remote sensing

MDPI

Article

Comparing the Performance of Neural Network and Deep Convolutional Neural Network in Estimating Soil Moisture from Satellite Observations

Lingling Ge [1], Renlong Hang [1], Yi Liu [2] and Qingshan Liu [1,*]

[1] Jiangsu Key Laboratory of Big Data Analysis Technology, School of Information and Control, Nanjing University of Information Science and Technology, Nanjing 210044, China; linglingge@nuist.edu.cn (L.G.); renlonghang@nuist.edu.cn (R.H.)

[2] School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China; yi.liu@nuist.edu.cn

* Correspondence: qsliu@nuist.edu.cn

Received: 16 July 2018; Accepted: 16 August 2018; Published: 21 August 2018



Abstract: Soil moisture (SM) plays an important role in hydrological cycle and weather forecasting. Satellite provides the only viable approach to regularly observe large-scale SM dynamics. Conventionally, SM is estimated from satellite observations based on the radiative transfer theory. Recent studies have demonstrated that the neural network (NN) method can retrieve SM with comparable accuracy as conventional methods. Here, we are interested in whether the NN model with more complex structures, namely deep convolutional neural network (DCNN), can bring about further improvement in SM retrievals when compared with the NN model used in recent studies. To achieve this objective, the same input data are used for the DCNN and NN models, including L-band Soil Moisture and Ocean Salinity (SMOS) brightness temperature (TB), C-band Advanced Scatterometer (ASCAT) backscattering coefficients, Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) and soil temperature. The target SM used to train the DCNN and NN models is the European Center for Medium-range Weather Forecasts Re-Analysis Interim (ERA-Interim) product. The experiment consists of two phases: the learning phase from 1 January to 31 December 2015 and the testing phase from 1 January to 31 December 2016. In the learning phase, we train the DCNN and NN models using the ERA-Interim SM. When evaluation between DCNN and NN against in situ measurements in the testing phase, we find that the temporal correlations between DCNN SM and in situ measurements are higher than those between NN SM and in situ measurements by 6.2% and 2.5% on ascending and descending orbits, respectively. In addition, from the perspective of temporal and spatial dynamics, the simulated SM values by DCNN/NN and the ERA-Interim SM agree relatively well at a global scale. Results suggest that both NN and DCNN models are effective in estimating SM from satellite observations, and DCNN can achieve slightly better performance than NN.

Keywords: soli moisture retrieval; neural network; deep convolutional neural network

1. Introduction

Soil moisture (SM) is an essential climate variable [1] and one of the most important drivers of the hydrological cycle, as it governs the redistribution of precipitation between infiltration and runoff [2]. It is also a key variable in better understanding of the land–atmosphere interactions [3]. Additionally, due to the soil characteristics and water content, events such as rainstorms can lead to floods or landslides. Having accurate and timely SM data will give better weather forecasts and better predictions of hazardous events [4].

SM data can be obtained from in situ measurements and satellite observations. In situ measurements provide accurate SM information and enable frequent acquisitions of data each day, but only for individually discrete locations. Satellite observations can acquire the large-scale SM dynamics at reasonable temporal intervals. Microwave observations have proven to be one of the most promising remote sensing approaches to monitor SM at the global scale [5–7].

As an exclusively designed satellite, Soil Moisture and Ocean Salinity (SMOS) has been widely used for SM retrieval [8–10]. The operational SMOS SM retrieval algorithm is based on a forward model [11]. It firstly uses the radiation transfer equation to simulate brightness temperature (TB). Then, the estimated SM values are obtained using an iterative method to minimize the difference between the simulated TB and the observed one by SMOS satellite. This model is considered to be local and time-consuming as each new observation needs a minimization process [12].

Different from the forward model, the inverse model provides another avenue to retrieve SM. Among various inverse models, neural network (NN) is a commonly used one. The NN method is a fully-connected feedforward model, which can be trained to learn the relationship between the satellite observations and the target SM values. To train the NN model, many works attempted to use the in situ measurements as the target data [2,10,13,14]. However, it is difficult to train an effective model that can be applied to the global SM retrieval when using a few in situ measurements as the target data. Thus, some works chose the simulated values by the radiative transfer model [15] or the estimated ones by the global land surface model [16,17] as the target data to train NN. In [12,18,19], NN was trained to describe the link between the satellite observations and the target SM values come from the Medium-range Weather Forecasts (ECMWF) model.

As an extension of NN, deep learning, especially a deep convolutional neural network (DCNN), has attracted increasing attention in the field of remote sensing recently [20,21]. The goal of deep learning is to automatically construct the complex relationship from input to output in a hierarchical manner. Though the DCNN method is promising, it has not been applied in estimating SM from satellite observations. To our knowledge, it is the first attempt to retrieve SM from satellite observations using a DCNN method. The objective of this study is to investigate whether the DCNN can bring about improvement in SM retrievals when compared with the NN method used in recent studies. If so, our second objective is to further identify the advantages and weakness of DCNN and NN methods in retrieving SM from satellite observations.

2. Datasets

In this section, we describe the data sets used in this paper in detail, including SMOS observations, Advanced Scatterometer (ASCAT) backscattering coefficients, Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI), ECMWF data and in situ SM measurements.

2.1. SMOS Data

The SMOS mission is a joint European Space Agency, National Centre for Space Studies, and Centro para el Desarrollo Tecnologico Industrial program that was launched on 2 November 2009 [11]. The SMOS satellite carries interferometric radiometer that operates at L-band (1.4 GHz). It generates records of TB over incidence angles from 0° to 65° with a spatial resolution in the range of 30–50 km [4]. The multi-angular TB is binned in incidence angle bins and the average width of incidence angle bins is 5°. The SMOS satellite flies in a sun-synchronous orbit and its time of equator overpass is 6:00 a.m. (ascending node) and 6:00 p.m. (descending node).

The SMOS Level 3 (SMOSL3) gridded multi-angular TB data from both ascending and descending overpasses are provided by Centre Aval de Traitement des Données SMOS (CATDS) in France. SMOSL3 TB product provides multi-angular TB data set in a 0.25° Equal Area Scalable Earth (EASE) grid [22]. Consistent with the work in [12], the SMOSL3 TB data in both horizontal and vertical

polarizations with incidence angles ranging from 20° to 60° is used. In addition, ascending and descending orbits data are processed separately.

In addition, the probability of having a radio frequency interferences (RFI) at each grid point provided by CATDS is also used because the RFI might perturb SMOS mission in certain areas of the world [23]. Following the work in [24], grid points with a cumulative RFI probability higher than 20% are filtered during data preprocessing.

2.2. ASCAT Data

ASCAT is a real aperture radar operating at 5.255 GHz (C-band) on board EUMETSAT's Meteorological Operational (MetOp) satellite [25]. It measures on both sides of the subsatellite track, producing two swaths of surface radar backscatter data with good radiometric accuracy and stability [26]. The incidence angles of ASCAT range from 25° to 65° with a spatial resolution of 25 km. The MetOp satellite also runs in a sun-synchronous orbit, but its equator overpass time is approximately 9:30 p.m. (ascending overpass) and 9:30 a.m. (descending overpass).

ASCAT is an active microwaves sensor and has already been used to retrieve SM [26,27]. Here, we use ASCAT backscattering coefficients as auxiliary data. Before using the ASCAT data, the backscattering coefficients are needed to be normalized to a standard incidence angle. Motivated by the processing chain of the European remote sensing satellite and ASCAT SM products in [28], we choose the middle of the observed incidence angle 40° as the standard incidence angle. According to previous studies [26,29], the relationship between backscatter and incidence angle is linear over a large range of angles. As a result, we use linear regression to obtain the standard backscattering coefficients at the incidence angle of 40°. Since the backscattering coefficients can be measured several times per day, we compute a daily average value.

