



Cloud detection in satellite images with classical and deep neural network approach: A review

Rachana Gupta¹ · Satyasai Jagannath Nanda¹

Received: 7 December 2020 / Revised: 28 July 2021 / Accepted: 4 January 2022 /
Published online: 11 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

This article introduces a review on implementations of various methods to perform cloud detection and its related applications such as detection of cloud shadow, types of cloud and cloud removal from multi-spectral satellite images. The cloud detection concept started with the basic sensitive parameters of clouds. These parameters have been reported based on albedo, spectral and textural parameters in the decade of 1980. With this parameters, a new era of Neural Network (NN) approach has been stimulated from 1992 for cloud classification. A summary of their empirical results are provided for various published works based on NN approach from 1970 to 2020. Moreover, the present article embodies experimental analysis of cloud detection using NN based classifier on multi-spectral satellite images with distinguish mathematical model of learning rules and number of hidden layers. The analysis is verified on L8, AVHRR, NOAA and GOES satellite images. The result demonstrates improved performance of NN approach with two layer perceptron architecture with Levenberg-Marquardt learning rule for cloud detection in terms of ellapse time. However, potential of self-organizing feature map (SOFM), an unsupervised NN approach, is observed in terms of accuracy over supervised learning architecture. The cloud detection algorithm is further discussed with convolutional neural network (CNN) as a deep learning algorithm to extract the local and global features from limited number of spectral bands to raise the performance accuracy of the approach.

Keywords Cloud detection · Textural feature · Neural network · Evolutionary algorithm · Convolutional neural network · Remote sensing

✉ Rachana Gupta
2015rec9517@mnit.ac.in

Satyasai Jagannath Nanda
sjnanda.ece@mnit.ac.in

¹ Department of Electronics and Communication Engineering, Malaviya National Institute of Technology, Jaipur 302017, Rajasthan, India

1 Introduction

The open data policy under Copernicus programme from NASA has acquired extensive attention in the field of remote sensing in 2008. The Copernicus programme began to provide the free Landsat satellite data. This is to combine the satellite data with the analytic tools in order to increase the economic growth. However; inspite of having larger modified softwares to pre-process the satellite data, there are problems where cloud covered regions often trouble the analysis [73]. Therefore, cloud detection is found to be an important initial approach to get the information regarding climatic studies and different properties of earth and atmosphere. This is due to the fact that change in cloud, affects the energy-budget between earth and atmosphere and hence affects the water exchange that describes the climate [12, 27, 108]. Moreover, cloud and its shadow obscures the specific scene more brighter and darker, respectively which affects various tasks such as crop monitoring, classification of marine habitats, land use region monitoring, environmental monitoring, geographic mapping and target recognition. Therefore, it is a vital pre-processing step to correctly and efficiently classify clouds, before satellite imagery used for further analysis. For this reason that cloud' statistics were utilized for atmospheric research since the first remote sensing imagery came into picture [4, 180].

Most of the methods based on cloud' statistics are observed to be focused around spectral signature using pixel-to-pixel processing. Distinct spectral features of clouds are observed in satellite imagery depend on its brighter and colder characteristics. This leads to misclassification of cloud with pixels having similar spectral signatures, for example, sand in deserts and snow/ice. Moreover, clouds having diverse shape also hampers the accuracy of its detection [31]. Therefore, various studies have been employed using physical parameters of cloud like cloud density, shape of clouds, optical properties and brightness temperature. However, higher spectral heterogeneity of clouds and its larger temperature variability with underlying surfaces cause difficult to extract cloud contaminated pixels from satellite images. The main reason of this difficulty is the utilization of threshold based approach which is sensitive in various atmospheric conditions for the detection and removal of distinguish characteristics of cloud. These characteristics of cloud are divided into low-level, mid-level and high-level based on its height from ground level, represented in Fig. 1.

Cirrocumulus, cirrus and cirrostratus are high-level clouds. Cirrocumulus cloud can be seen in winter indicates fair but cold weather. Its appearance is small rounded puffs, long rows and usually find in white or gray in color. Similarly, cirrus cloud indicates fair weather and can be seen as long, thin, wispy white stream. Whereas, cirrostratus cloud can be found in 12 to 24 hours before rainfall or snowstorm. The another characteristics of cloud comes under mid-level which includes altocumulus and altostratus. It generally extend over the entire sky like a thin sheet of cloud. Altocumulus cloud produces in group as grayish-white with 1 km thick and one portion darker as compare to other. It indicates warm or humid morning and thunderstorm in late noon time. While altostratus indicates a storm with rain or might be snow on its way. They can be seen as gray or blue-gray cloud and generally expanded in the entire sky. Cumulonimbus cloud with vertical growth of 10 km high indicates the thunderstorm which is associated with heavy rain, lightening, snow and sometimes tornado. In the presence of cumulonimbus cloud, if rainfall hits to the earth's surface then it converts into nimbostratus cloud. Stratus, cumulus, and stratocumulus are the other categories that come under low-level cloud. Stratus cloud appears gray in color like fog and cover almost entire sky and don't reach the earth's surface. It indicates fall of little mist or drizzle. Cumulus cloud in the form of cotton balls with sharp outlines, flat base and with



Fig. 1 Cloud chart based on their heights from the ground level [119]

vertical growth show fair or stormy weather. Stratus cloud with low and lumpy shape with gray color indicates light rainfall [67].

