to speed up the code, we selected total 10,000 samples in March 2012 and the samples are equally distributed among the four categories.

A. Deep learning model

A deep learning model simulates the way biological nervous systems (e.g., human brain) process information [21]. A deep learning model composes multiple layers of neurons. In this study, we used Deep Neural Networks (DNN) deep learning model. The first layer contains input predictors (in this case, 16 radiometric channels and four geometric parameters), and the last layer contains output responses (in this case, the classification of the pixel types). Between the input layer and the output layer are one or more hidden layers interconnected with each other by hidden neurons. Each layer extracts features of the input for classification. The use of multiple hidden layers allows the construction of hierarchical features at different levels of resolution [22]. The choice of the number of hidden layers and the number of hidden neurons in each layer is often guided by background knowledge and experimentation. It is most common to have a reasonably large number of hidden neurons and train them with regularization. In this study, we determine the number of hidden layers and the number of hidden neurons by trial and error. We use L2 regularization in each layer to suppress the large weights and result in a model that is more stable and less like to overfit the training data. The activation function and the learning algorithm in deep learning models are also selected by trial and error.

B. Feature selection

Another goal of this study is to find the important input features for dust detection. We use two approaches to conduct feature selection in the deep learning model: a shuffling procedure and a genetic algorithm. The shuffling procedure gives an importance order of the features, while the genetic algorithm selects a subset of the features that can generate the optimal prediction performance.

1) Shuffling procedure: The procedure is first to get a benchmark test accuracy by training the model once and then predict multiple times while randomizing each variable in the test set. The difference of the benchmark test accuracy and the test accuracy after permuting the variable, meaning with and without the help of this variable, is used as an importance measure (i.e., permutation importance). If the accuracy after randomizing a variable is lower than the benchmark test accuracy, it is an important variable. On the other hand, if nothing changes or the accuracy is higher than the benchmark, it is a useless variable. We randomize 50 times and get an average test accuracy for each variable and compare with the benchmark test accuracy.

2) Genetic algorithm: Genetic algorithm is a directed random search technique that simulates the natural selection and evolution process [23]. Because it can be directly integrated to existing simulations and models, genetic algorithm has been widely used for many optimization problems which have a large number of parameters and their analytical solutions are hard to derive [24]. Rationally, genetic algorithm has also been used to optimize deep learning models [25], [26]. Here, we use a genetic algorithm to select a subset of features in our deep learning model. A sequence indicating whether a features is selected or not is defined as a genome.

1) Initialization: we create a certain number of deep learning models with randomly generated genomes to be the population of the first generation.

2) Fitness evaluation: we train each model in the population and evaluate its performance on the test set using classification accuracy.
3) Selection: we rank all models in the population by accuracy and keep 20% of the top-ranked models to become part of the next generation to breed children. we also randomly keep 10% of the rest of the models. This helps find potentially successful combinations between worse-performers and top-performers, and also helps avoiding stuck in local maximum.
4) Crossover: crossover is the combination process from two members of a population to generate one or more children. Besides the top 20% models and the randomly kept 10% non-top models, to keep our population of 30 models, 21 children are generated for breeding in each generation.
5) Mutation: we randomly mutate some of the genomes on some of the kept models.
6) Genome replacement: genomes of the previous generations are replaced using the genomes after crossover and mutation.
7) Step 2 to step 6 are repeated for multiple generations until the model performance converges, i.e., the test accuracy will not get any better. The best performed genome in the final generation is the selected best deep learning model, and the best performed genome indicates the selected subset of the features.

IV. Empirical Evaluation

A. Deep learning model

Our final developed deep learning model has five hidden layers with 512 neruons in each layer. The activation function in each layer is "relu", which is found to achieve better results than other functions. And we use the optimizer "adam" to train the model. Figure 1 shows the training and validation loss on the test set. The validation loss is very close to the training loss, indicating no over-fitting occurred. The classification accuracy on the test dataset is 71.1%.

B. Feature selection by shuffling procedure

Figure 2 shows the test accuracy when shuffling each feature in the model. The red horizontal line indicates the benchmark test accuracy (68.3%) when no variable is shuffled. Note the accuracy is different from 71.1% in Section IV-A. This is due to the randomness in the training process. Neural networks and genetic algorithm are both stochastic, which means they make use of randomness (e.g., random weights initialized in the deep learning model, population random generated in the genetic algorithm) and therefore each time can produce different results. Each box in Figure 2 indicates the distribution of the test accuracy of each variable being shuffled 50 times. Overall, the average test accuracy of all the variables are smaller than the benchmark test accuracy, which means all the variables contribute to the model to a certain degree. But some are more important than others. The most important features are the geometric information of solar zenith angle (18), view zenith angle (20), and solar azimuth angle (17), followed by channels 15 and 16 (11 and 12 μm) which are consistent with the NDDI and RSB methods from physical retrievals mentioned in Section II. Because dust particles are non-sphere particles, so their reflectance are different from different view angles, so are the emittance.