2.3. MODIS NDVI

NDVI is related to SM in many ecosystems. Paloscia et al. [30] acquired different accuracies with or without using NDVI for SM retrieval. They found that NDVI information can improve the retrieval performances. Accordingly, we use the MODIS NDVI product MYD13C1 in this paper. The MYD13C1 is provided as a Level 3 product projected on a 0.05° geographic Climate Modeling Grid (CMG). Since the SMOSL3 TB data are set in a 0.25° EASE grid, we use a nearest-neighbor interpolation method to change the 0.05° CMG grid to the 0.25° EASE grid.

2.4. ECMWF Data

The European Center for Medium-range Weather Forecasts Re-Analysis Interim (ERA-Interim) product from ECMWF public datasets web is used here. It is a global atmospheric reanalysis from 1979, continuously updated in real time [31]. It is produced by a data assimilation system based on a 2006 release of the ECMWF Integrated Forecasting System. As discussed in [32–34], the Tiled ECMWF Scheme for Surface Exchanges over Land (TESSEL) land-surface model is used for the ERA-Interim model. Specifically, the ERA-Interim model uses the TESSEL [35,36] scheme to evolve the thermal and water storage in the four layers of soil and snow during the forecast. The volumetric water content in the first layer of soil is used as the target SM values, hereafter is referred to as ERA-interim SM. Note that, for a good match for the training, the ERA-interim SM with a resolution of $0.25° \times 0.25°$ is selected. In addition, the simulated SM values from the ERA-interim product can be acquired four times (12:00 a.m., 6:00 a.m., 12:00 p.m., 6:00 p.m.) a day. In order to be consistent with the time of SMOS acquisitions, ERA-Interim SMs at the time of 6:00 a.m. and 6:00 p.m. are selected. In addition, the soil temperature and snow depth in the first layer of soil are also used to filter out the regions with frozen soils or snow.

2.5. In Situ SM Measurements

In situ SM measurements from Soil Climate Analysis Network (SCAN) are used to assess the performance of SM retrieval results. The SCAN provided by the U.S. Department of Agriculture, Natural Resources Conservation Service contains 232 stations. Each station measures SM every half hour. In order to be consistent with the time of SMOS acquisitions, in situ measurements of SM from SCAN sites at the time of 6:00 a.m. and 6:00 p.m. are collected, separately. In addition, the depths of SM measurements obtained by SM sensors vary from 0.02 m to 2.03 m [37]. Here, we use the in situ SM measurements in the 0–5 cm depth range, which is consistent with that of ERA-Interim SM [38].

3. Methodology

The goal of SM retrieval is to estimate SM values when given satellite observations. From the machine learning point of view, this process can be considered as a regression problem, which usually consists of two phases: the learning phase and the testing phase. In the learning phase, we need to collect the collocated input and output data to construct a training set. We put SMOSL3 TB data, ASCAT backscattering coefficients, MODIS NDVI and soil temperature together to generate a 19-dimensional input vector for each grid all over the globe, where the first 16 inputs correspond to SMOSL3 TB values in both horizontal and vertical polarizations with incidence angles ranging from 20° to 60° (the centers of angle bins are 22.5°, 27.5°, 32.5°, 37.5°, 42.5°, 47.5°, 52.5°, 57.5° respectively), and the other three inputs are the average backscattering coefficient at the incidence angle 40°, the MODIS NDVI product MYD13C1, as well as the soil temperature in the first layer of soil; meanwhile, we use the ERA-Interim SM as the 1-dimensional output value. This training set is used to train the regression model to learn the fitting coefficients. In the testing phase, the trained model is used to predict/estimate the SM value once given input data.

3.1. Neural Network

As one of the most widely used regression models in the field of machine learning, NN has been successfully applied to SM retrieval [12,19,39]. Figure 1 shows an example of NN. It consists of three layers: the input layer, the hidden layer, and the output layer. The numbers of nodes in each layer are respectively set as 19, 10 and 1. Here, we choose 10 nodes in the hidden layer empirically because it can achieve satisfying retrieval results.

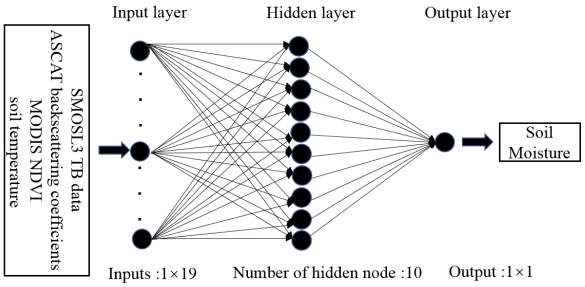


Figure 1. The architecture of the NN model.

Assume that the input vector is x, then the output h of the hidden layer can be described as:

$$h = \sigma(w_1 x + b_1), \tag{1}$$

where w_1 is the weight between the input node and the hidden node, b_1 is the bias and σ is the nonlinear activation function. Similarly, the estimated SM $\hat{y}$ of the output layer is:

$$\hat{y} = \sigma(w_2 h + b_2), \tag{2}$$

where w_2 is the weight between the hidden node and the output node, and b_2 is the bias. Without loss of generality, we also use sigmoid function as the activation function:

$$\sigma(x) = \frac{1}{1+e^{-x}}. \tag{3}$$

In the learning phase, NN aims to minimize the loss function between the estimated SM $\hat{y}$ and the corresponding ERA-Interim product y. This can be addressed by a classical back-propagation algorithm [40,41]. In this paper, we use the gradient backpropagation with the Levenberg–Marquardt algorithm [42]. In addition, we use the Min-Max normalization method to scale all data into the range [0,1] because it can avoid unnecessary numerical problems and make the network converge quickly. It is realized by the following:

$$d_{norm} = \frac{d - d_{min}}{d_{max} - d_{min}}, \tag{4}$$

where d is the original input data, d_{max} is the maximum value of the original input data, d_{min} is the minimum value of the original input data, and d_{norm} is the normalized value of d.

In the predicting phase, we can feed the given data to the input layer of NN, and acquire the output of the hidden layer, which is then used as an input of the output layer to get the estimated SM value.

3.2. Deep Convolutional Neural Network

Different from NN, deep NN (DNN) extends the number of hidden layers to two or more, thus improving the learning and generalization abilities of networks. As a kind of DNN, DCNN has been popularly employed in the field of remote sensing [43–45]. Inspired from them, we also design a DCNN for SM retrieval. The architecture of the designed network is shown in Figure 2, which consists of an input layer, three convolutional layers, one fully-connected layer, and an output layer. The same as NN, the input data of DCNN is also a 19-dimensional vector. Thus, we use a one-dimensional convolutional operator in each convolutional layer. The size and the number of convolutional kernel are 1×5 and 12, respectively. We select three convolutional layers because it can achieve better results than more numbers of convolutional layers, which will be discussed in Section 4.3.

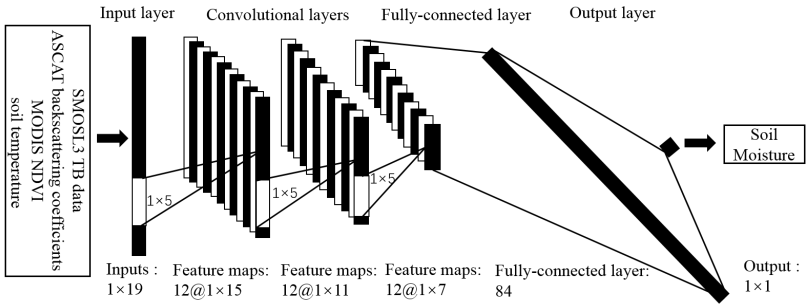


Figure 2. The architecture of the DCNN model.

The core module of DCNN is the convolutional operator. For example, the $(l+1)$-th layer output after the convolutional operator can be written as:

$$x^{l+1} = \phi(x^l * w^{l+1} + b^{l+1}), \tag{5}$$

where '$*$' is the convolutional operator, x^{l+1} is the output of the $(l+1)$-th layer, x^l is the input of the $(l+1)$-th layer, w^{l+1} is the filter weight in the $(l+1)$-th layer, and b^{l+1} is the bias in the $(l+1)$-th layer.