The only problem in detecting distinguish characteristics of cloud is its responses towards the change in climate are complicated and hence tough to compute the energy as well as water balance [72, 108, 139]. Therefore, mechanism of cloud is an essential area of study to incorporate more precise representation of its behavior. The number of analysis are done over seasonal cloud patterns and weather conditions but their relativity to earth radiation budget (ERB) calculation is limited [194]. The accurate understanding regarding mechanics of cloud formation and its growth is still unresolved but many theories are suggested by explaining the cloud structure using micro-physics of single droplets. Over sixty years of review, satellite data has presented to us an abundance of learning for understanding the atmosphere from space. Consequently, the approach of weather satellite technology increases the possibility of accurate study of clouds. In addition to the type division, clouds can further be classified in species, and sometimes divided further into varieties, which define the special characteristics of the clouds, and are related to the cloud transparency and the arrangements of the macroscopic cloud elements. Thus, with knowledge of cloud type, size, motion and development, one can calculate and predict the presence of cloud over underlying earth surfaces.

Literature reveals the huge number of cloud detection methods which are basically developed in two categories. The one is threshold based algorithm and another category depends on classification based approach. Threshold based approaches are based on spectral and textural information of clouds which is captured through the available channels of the sensors. But it is further observed that threshold method failed to give accuracy in the presence of

complex underlying surfaces to distinguish various types of clouds. On the other side, due to the breakthrough in the field of pattern recognition and machine learning algorithms, classification based approach started to apply on cloudy satellite images. These algorithms select the favorable features of clouds to optimize the model parameters through training data. However, these favorable features required to be extracted manually after large computation to gain better detection accuracy. Therefore, the use of neural network approach came into the picture to extract features automatically in order to distinguish cloud contaminated pixels in satellite images. Further, the researcher focused around the task of differentiating thin clouds because the information of thin clouds mixes with the information of underlying regions. Hence, the concept of super-pixel is used with the deep neural network approach [143]; but the approach ignores the multistage con-textures during classification. Thereafter, multi-scale deep neural network has been developed in order to extract both local as well as global features of clouds and segment the cloudy regions.

In features of development and limitation of cloud detection algorithm discussed above, present work discusses the various approaches on cloud detection with following major contributions.

- Algorithmic advancement of cloud detection is discussed based on spectral and textural parameters.
- Experimental simulation on existing neural network approaches with spectral and textural parameters is employed as classifier for cloud detection.
- The growth of neural network approach is further extended towards deep neural network for cloud detection in satellite images.

As a confine of research on cloud detection approaches, this paper aims to the growth of cloud detection algorithm by sinking the knowledge from the literature review and classifying the articles from 1970 to 2020. The reason for selecting this duration is that the information and communication technology of this era has played a vital role, not only in the direction of remote sensing but also in the potential to accumulate data from online databases. This literature review started on October 2016 and it was built on a search in the keyword index, abstract and conclusion of articles under IEEE transactions on Geo-science and Remote Sensing, Royal Meteorological Society, Monthly Weather Review, AGU Journals, Journal of Applied Meteorology, International Journal of Remote Sensing, Remote Sensing of Environment, Journal of Geophysical Research, Journal of Climate and Applied Meteorology, ISPRS Journal of Photogrammetry and Remote Sensing, Journal of the Atmospheric Sciences and Journal of Climate. It was observed that limited number of research articles was easy reach to the thematic point and hence additional sources are added in this work. Based on the scope of approximately 200 selected journal, books and thesis are collectively represented in a wider range of methodologies in algorithmic advancement and experimental analysis on cloud detection approach. Suitable papers are summarized and enable us to derive conclusions and recommendations for further investigations while evaluating and comparing different methods.

The remainder of present article is arranged as follows. Section 2 presents the algorithmic advancement on cloud detection in detail. Section 3 conducts a critical experimental analysis with Existing Techniques on Cloud Detection. Section 4 presents the simulation results on existing neural network approach to detect clouds using various supervised learning rules. Section 5 discusses the algorithm that demonstrates the growth of neural network approach towards deep neural network for cloud detection. Lastly, the concluding remarks are given in Section 5.