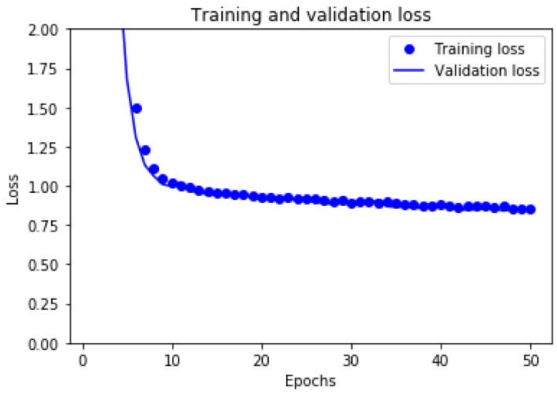

Fig. 1. Training and validation loss on the test dataset.

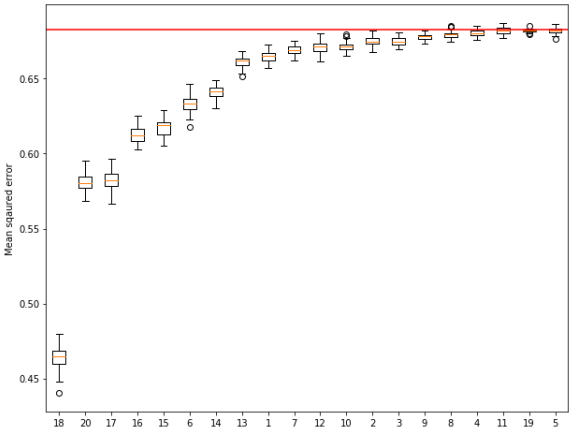

Fig. 2. Test accuracy after shuffling each feature.

C. Feature selection by genetic algorithm

Table I shows the selected features by genetic algorithm and their performance. We test scenarios with

TABLE I
FEATURES SELECTED BY GENETIC ALGORITHM

Population size	Number of generations	Selected features	Best test accuracy
8	4	2, 3, 6, 8, 9, 11, 12, 14, 15, 17, 18, 19	67.8%
16	4	2, 4, 8, 11, 13, 14, 15, 17,18, 19, 20	68.1%
32	4	1, 2, 3, 4, 5, 9, 10, 12, 13, 14, 16, 17, 18, 19, 20	70.3%
32	8	1, 3, 6, 8, 9, 10, 11, 17, 18, 19, 20	70.1%
64	8	1, 5, 7, 9, 12, 15, 16, 17, 18, 19, 20	71.5%

different population size and different number of generations. The results show that, in general, when we have a larger population size and more generations, the genetic algorithm is able to find a better solution. When we have a population size as 64 and 8 generations, the test accuracy can achieve 71.5%, which is even better than using all the features (71.1%). This proves that genetic algorithm is able to find a subset of features that can generate comparable results with all available features.

For the selected features in Table I, the geometric information are important regardless of the population size and number of generations. This is consistent with the result from the shuffling procedure and is physically reasonable due to the non-spherical radiative properties of dust particles. The geometric matrix, however, are not adequately characterized in the physical based classification methods. Shi et al. also showed that machine learning methods that include geometric angles have a higher prediction accuracy than the physical based methods [12].

The channels used in physical based models are all selected as important features in Table I, suggesting the deep learning model is capable of classifying dust and cloud in a physically feasible way.

Note that the prediction accuracy from our deep learning model is lower than that by Shi et al. in [12]. This is due to several reasons. First, they use 36 channels from MODIS, while we use 16 channels and 4 geometric parameters from VIIRS. Second, they only classify dust pixels while we classify both dust and cloud pixels. Besides, they use the training data and test data from the same day, which is likely to have higher prediction accuracy because of the similar dust and atmospheric properties. While in our study, the training and test datasets are randomly chosen and are not necessarily from the same day or over a specified region. To further improve our deep learning model, a more extensive dataset is required.