Different from NN, DCNN often uses Rectified Linear Units (ReLU) as the activation function ϕ, which can be defined as:

$$\phi(x) = max(0, x). \tag{6}$$

After the last convolutional layer, the feature maps are fed into the next layer via a fully-connected operator. The output h of the fully-connected layer can be described as:

$$h = \phi(x * w_1 + b_1), \tag{7}$$

where w_1 is the weight between the last convolutional layer and fully-connected layer, b_1 is the bias. After the fully-connected layer, the estimated SM $\hat{y}$ in the output layer can be obtained by:

$$\hat{y} = \phi(w_2 h + b_2), \tag{8}$$

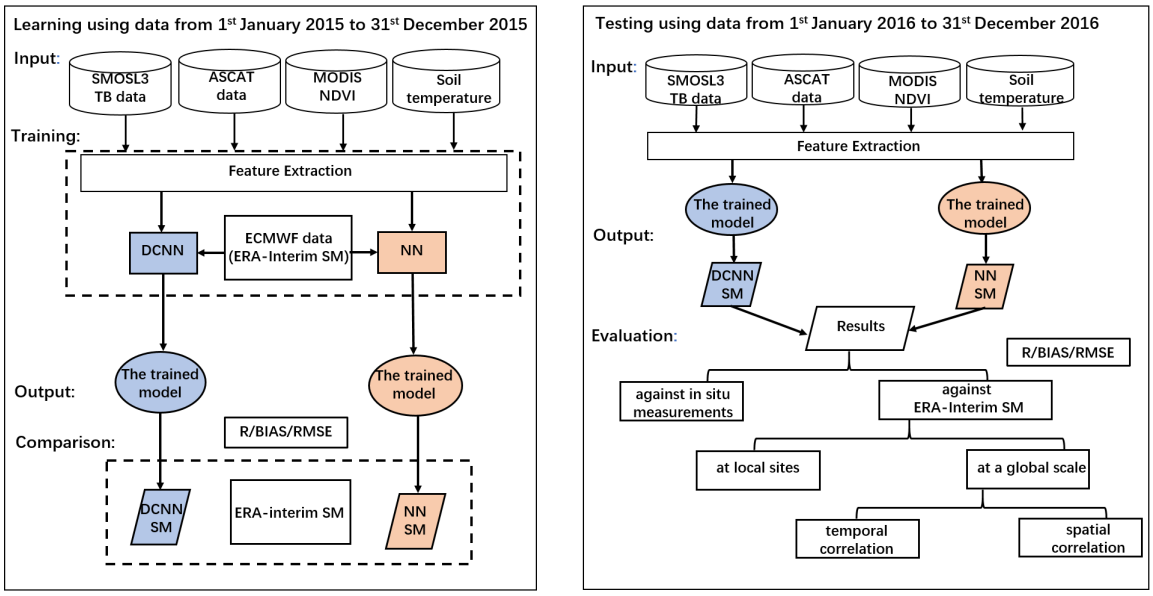
where w_2 is the weight between the the fully-connected layer and the output layer, b_2 is the bias. Note that the parameter optimization process of DCNN is the same as NN.

4. Experiments

In this section, we introduce the experimental settings, data preprocessing and hyperparameters selection of DCNN, and then compare the performance of DCNN model with the NN model.

4.1. Experimental Settings

As shown in Figure 3, our experiment consists of two phases: the learning phase from 1 January to 31 December 2015 and the testing phase from 1 January to 31 December 2016. In the learning phase, preprocessing is first conducted, and then hyperparameters are selected. After the training, SM estimates are obtained from NN and DCNN methods, respectively, and we compare their performance against ERA-Interim SM. In the testing phase, the trained models obtained from learning phases are applied to the testing data, and then we evaluate the retrieval quality of DCNN and NN by comparing with both in situ measurements and ERA-Interim SM. The evaluation against ERA-Interim SM also contains two parts, at the grid cells where the in situ stations are located and at a global scale.



(**a**) (**b**)

Figure 3. Flow chart of experiments. (**a**) learning phase; (**b**) testing phase.

Without loss of generality, we use three standard evaluation metrics: the Pearson's correlation coefficient R to evaluate the goodness of fit, the bias (mean retrieved SM minus mean target SM) $BIAS$ and the root mean square error $RMSE$ to evaluate the accuracy of the estimations:

$$R = \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}_i})(y_i - \overline{y_i})}{\sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}_i})^2 \sum_{i=1}^{n}(y_i - \overline{y_i})^2}, \tag{9}$$

$$BIAS = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n}, \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}, \tag{11}$$

where $\hat{y}_i$ and y_i are the retrieved SM value and the target SM value, respectively. n is the number of all retrieved SM values.

4.2. Data Preprocessing

Consistent with the work in [12,18], we conduct data filtering before training, for example, removing grids that might be affected by water bodies, frozen soil, RFI, etc. For the whole data set, grid cells with a latitude higher than 75° or lower than −60° are filtered out. Grid cells with frozen soil whose soil temperature is lower than 273.15 K or covered by snow are also filtered out. Grid cells with TB lower than 50 K or higher than 400 K are also filtered out. Additionally, SM values are very low and its uncertainties are possibly high in some regions, such as the Sahara desert. Grid cells corresponding to these areas are also removed for the whole data set. For the ECMWF data, grid points with SM values above 0.5 m^3/m^3 are filtered out because those grid points are highly possible to be water bodies or wet soil. More importantly, grid cells corresponding to the SCAN sites are also removed from this subset because they will be used to perform independent tests.

In summary, we collect a total of $\sim 6 \times 10^6$ data from 1 January 2015 to 31 December 2016, which includes SMOSL3 TB for incidence angles from 20° to 60°, ASCAT backscattering coefficients, MODIS NDVI, soil temperature and ERA-Interim SM. From this data set, data (about 3×10^6) from 1 January to 31 December 2015 are used to train models. The rest ones (about 3×10^6) from 1 January to 31 December 2016 are used to test the retrieval performance.

4.3. Hyperparameter Selection

A hyperparameter is a parameter that sets a value before a learning process, rather than the parameter obtained through learning process. For the DCNN model, there are four important hyperparameters need to be fixed, including the learning rate α, the size of batch ω, the size of filter $1 \times c$ and the number of convolutional layers l. We choose them from the following candidate sets: $\alpha \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$, $\omega \in \{20, 50, 100, 200\}$, $c \in \{4, 5, 6, 7\}$, $l \in \{2, 3, 4, 5\}$.

To show the effects of c on the retrieval performance, we fix other parameters and select c from $\{4, 5, 6, 7\}$. It is interesting to see that, as c increases, R first increases and then decreases when c exceeds 5 in Figure 4c. In contrast, $RMSE$ first decreases and then increases in Figure 4d. Therefore, the optimal size of filter is 1×5.

Similarly, the optimal number of convolutional layers $l = 3$ can be obtained from Figure 4e,f. Afterwards, we continue to train the DCNN model on the training data with the above hyperparameters, and then we will obtain the trained model of DCNN.