2 Algorithmic advancement on cloud detection

This section presents various studies that incorporate the algorithms which were not highly customized; instead the focus was on cloud detection algorithm and its related works by evaluating under a given parameters. Its result of literature is shown graphically using web knowledge in Fig. 2.

2.1 Spectral parameters

The properties of cloud in terms of its spectral parameters like irradiance and reflectance values demonstrate the thickness, height and amount of cloud [32, 114]. Bowker et al. [14] studied spectral reflectance information of 156 targets using which classification algorithms were designed to differentiate cloud against various earth's surfaces. Similarly, different studies has been observed where individual earth's surface is taken into account against cloud. For example, Greaves and Chang; Liljas; Crane and Anderson [29, 46, 100] assessed the reflectivity at $3.7\mu\text{m}$ to distinguish water and ice crystal cloud. Raschke et al. [132] used the spectral information to differentiate cloud against snow/ice in the polar region with AVHRR data. Similarly, many efforts have been made to extract spectral parameters such as,

- Ackerman et al. [1] assessed the analysis on cloud mask algorithm for MODIS data and found that $0.86\mu\text{m}$ is sensitive to detect all types of cloud in absence of sun-glint. The algorithm found difficulty in detecting cloud during night time in polar region.
- Lo and Johnson; Kuhn [84, 105] evaluated the absorption region of water vapor over Nimbus II satellite. The analysis found that the wavelength range of $6.4\text{--}6.9\mu\text{m}$ is used to recognize water vapor over clear atmosphere.
- The ratio of near-infrared (NIR) band to visible (VIS) band has been utilized to discriminate land area with threshold values 1.6 and sea surfaces with 0.75. This is due to the fact that VIS is more reflective than NIR band. However, this ratio technique was found ineffective in the existence of sun-glint [11, 100, 137].
- Raschke et al. [29] observed the similar reflectance property of cloud and snow/ice in VIS band and hence it is challenging to extract cloud contaminated region in this range. However, more reflectance of cloud is observed in the water absorption band ($1.5\mu\text{m}$)

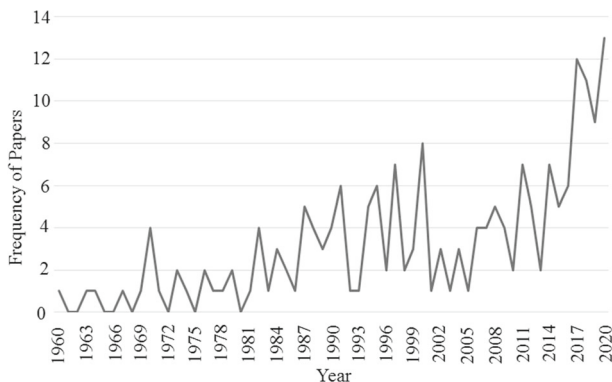


Fig. 2 Growth of popularity of methods in cloud classification over the past decades

whereas more absorption for snow pixel is seen in this band. Hence $1.5 \mu\text{m}$ can be used to differentiate cloud against snow.

- Grant and Hunt [45] studied that ice cloud can be discriminated in 10–11 μm band against water territory due to 10% difference in its refractive index.
- Crane and Anderson [29] used 1.6 μm channel of AVHRR satellite to discriminate cloudy region over snow/ice.
- Hulley and Hook [65, 99] introduced new ASTER cloud mask algorithm (NACMA) to differentiate cloud against snow/ice/deserts/or thin cloud from Landsat-7, MODIS, and AVHRR data. To check the validity of the algorithm, normalized difference snow index (NDSI) was used and found that its value greater than 0.4 shows the presence of snow with reflectance greater than 0.11.
- Jedlovec et al. [71] used bispectral composite threshold (BCT) algorithm over GOES-12 data using 11 μm and 3.9 μm channels. Better performance of the algorithm is observed for day time cloud detection during summer season because sun-glint produces better contrast between cloud and earth's surfaces in the infrared channel.
- Shenk et al. [145] investigated that 2.3 μm is sensitive channel for cloud detection and particularly cirrus cloud can be discriminated at 6.5–7.0 μm range of wavelength.
- Szejwach [153] observed that cirrus cloud is colder in 5.7–7.1 μm and 10.5–12.5 μm ranges.
- Yang et al. [174] developed automated ground based cirrus cloud detection scheme because of non-uniformity of cloud brightness in the presence of sun-glint. Hence background subtraction method is used named background subtraction adaptive threshold (BSAT) method to detect cirrus cloud and eliminate illumination effect.
- Band ratio of Red channel to Blue channel is used which gives better results for thick cirrus cloud but fail in case of thin cirrus. Hence combination of BSAT and adaptive threshold method is incorporated with band ratio method to extract cirrus cloud accurately.
- Bell and Wong; Liljas [11, 100] found that thin cirrus cloud is detected by taking the brightness temperature (BT) difference at 11 μm and 12 μm .
- The BT difference between 11 μm and 3.7 μm is used to detect low stratus cloud during night time [11, 100, 137].
- Li et al. [96] used hybrid thresholding algorithm (HYTA) using band ratio technique of blue to red band with minimum cross entropy thresholding. The challenge to the technique is to discriminate cloud in the presence of sun-glint which required additional thresholding techniques.
- On the same application, day time [55] and night time [56] cloud detection algorithm was developed over FengYun (FY-3A)/VIRR polar-orbiting meteorological satellite data using BT difference of 1.38 and 1.6 μm and 3.7 and 12 μm to differentiate high and low cloud, respectively.
- Gupta and Panchal [52] realized day time cloud detection algorithm using decision tree approach in the VIS band. The results were compared with FY-3A/VIRR official cloud mask and achieved more accurate detection rate.