V. DISCUSSION

The selections of some of the channels are obvious from physics point view. The reflectance of dust in visible wavelength (0.4-0.7 μm) increase steadily, so the reflectance and differences in visible channels (e.g., channel 1 to channel 5) can be indications of dust aloft. Dust layers usually transport at an altitude of 5 km, a height that above most of water vapor, so the dust particles have higher reflectance in channel 9 (1.38 μm) than cloud and this provides the possibility to separate dust pixels from cloudy pixels. From radiative transfer calculations in [13], the ratio of channel 4 (0.54 μm) to channel 7 (0.86 μm) is greater than 1 when no cloud or dust is present in the pixel, and the difference in the two channels is mainly contributed by molecular scattering. Because shorter VIIRS wavelength often saturates in cloud filled pixels, this ratio is near 1 for cloudy pixels. For dust, the reflectance is lower in shorter wavelength, so this ratio will decrease to below 1 in dust pixels. With the increase of dust optical depth, this ratio decreases. Fine dust particles have different emissivites in channel 14 (8.6 μm) and channel 15 (11 μm). This produces brightness temperature differences in the two channels. In window regions, Ackerman found that the difference of brightness temperature in channel 15 (11 μm) and channel 16 (12 μm) is negative (less than 1K) for dust, among other aerosol types [10]. This threshold has been used in BTD method ever since it was proposed. The channels used in physical-based methods are all selected as important features in our model.

Besides those, some other channels also appear to be significant, for example, channels 8 and 11. Although it is not clear how, this may suggest that channels 8 and 11 also contain useful information to identify dust, and this can provide directions for future satellite dust detection retrieval developments.

Identify a limited number of features is valuable for dust detection. First, with fewer data to train the model, the computational efficiency of predicting the dust storm will significantly be improved. Real-time prediction is vital for providing in-time prevention and warning and thus reduce the harmful impacts caused by dust storms. Second, we can implement the geometric angles that identified as important features into physical based methods, which will improve the accuracy of these methods. To further enhance the deep learning model,

more data points are needed. With the additional training data, more CPU nodes or GPU will likely be required to facilitate the computation.

VI. Conclusion

The quantitative monitoring of dust storm is of great significance for disease prevention, environmental protection, and sustainable development. In this study, a deep learning model was trained and used to classify dust and cloud using VIIRS and CALIPSO data. The deep learning model achieved a benchmark prediction accuracy of 71.1%. We use a shuffling procedure to identify the importance of features in the model and a genetic algorithm to conduct feature selection. Through careful tuning of population size and number of generations, the model can predict with comparable accuracy using a subset of the variables. The selected subsets consist of the channels used in the physical based classification methods, and the geometric angles are always shown importance in the subsets for duct detection.

References

[1] O. Alizadeh Choobari, P. Zawar-Reza, and A. Sturman. The global distribution of mineral dust and its impacts on the climate system: A review. *Atmospheric Research*, 138:152 – 165, 2014.

[2] C. Textor, M. Schulz, S. Guibert, S. Kinne, Y. Balkanski, S. Bauer, T. Berntsen, T. Berglen, O. Boucher, M. Chin, F. Dentener, T. Diehl, R. Easter, H. Feichter, D. Fillmore, S. Ghan, P. Ginoux, S. Gong, A. Grini, J. Hendricks, L. Horowitz, P. Huang, I. Isaksen, I. Iversen, S. Kloster, D. Koch, A. Kirkevåg, J. E. Kristjansson, M. Krol, A. Lauer, J. F. Lamarque, X. Liu, V. Montanaro, G. Myhre, J. Penner, G. Pitari, S. Reddy, Ø. Seland, P. Stier, T. Takemura, and X. Tie. Analysis and quantification of the diversities of aerosol life cycles within aerocom. *Atmospheric Chemistry and Physics*, 6(7):1777–1813, 2006.

[3] Ina Tegen and Andrew A. Lacis. Modeling of particle size distribution and its influence on the radiative properties of mineral dust aerosol. *Journal of Geophysical Research: Atmospheres*, 101(D14):19237–19244, 1996.

[4] H. Yu, Y. J. Kaufman, M. Chin, G. Feingold, L. A. Remer, T. L. Anderson, Y. Balkanski, N. Bellouin, O. Boucher, S. Christopher, P. DeCola, R. Kahn, D. Koch, N. Loeb, M. S. Reddy, M. Schulz, T. Takemura, and M. Zhou. A review of measurement-based assessments of the aerosol direct radiative effect and forcing. *Atmospheric Chemistry and Physics*, 6(3):613–666, 2006.

[5] Jasper F. Kok, David A. Ridley, Qing Zhou, Ron L. Miller, Chun Zhao, Colette L. Heald, Daniel S. Ward, Samuel Albani, and Karsten Haustein. Smaller desert dust cooling effect estimated from analysis of dust size and abundance. *Nature Geoscience*, 10(4):274–278, 2017.