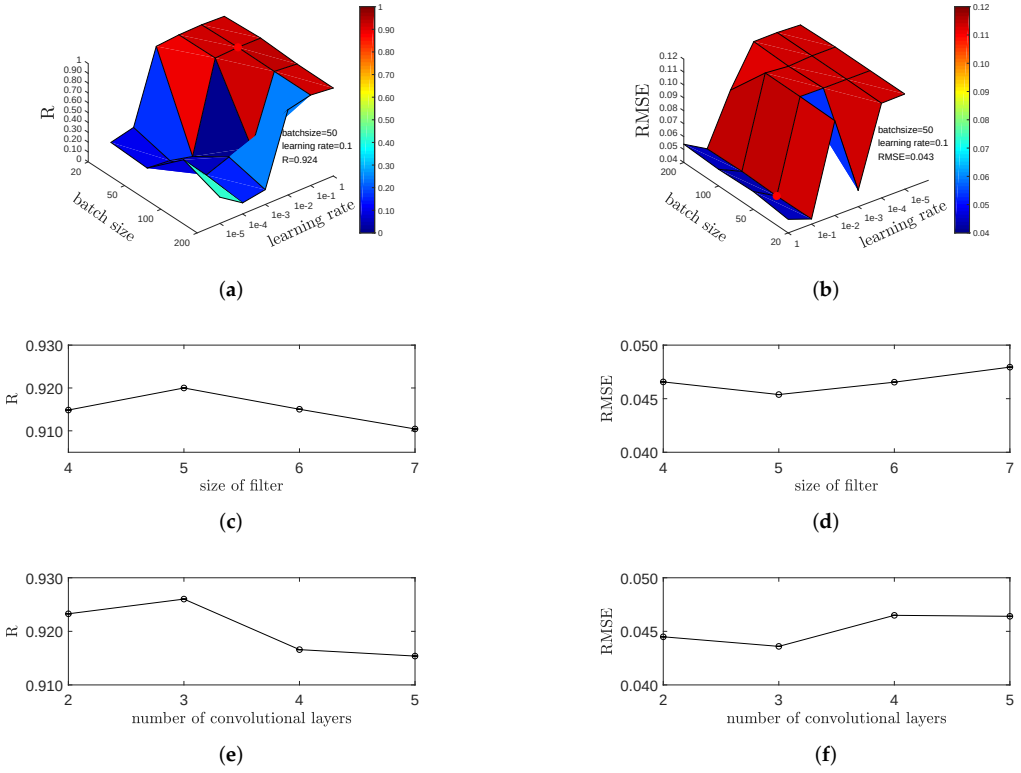


Figure 4. Effects of four hyperparameters on the retrieval performance on the same training data set. (**a**) *R* values achieved by applying different batch sizes and learning rates; (**b**) *RMSE* values achieved by applying different batch sizes and learning rates; (**c**) *R* values achieved by applying different sizes of filters; (**d**) *RMSE* values achieved by applying different sizes of filters; (**e**) *R* values achieved by applying different numbers of convolutional layers; (**f**) *RMSE* values achieved by applying different numbers of convolutional layers.

4.4. Performance Comparison

Our comparison analysis focuses on the testing phase, i.e., data from 1 January to 31 December 2016. We apply the trained DCNN/NN models to the testing data, and then compare the SM retrievals against both in situ measurements and the ERA-Interim SM.

First, we make a full validation against the in situ measurements from SCAN. In general, we compute the temporal correlation between different SM retrievals (DCNN SM, NN SM and ERA-Interim SM) and the in situ measurements in the testing phase. Here, the temporal correlation is the Pearson correlation coefficient of time series for each site. Specifically, we collect in situ SM measurements and the closest EASE grid cell of different SM retrievals at the time of SMOS overpasses to generate time series for each site. Note that only time series of different SM retrievals and the in situ measurements for each site are all available that can be used to compute the temporal correlation. During the experiment, we totally use 197 time series of the SCAN sites to evaluate the retrieval quality of DCNN/NN SM and ERA-Interim SM. For each site, R, $BIAS$ and $RMSE$ values between different SM retrievals and in situ measurements are computed. The min, mean, median and max values of the above metrics are recorded. Furthermore, we also compute R, $BIAS$, $RMSE$ between DCNN/NN SM and ERA-Interim SM over the grid cells where sites from SCAN are located.

Secondly, we evaluate DCNN/NN SM against ERA-Interim SM from the perspective of temporal dynamics. In order to validate whether DCNN and NN capture the temporal variability of the ERA-Interim SM, we compute R, $BIAS$ and $RMSE$ between the time series of DCNN/NN SM and ERA-Interim SM for each gird at a global scale.

Thirdly, we evaluate DCNN/NN SM against ERA-Interim SM from the perspective of spatial distribution. In order to validate whether NN and DCNN capture the spatial characteristics of ERA-Interim SM, we also compute the R, $BIAS$ and $RMSE$ between global SM maps at the annual and monthly time scale. Note that the global SM maps are transformed into 1D vectors before computing spatial correlation.

5. Results

5.1. Comparison against In Situ Measurements

When compared with in situ measurements, we compute the temporal correlation between different SM retrievals (DCNN SM, NN SM and ERA-Interim SM) and the in situ measurements in the testing phase. It is found that DCNN SM agrees better with in situ measurements than NN SM (Table 1). The mean value of R between DCNN SM and in situ measurements is higher than that between NN SM and in situ measurements by 6.2% for SMOS ascending overpass, while the mean value of R between DCNN SM and in situ measurements is higher than that between NN SM and in situ measurements by 2.5% for SMOS descending overpass (Table 1). Furthermore, box plots in Figure 5a show that the median values (the red line) of R between DCNN SM and in situ measurements are also slightly higher than those between NN SM and in situ measurements. Figure 5b,c demonstrate similar findings as Figure 5 from the other evaluation metrics $BIAS$ and $RMSE$.

In addition, significance tests are conducted between the temporal correlation of DCNN SM and that of NN SM. According to the p-value ($p < 0.05$) obtained by the significance test method, there is a statistically significant difference between the R value of DCNN SM and NN SM. In summary, the comparison against in situ measurements shows the good performance of DCNN/NN SM, and DCNN SM gives better results than NN SM. In addition, the performance of DCNN/NN SM retrievals for ascending overpass is generally better than the descending overpass. This may be caused by the convective rainfall that usually takes place in the afternoon.

Table 1. The min, mean, median and max values of R, $BIAS$, $RMSE$ between the time series of DCNN/NN SM and ERA-Interim SM against in situ measurements from SCAN sites in the testing phase during 1 January to 31 December 2016. Data from ascending and descending orbits are processed separately.

	ASCENDING Orbit			**DESCENDING Orbit**		
	DCNN SM	**NN SM**	**ERA-Interim SM**	**DCNN SM**	**NN SM**	**ERA-Interim SM**
			R			
MIN	0.117	0.111	0.102	0.109	0.103	0.107
MEAN	0.564	0.531	0.556	0.530	0.517	0.562
MEDIAN	0.593	0.548	0.567	0.544	0.500	0.573
MAX	0.966	0.947	0.955	0.953	0.955	0.967
			BIAS			
MIN	0.000	0.009	0.000	0.000	0.001	0.002
MEAN	0.036	0.097	0.035	0.037	0.107	0.040
MEDIAN	0.033	0.096	0.028	0.030	0.111	0.034
MAX	0.129	0.211	0.128	0.115	0.238	0.120
			RMSE			
MIN	0.035	0.067	0.033	0.025	0.065	0.019
MEAN	0.099	0.140	0.098	0.102	0.156	0.102
MEDIAN	0.098	0.138	0.097	0.101	0.154	0.101
MAX	0.168	0.229	0.180	0.185	0.263	0.183

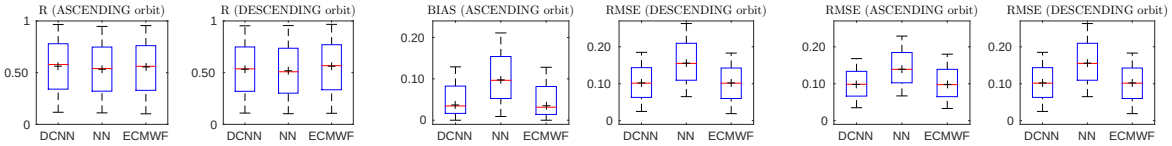


Figure 5. Box plots for (**a**) Pearson correlation coefficient, (**b**) $BIAS$ and (**c**) $RMSE$ between DCNN SM/NN SM/ERA-Interim SM time series and the in situ measurements. The blue boxes contain the middle 50% of the data, and the red line represents the median value of the distribution. The upper edge of the box indicates the 75th percentile of the data (Q3), and the lower edge indicates the 25th percentile (Q1). The mean values are also shown as black crosses. The upper and lower bars represent the minimum and maximum values of the distribution excluding outliers. Points are considered as outliers if they are larger than Q3 + 1.5 × IQR or smaller than Q1−1.5 × IQR, where inter quartile range (IQR) means Q3−Q1.

5.2. Temporal Correlation

Table 2 shows R, $BIAS$, $RMSE$ values between DCNN/NN SM and ERA-Interim SM over the grid cells where sites from SCAN are located in the testing phase. It is found that the mean values of $RMSE$ between DCNN SM and ERA-Interim SM are lower than that between NN SM and ERA-Interim SM by 67.6% and 69.7% for ascending and descending overpass, respectively. It is also found that the mean values of $BIAS$ between DCNN SM and ERA-Interim SM are lower than that between NN SM and ERA-Interim SM by 88.9% and 90.4% for ascending and descending overpass, respectively. Additionally, the mean value of R between the time series of DCNN SM and ERA-Interim SM is higher than that between NN SM and in situ measurements by 0.027 for ascending overpass, and the mean values of R between the time series of DCNN SM and ERA-Interim SM is higher than that between NN SM and in situ measurements by 0.05 for descending overpass.

Table 2. The min, mean, median and max values of R, $BIAS$, $RMSE$ between DCNN/NN SM and ERA-Interim SM in the testing phase. Data from ascending and descending orbits are processed separately.

	ASCENDING Orbit		**DESCENDING Orbit**	
	DCNN SM	**NN SM**	**DCNN SM**	**NN SM**
		R		
MIN	0.635	0.606	0.573	0.515
MEAN	0.663	0.636	0.600	0.550
MEDIAN	0.672	0.644	0.599	0.550
MAX	0.682	0.656	0.617	0.565
		BIAS		
MIN	0.009	0.090	0.000	0.102
MEAN	0.010	0.090	0.010	0.104
MEDIAN	0.011	0.094	0.011	0.104
MAX	0.011	0.096	0.022	0.106
		RMSE		
MIN	0.035	0.108	0.035	0.119
MEAN	0.036	0.111	0.037	0.122
MEDIAN	0.036	0.111	0.037	0.122
MAX	0.037	0.112	0.038	0.124

Furthermore, the SM estimates from DCNN and NN models in the testing period are compared with ERA-Interim SM at a global scale (Table 3). It is found that the mean values of $RMSE$ between DCNN SM and ERA-Interim SM are lower than that between NN SM and ERA-Interim SM by 81.9%

and 79.6% for ascending and descending overpass, respectively. It is also found that the mean values of $BIAS$ between DCNN SM and ERA-Interim SM are lower than that between NN SM and ERA-Interim SM by 92.4% and 89.4% for ascending and descending overpass, respectively. The aforementioned findings indicate that $BIAS$ and $RMSE$ can be improved by using DCNN when compared with the ERA-Interim SM at a global scale. In addition, the mean values of R between DCNN SM and ERA-Interim SM are higher than that between NN SM and ERA-Interim SM by 0.006 and 0.01 for ascending and descending overpass respectively, which shows that DCNN can keep better temporal variability of ERA-Interim than NN.

Table 3. Temporal correlation between DCNN/NN SM and ERA-Interim SM in testing phase. Data from ascending and descending orbits are processed separately.

	ASCENDING Orbit		**DESCENDING Orbit**	
	DCNN SM	**NN SM**	**DCNN SM**	**NN SM**
R	0.576	0.570	0.568	0.558
$BIAS$	0.014	0.184	0.019	0.180
$RMSE$	0.036	0.199	0.040	0.196

Figure 6 demonstrates the spatial distributions of R, $BIAS$ and $RMSE$ between DCNN/NN SM and ERA-Interim SM. From Figure 6a,b, we can see that most regions have high positive correlations, especially over western North America, eastern South America, central Asia, southern Africa and southern Australia, while some regions, e.g., northern Africa and northern Australia, have relatively low positive correlations. In addition, from the difference of R between DCNN SM and NN SM in Figure 6c, it is found that DCNN achieves slightly better performance than NN in some regions, such as western North America, eastern South America, central Asia and Australia. It is also observed in Figure 7 that these regions have relative longer time series than other regions.

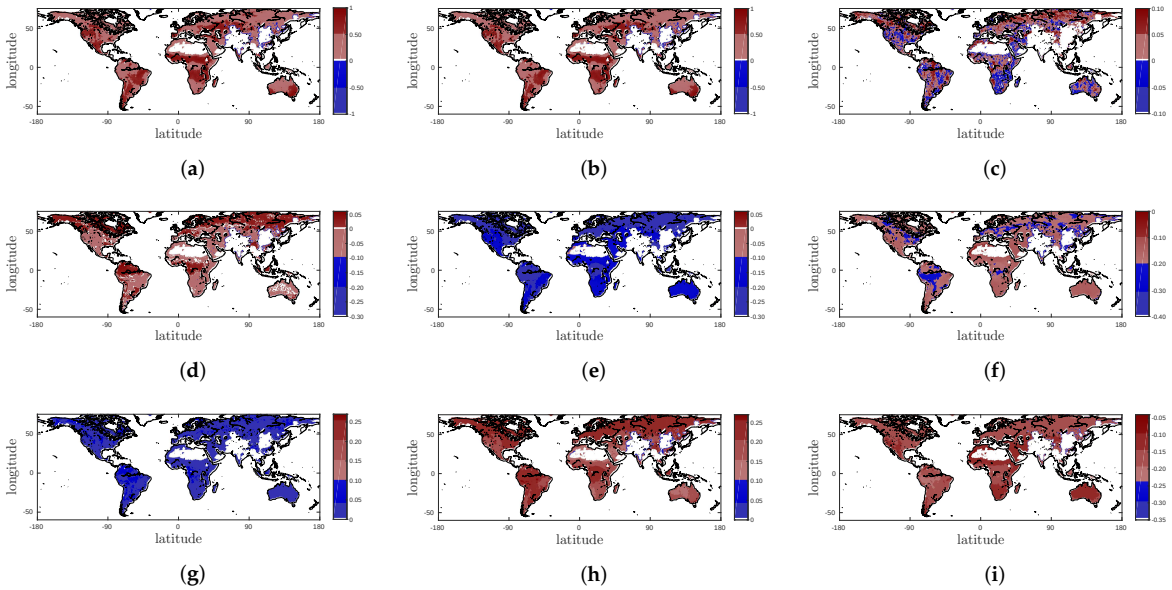


Figure 6. Spatial distribution of R (the top panels), $BIAS$ (the middle panels), $RMSE$ (the bottom panels) between the time series of DCNN/NN SM and ERA-Interim SM for SMOS ascending overpass in the testing phase. (**a**) R between DCNN SM and ERA-Interim SM; (**b**) R between NN SM and ERA-Interim SM; (**c**) difference in R (a minus b); (**d**) BIAS between DCNNM SM and ERA-Interim SM; (**e**) $BIAS$ between NN SM and ERA-Interim SM; (**f**) difference in $BIAS$ ($|d|$ minus $|e|$); (**g**) $RMSE$ between DCNN SM and ERA-Interim SM; (**h**) RMSE between NN SM and ERA-Interim SM; (**i**) difference in $RMSE$ (g minus h).

Comparing Figure 6d,e, we can see that bias between DCNN/NN SM and ERA-Interim SM are different over the same region. The bias between DCNN SM and ERA-Interim SM in most regions ranges from −0.10 to 0.05 (in rust red), whereas the bias between NN SM and ERA-Interim SM in most regions ranges from −0.30 to 0.10 (in blue). In addition, from the difference of absolute $BIAS$ between DCNN SM and NN SM in Figure 6f, it is observed that the bias between DCNN SM and ERA-Interim SM is smaller than that between NN SM and ERA-Interim SM.

As clearly shown in Figure 6g,h, the values of $RMSE$ between DCNN/NN SM and ERA-Interim SM are small. The values of $RMSE$ between DCNN SM and ERA-Interim SM are relatively smaller than that between NN SM and ERA-Interim (Figure 6i), which indicates that DCNN is more effective in retrieving SM.

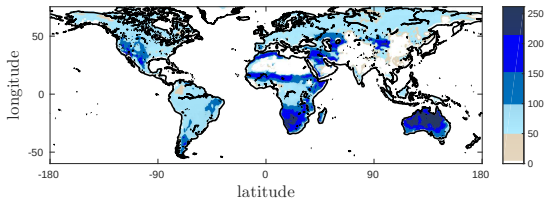


Figure 7. Numbers of training samples for each grid for SMOS ascending overpass in the training phase.