[6] Q. Song, Z. Zhang, H. Yu, S. Kato, P. Yang, P. Colarco, L. A. Remer, and C. L. Ryder. Net radiative effects of dust in the tropical north atlantic based on integrated satellite observations and in situ measurements. *Atmospheric Chemistry and Physics*, 18(15):11303–11322, 2018.

[7] S Twomey. Pollution and the planetary albedo. *Atmospheric Environment (1967)*, 8(12):1251–1256, 1974.

[8] Bruce A Albrecht. Aerosols, cloud microphysics, and fractional cloudiness. *Science*, 245(4923):1227–1230, 1989.

[9] SA Ackerman, S Platnick, PK Bhartia, B Duncan, T L'Ecuyer, A Heidinger, G Skofronick-Jackson, N Loeb, T Schmit, and N Smith. Satellites see the world's atmosphere. *Meteorological Monographs*, 59:4–1, 2019.

[10] Steven A Ackerman. Using the radiative temperature difference at 3.7 and 11 μm to tract dust outbreaks. *Remote Sensing of Environment*, 27(2):129–133, 1989.

[11] S. D. Miller. A consolidated technique for enhancing desert dust storms with MODIS. *Geophysical Research Letters*, 30(20), oct 2003.

[12] Peichang Shi, Qianqian Song, Janita Patwardhan, Zhibo Zhang, Jianwu Wang, and Aryya Gangopadhyay. A hybrid algorithm for mineral dust detection using satellite data. *Proceedings of the 15th IEEE International Conference on e-Science (e-Science2019)*, pages 39–46, 2019.

[13] J. K. Roskovensky and K. N. Liou. Differentiating airborne dust from cirrus clouds using modis data. *Geophysical Research Letters*, 32(12), 2005.

[14] Yang Liu, Ronggao Liu, and Xiao Cheng. Dust detection over desert surfaces with thermal infrared bands using dynamic reference brightness temperature differences. *Journal of Geophysical Research: Atmospheres*, 118(15):8566–8584, 2013.

[15] J. J. Qu, X. Hao, M. Kafatos, and L. Wang. Asian dust storm monitoring combining terra and aqua modis srb measurements. *IEEE Geoscience and Remote Sensing Letters*, 3(4):484–486, Oct 2006.

[16] Mehdi Samadi, Ali Darvishi Boloorani, Seyed Kazem Alavipanah, Hossein Mohamadi, and Mohamad Saeed Najafi. Global dust detection index (gddi); a new remotely sensed methodology for dust storms detection. *Journal of Environmental Health Science and Engineering*, 12(1):20, Jan 2014.

[17] Murguia M.I.C. Tilton J.J. Rivas-Perea P., Rosiles J.G. Automatic dust storm detection based on supervised classification of multispectral data. *Melin P., Kacprzyk J., Pedrycz W. (eds) Soft Computing for Recognition Based on Biometrics. Studies in Computational Intelligence, vol 312*, 2010.

[18] Han Tao, Li Yaohui, Han Hui, Zhang Yongzhong, and Wang Yujie. Automatic detection of dust storm in the northwest of china using decision tree classifier based on modis visible bands data. In *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.*, volume 5, pages 3603–3606, July 2005.

[19] Mario I. Chacon-Murguía, Yearim Quezada-Holguín, Pablo Rivas-Perea, and Sergio Cabrera. Dust storm detection using a neural network with uncertainty and ambiguity output analysis. In José Francisco Martínez-Trinidad, Jesús Ariel Carrasco-Ochoa, Cherif Ben-Youssef Brants, and Edwin Robert Hancock, editors, *Pattern Recognition*, pages 305–313, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[20] Amir Hossein Souri and Sanaz Vajedian. Dust storm detection using random forests and physical-based approaches over the middle east. *Journal of Earth System Science*, 124(5):1127–1141, Jul 2015.

[21] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.

[22] Trevor Hastie, Robert Tibshirani, Jerome Friedman, and James Franklin. The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85, 2005.

[23] John Henry Holland et al. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[24] Duc Pham and Dervis Karaboga. *Intelligent optimisation techniques: genetic algorithms, tabu search, simulated annealing and neural networks*. Springer Science & Business Media, 2012.

[25] PG Benardos and G-C Vosniakos. Optimizing feedforward artificial neural network architecture. *Engineering Applications of Artificial Intelligence*, 20(3):365–382, 2007.

[26] Marylyn D Ritchie, Bill C White, Joel S Parker, Lance W Hahn, and Jason H Moore. Optimizationof neural network architecture using genetic programming improvesdetection and modeling of gene-gene interactions in studies of humandiseases. *BMC bioinformatics*, 4(1):28, 2003.