Additionally, according to the previous study in [12], there are some connection between the temporal correlation and the variance of the time series for the target data. From Figures 6a,b and 8, we find that regions with a lower temporal correlation are those areas whose variance of ERA-Interim SM time series are relatively low, while regions with a higher temporal correlation are those regions whose variance of ERA-Interim SM time series are relatively high.

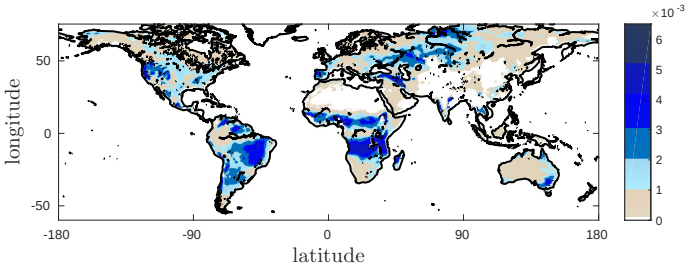


Figure 8. Spatial distribution for variance of the time series for ERA-Interim SM for ascending overpass in the testing phase.

5.3. Spatial Correlation

Table 4 lists spatial correlation between DCNN/NN SM and ERA-Interim SM. It can be observed that the $BIAS$ and $RMSE$ values between DCNN SM and ERA-Interim SM are similar to those between NN SM and ERA-Interim, but R values between DCNN SM and ERA-Interim SM are slightly higher than those between NN SM and ERA-Interim SM. Figure 9 shows the spatial distribution of yearly average DCNN SM, NN SM and ERA-Interim SM for ascending overpass in the testing phase. From a global point of view, we can observe that DCNN and NN can achieve similar results to ERA-Interim, which indicates the effectiveness of DCNN and NN retrieval methods. From a local point of view, we can see that DCNN can achieve a little better result than NN when compared to ERA-Interim, especially in western North America and eastern South America.

Table 4. Spatial correlation between DCNN/NN SM and ERA-interim SM in testing phase from 1 January to 31 December 2016. Data from ascending and descending orbits are processed separately.

	ASCENDING Orbit		**DESCENDING Orbit**	
	DCNN SM	**NN SM**	**DCNN SM**	**NN SM**
R	0.927	0.924	0.926	0.922
BIAS	0.006	0.006	0.009	0.009
RMSE	0.043	0.043	0.045	0.048

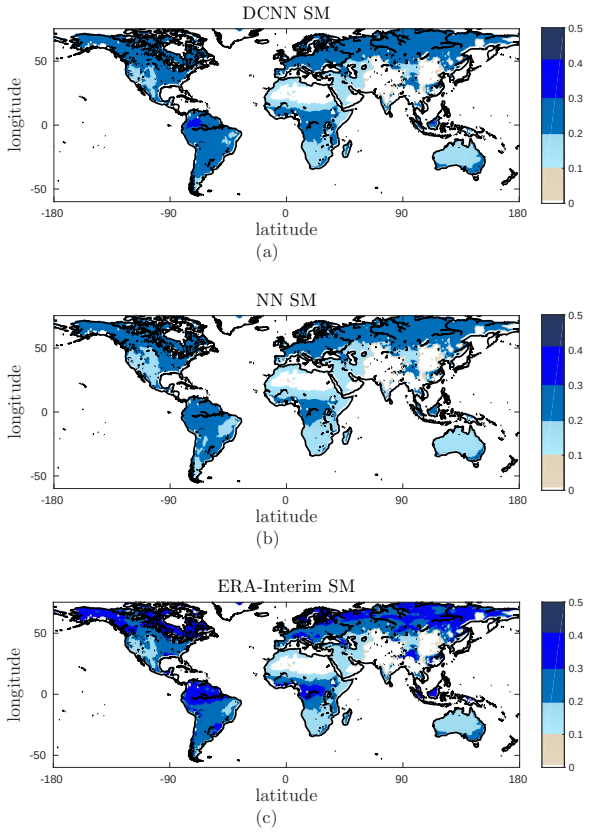


Figure 9. Spatial distribution of the yearly average SM values by DCNN and NN and the corresponding ERA-Interim SM for ascending overpass in testing phase from 1 January to 31 December 2016. (**a**) yearly average DCNN SM; (**b**) yearly average NN SM; (**c**) yearly average ERA-Interim SM.

Although similar spatial correlations between DCNN/NN SM and ERA-interim SM is found at the annual scale, their agreements vary in different months shown in Figure 10. First, the number of data used in each month is different. Specifically, the data in January, February, November, December are apparently fewer than those in other months. Second, DCNN achieves better retrieval results than NN in February, May, September, October, November and December; NN achieves better retrieval results in January, July and August; DCNN and NN achieve similar results in March, April, and June.

In summary, DCNN and NN obtain similar retrieval results as compared to ERA-interim SM, while DCNN SM is closer to ERA-interim SM than NN SM at a global scale. In addition, DCNN achieves better or comparable retrieval results than NN in most times of the year 2016, which suggests that the retrieval performance of DCNN is slightly better than that of NN.

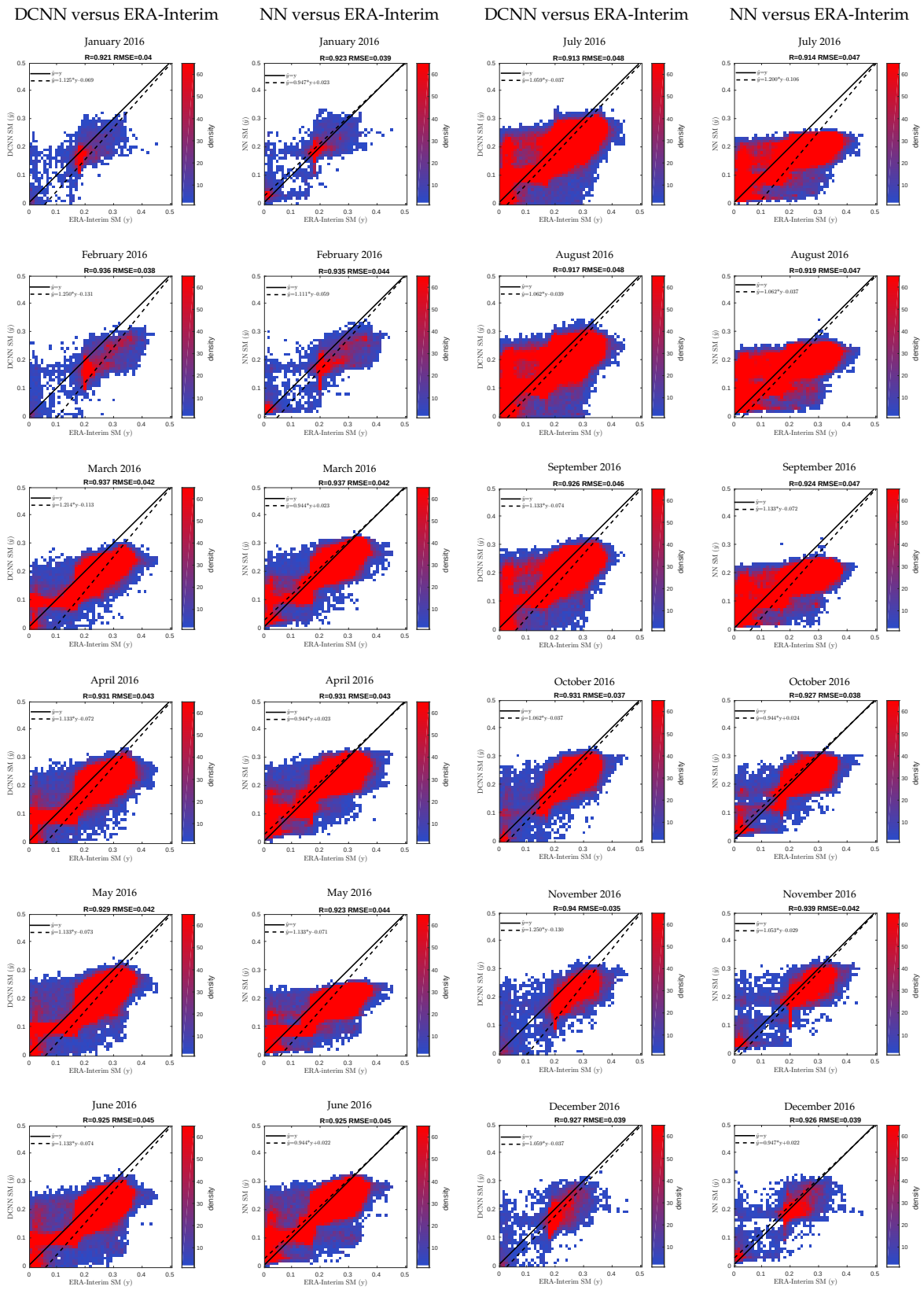


Figure 10. Scatter plot in density of the retrieved SM by DCNN versus ERA-Interim SM (the first and third column) and NN versus ERA-Interim SM (the second and fourth column) when SMOS ascending overpass in the testing phase. The solid line represents the ideal retrieval results, while the dashed line represents the regression.

6. Discussion

DCNN has been widely used in the field of image recognition, speech recognition and computer vision fields [46–48], where it can achieve better performance than the traditional NN model. Recently, it has also been introduced to remote sensing data analysis, such as remote sensing scene classification [43–45] and object detection [49–51]. However, its superiority has not been explored in SM retrieval. In this paper, we attempt using it to estimate SM from satellite observations and compare its performance with NN.

The experimental results show that the temporal and spatial correlations between NN (DCNN) SM and ERA-Interim SM are 0.558–0.570 (0.568–0.576) and 0.922–0.924 (0.926–0.927), respectively. These results indicate that DCNN is slightly better than NN in SM retrieval, especially in the regions (such as western North America, eastern South America, central Asia and Australia) where there exist more numbers of training samples. This can be explained by the fact that DCNN has more powerful learning ability.

Nevertheless, the DCNN model has a more complex structure than NN model. As a result, it usually costs much more time to estimate SM than NN. For example, when the trained NN and DCNN models are applied to the data set in September 2016 which contains about 5.8×10^5 samples, NN needs about 10 s to retrieve SM values while DCNN needs about 60 s. Note that our experiments were carried on a personal computer (Intel Core 3.60 GHz processor with 32 GB random access memory). The software implementation was performed using MATLAB(MathWorks, Inc.).

In addition, the simulated SM values by DCNN and NN are easily affected by the target data during the training period. In our study, we select the ERA-Interim product as the target data. ERA-Interim SM is a reliable SM estimate, but it still exists some deviation which may directly affect the performance of models. In addition, we only use the in situ measurements from SCAN sites, which may exist uncertainties when evaluated against different SM retrievals. In future research, we will use other reliable SM estimates and more in situ measurements to further evaluate the performance of DCNN.

7. Conclusions

In this paper, we compare the performance of DCNN and NN in estimating SM from satellite observations. The same input data are used for the DCNN and NN models, including L-band SMOSL3 TB data, ASCAT backscattering coefficients, MODIS NDVI and soil temperature. The target SM data used to train the DCNN and NN model is the ERA-Interim product. We apply the trained DCNN/NN models to the testing data, and then compare the different SM retrievals against both in situ measurements and the ERA-Interim SM. The experimental results indicate that DCNN works very well and gives better results than NN when evaluated against in situ measurements from SCAN sites in the testing phase. Furthermore, the estimated SM values by DCNN/NN and ERA-Interim SM agree relatively well at a global scale. In summary, both NN and DCNN methods are effective in estimating SM from satellite observations and the performance of DCNN is slightly better than NN.

Author Contributions: Q.L. proposed the algorithm. R.H. and L.G. conceived and designed the experiments. L.G. performed the experiments. R.H. and Y.L. supervised the study, analyzed the results and gave insightful suggestions for the manuscript. R.H. and L.G. drafted the manuscript. All coauthors contributed to the revision of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61532009, in part by the Foundation of Jiangsu Province of China under Grant 18KJB520032, and in part by the Nanjing University of Information Science and Technology (NUIST) startup grant 2243141701020.

Acknowledgments: The SMOS data in this paper come from the "Centre Aval de Traitement des Données SMOS" (CATDS), operated for the "Centre National d'Etudes Spatiales" (CNES, France) by IFREMER (Brest, France). The authors would like to thank NASA Earth Observing System Data and Information System (EOSDIS) Land

Processes Distributed Active Archive Center (LP DAAC) for providing MODIS NDVI MYD13C1 product at https://doi.org/10.5067/MODIS/MYD13C1.006.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Secretariat, G. Implementation plan for the global observing system for climate in support of the UNFCCC (2010 Update). In Proceedings of the Conference of the Parties (COP), Copenhagen, Denmark, 7–19 December 2009; 150p.
2. Gruber, A.; Paloscia, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Smolander, T.; Pulliainen, J.; Mittelbach, H.; Dorigo, W.; Wagner, W. Performance inter-comparison of soil moisture retrieval models for the MetOp-A ASCAT instrument. Geoscience and Remote Sensing Symposium (IGARSS). In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 2455–2458.
3. Hirschi, M.; Mueller, B.; Dorigo, W.; Seneviratne, S.I. Using remotely sensed soil moisture for land–Atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability. *Remote Sens. Environ.* **2014**, *154*, 246–252. [CrossRef]
4. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.J.; Font, J.; Reul, N.; Gruhier, C.; et al. The SMOS mission: New tool for monitoring key elements ofthe global water cycle. *Proc. IEEE* **2010**, *98*, 666–687. [CrossRef]
5. Wagner, W.; Blöschl, G.; Pampaloni, P.; Calvet, J.C.; Bizzarri, B.; Wigneron, J.P.; Kerr, Y. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. *Hydrol. Res.* **2007**, *38*, 1–20. [CrossRef]
6. Kerr, Y.H. Soil moisture from space: Where are we? *Hydrogeol. J.* **2007**, *15*, 117–120. [CrossRef]
7. Liu, S.F.; Liou, Y.A.; Wang, W.J.; Wigneron, J.P.; Lee, J.B. Retrieval of crop biomass and soil moisture from measured 1.4 and 10.65 GHz brightness temperatures. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1260–1268.
8. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Martinuzzi, J.; Font, J.; Berger, M. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1729–1735. [CrossRef]
9. McMullan, K.; Brown, M.A.; Martín-Neira, M.; Rits, W.; Ekholm, S.; Marti, J.; Lemanczyk, J. SMOS: The payload. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 594–605. [CrossRef]
10. Jackson, T.J.; Bindlish, R.; Cosh, M.H.; Zhao, T.; Starks, P.J.; Bosch, D.D.; Seyfried, M.; Moran, M.S.; Goodrich, D.C.; Kerr, Y.H.; et al. Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the US. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1530–1543. [CrossRef]
11. Kerr, Y.H.; Waldteufel, P.; Richaume, P.; Wigneron, J.P.; Ferrazzoli, P.; Mahmoodi, A.; Al Bitar, A.; Cabot, F.; Gruhier, C.; Juglea, S.E.; et al. The SMOS soil moisture retrieval algorithm. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1384–1403. [CrossRef]
12. Rodríguez-Fernández, N.J.; Aires, F.; Richaume, P.; Kerr, Y.H.; Prigent, C.; Kolassa, J.; Cabot, F.; Jiménez, C.; Mahmoodi, A.; Drusch, M. Soil moisture retrieval using neural networks: Application to SMOS. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5991–6007. [CrossRef]
13. Albergel, C.; De Rosnay, P.; Gruhier, C.; Muñoz-Sabater, J.; Hasenauer, S.; Isaksen, L.; Kerr, Y.; Wagner, W. Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sens. Environ.* **2012**, *118*, 215–226. [CrossRef]
14. Panciera, R.; Walker, J.P.; Kalma, J.D.; Kim, E.J.; Saleh, K.; Wigneron, J.P. Evaluation of the SMOS L-MEB passive microwave soil moisture retrieval algorithm. *Remote Sens. Environ.* **2009**, *113*, 435–444. [CrossRef]
15. Santi, E.; Paloscia, S.; Pettinato, S.; Fontanelli, G. A prototype ANN based algorithm for the soil moisture retrieval from L-band in view of the incoming SMAP mission. In Proceedings of the 2014 13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad), Pasadena, CA, USA, 24–27 March 2014; pp. 5–9.
16. Prigent, C.; Aires, F.; Rossow, W.B.; Robock, A. Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: Relationship of satellite observations to in situ soil moisture measurements. *J. Geophys. Res. Atmos.* **2005**, *110*. [CrossRef]
17. Aires, F.; Prigent, C.; Rossow, W. Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships. *J. Geophys. Res. Atmos.* **2005**, *110*.. [CrossRef]

18. Kolassa, J.; Aires, F.; Polcher, J.; Prigent, C.; Jimenez, C.; Pereira, J.M. Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology. *J. Geophys. Res. Atmos.* **2013**, *118*, 4847–4859. [CrossRef]
19. Rodríguez-Fernández, N.J.; Sabater, J.M.; Richaume, P.; de Rosnay, P.; Kerr, Y.H.; Albergel, C.; Drusch, M.; Mecklenburg, S. SMOS near-real-time soil moisture product: processor overview and first validation results. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5201–5216. [CrossRef]
20. Liu, Q.; Zhou, F.; Hang, R.; Yuan, X. Bidirectional-Convolutional LSTM Based Spectral-Spatial Feature Learning for Hyperspectral Image Classification. *Remote Sens.* **2017**, *9*, 1330.
21. Liu, Q.; Hang, R.; Song, H.; Li, Z. Learning Multiscale Deep Features for High-Resolution Satellite Image Scene Classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 117–126. [CrossRef]
22. Berthon, L.; Mialon, A.; Bitar, A.A.; Cabot, F.; Kerr, Y. SMOS CATDS level 3 Soil Moisture products. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 22–27 April 2012; Volume 14, p. 5456.
23. Oliva, R.; Daganzo, E.; Kerr, Y.H.; Mecklenburg, S.; Nieto, S.; Richaume, P.; Gruhier, C. SMOS radio frequency interference scenario: Status and actions taken to improve the RFI environment in the 1400–1427-MHz passive band. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1427–1439. [CrossRef]
24. Kerr, Y.H.; Al-Yaari, A.; Rodriguez-Fernandez, N.; Parrens, M.; Molero, B.; Leroux, D.; Bircher, S.; Mahmoodi, A.; Mialon, A.; Richaume, P.; et al. Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation. *Remote Sens. Environ.* **2016**, *180*, 40–63. [CrossRef]
25. Gelsthorpe, R.; Schied, E.; Wilson, J. ASCAT-Metop's advanced scatterometer. *ESA Bull.* **2000**, *102*, 19–27.
26. Bartalis, Z.; Wagner, W.; Naeimi, V.; Hasenauer, S.; Scipal, K.; Bonekamp, H.; Figa, J.; Anderson, C. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.* **2007**, *34*. [CrossRef]
27. Brocca, L.; Hasenauer, S.; Lacava, T.; Melone, F.; Moramarco, T.; Wagner, W.; Dorigo, W.; Matgen, P.; Martínez-Fernández, J.; Llorens, P.; et al. Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sens. Environ.* **2011**, *115*, 3390–3408. [CrossRef]
28. Wagner, W.; Lemoine, G.; Borgeaud, M.; Rott, H. A study of vegetation cover effects on ERS scatterometer data. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 938–948. [CrossRef]
29. Lindsley, R.D.; Long, D.G. Enhanced-resolution reconstruction of ASCAT backscatter measurements. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2589–2601. [CrossRef]
30. Paloscia, S.; Pettinato, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Reppucci, A. Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sens. Environ.* **2013**, *134*, 234–248. [CrossRef]
31. Berrisford, P.; Dee, D.; Poli, P.; Brugge, R.; Fielding, K.; Fuentes, M.; Kallberg, P.; Kobayashi, S.; Uppala, S.; Simmons, A. *The ERA-Interim Archive Version 2.0*; ERA Report Series 1; ECMWF: Shinfield Park, Reading, UK, 2011; 23p.
32. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [CrossRef]
33. Balsamo, G.; Albergel, C.; Beljaars, A.; Boussetta, S.; Brun, E.; Cloke, H.; Dee, D.; Dutra, E.; Muñoz-Sabater, J.; Pappenberger, F.; et al. ERA-Interim/Land: A global land surface reanalysis data set. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 389–407. [CrossRef]
34. Zhou, J.; Wen, J.; Wang, X.; Jia, D.; Chen, J. Analysis of the Qinghai-Xizang Plateau monsoon evolution and its linkages with soil moisture. *Remote Sens.* **2016**, *8*, 493. [CrossRef]
35. Viterbo, P.; Beljaars, A.C. An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Clim.* **1995**, *8*, 2716–2748. [CrossRef]
36. Viterbo, P.; Betts, A.K. Impact on ECMWF forecasts of changes to the albedo of the boreal forests in the presence of snow. *J. Geophys. Res. Atmos.* **1999**, *104*, 27803–27810. [CrossRef]
37. Schaefer, G.L.; Cosh, M.H.; Jackson, T.J. The USDA natural resources conservation service soil climate analysis network (SCAN). *J. Atmos. Ocean. Technol.* **2007**, *24*, 2073–2077. [CrossRef]

38. Al Bitar, A.; Leroux, D.; Kerr, Y.H.; Merlin, O.; Richaume, P.; Sahoo, A.; Wood, E.F. Evaluation of SMOS soil moisture products over continental US using the SCAN/SNOTEL network. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1572–1586. [CrossRef]
39. Kolassa, J.; Gentine, P.; Prigent, C.; Aires, F. Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis. *Remote Sens. Environ.* **2016**, *173*, 1–14. [CrossRef]
40. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
41. Rumelhart, D.E.; Durbin, R.; Golden, R.; Chauvin, Y. Backpropagation: The basic theory. In *Backpropagation: Theory, Architectures and Applications*; Psychology Press: London, UK, 1995; pp. 1–34.
42. Moré, J.J. The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical Analysis*; Springer: Berlin/Heidelberg, Germany, 1978; pp. 105–116.
43. Scott, G.J.; England, M.R.; Starms, W.A.; Marcum, R.A.; Davis, C.H. Training deep convolutional neural networks for land–cover classification of high-resolution imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 549–553. [CrossRef]
44. Hu, F.; Xia, G.S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* **2015**, *7*, 14680–14707. [CrossRef]
45. Castelluccio, M.; Poggi, G.; Sansone, C.; Verdoliva, L. Land use classification in remote sensing images by convolutional neural networks. *arXiv* **2015**, arXiv:1508.00092.
46. LeCun, Y.; Huang, F.J.; Bottou, L. Learning methods for generic object recognition with invariance to pose and lighting. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Related Articles, Washington, DC, USA, 27 June–2 July 2004; IEEE: Piscataway, NJ, USA; Volume 2, pp. II–104.
47. Abdel-Hamid, O.; Mohamed, A.-r.; Jiang, H.; Penn, G. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 25–30 March 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 4277–4280.
48. Lawrence, S.; Giles, C.L.; Tsoi, A.C.; Back, A.D. Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Netw.* **1997**, *8*, 98–113. [CrossRef] [PubMed]
49. Cai, Z.; Fan, Q.; Feris, R.S.; Vasconcelos, N. A unified multi-scale deep convolutional neural network for fast object detection. In Proceedings of the 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 354–370.
50. Maturana, D.; Scherer, S. Voxnet: A 3d convolutional neural network for real-time object recognition. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 922–928.
51. Szegedy, C.; Toshev, A.; Erhan, D. Deep neural networks for object detection. In *Advances in Neural Information Processing Systems*; NIPS Foundation: Montreal, CA, USA, 2013; pp. 2553–2561.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).