Surface albedo is the ratio of reflected radiation to its incident radiation on planet. The amount of reflected radiation not only demonstrates the characteristics of surface but also shows the spectral and angular distribution of irradiation [26]. Albedo parameters affect the climate by knowing the absorption radiation of planet. The variation in albedo due to irregular heating effect among land and ocean territories drives the effect of climatic information. Welch et al. [165] investigated that albedo is the cosine of solar zenith angle measured at

each pixel and scaled the gray level from 0-255 corresponds to 0-100%, respectively. The pixels appear brighter with more value of albedo and darker with less value of the albedo parameter [151]. Few contribution on cloud detection based on this parameter are,

- Climate is hot and humid in tropical region and has heavy rainfall. Hence, Kazantzidis et al. [78] estimated the effect of cirrus cloud with different solar zenith angles over solar irradiance to predict tropical cyclogenesis. This increases the accuracy in the tropical measurements as well as the impact of cirrus cloud on climate.
- The automatic cloud cover assessment (ACCA) has been proposed [54] for Landsat data but false results for cloud edges and transparent clouds are observed due to the absence of channel around $1.375 \mu\text{m}$. Thus ACCA algorithm was improved by changing its threshold parameter to minimize the error in cloud mask and improve the detection of optically thin cirrus cloud and its edges [159].
- Zhong et al. [189] developed object-oriented cloud and cloud-shadow matching (OCM) algorithm for charge-coupled device (CCD) sensors where the availability of channels are not enough. The OCM method uses the concept of modified ACCA approach for Landsat 7 image data to initialize the cloud map with omission error of 14.74%.

2.2 Textural parameters

Spectral features were used to evaluate the average brightness variation between the bands while textural features include identification of different parameters (such as rough, chunk, striped, ruffle, thin and many more contextual attributes) within a band in spatial domain.

Pickett [131] studied that color, appearance, orientation and size of a region can be utilized to distinguish various textures within the region. Therefore, Gutman et al. [53] proposed a focus on extension of spatial domain information into radiance information of VIS to extract cloudy region. Moreover, Knottenberg and Raschke [79] observed many applications like cloud and its shadow detection using visible imagery by distinguishing textural properties of clouds. However, Seze and Rossow [142] observed some situations like dust or cirrus cloud over the desert where least information of spatial variance is observed in the VIS channel. Similarly, various analysis is done for cloud detection using textural parameters such as,

- Cheng et al. [24] developed cloud removal algorithm based on similar pixel replacement from multi-temporal images using markov random field (MRF) [20, 34]. Better results were seen over SPOT/HRV images.
- Shenk et al. [145] studied 13 different variety of features such as height, appearance, albedo and fourier power spectrum (FPS) to obtain climatic conditions by discriminating various types of cloud till five years over Nimbus-3 medium resolution infrared radiometer (MRIR) low-spatial resolution data.
- The classification of stratocumulus, cumulus and cirrus clouds using textural features has been done over Landsat multi-spectral data [126]. However, stratocumulus misclassified with cumulus cloud because of their least difference in size, position, number of gaps and presence of different gaseous element across the radiation path.
- Chen and Pavlidis [138] investigated that magnitude spectrum is huge for smooth texture analysis with high spatial frequency and low value for rough texture analysis. Therefore, Kuenning et al. Rossow [83, 133] proposed a focus on extraction of rough and smooth textural features using entropy of particular range of wavelengths at low-level of altitude [43].

- Seze and Rossow [142] observed difficulty in distinguishing various types of cloud in the small inhomogeneous surface with low contrast case. This is due to the fact that inhomogeneous surface consists large spatial variability and hence mis-classifies as cloudy pixel. Therefore, various cloud detection techniques (e.g., [115, 134, 142, 150, 162]) have been discovered using threshold technique for low contrast region by computing the various statistics of cloud.
- [163] employed GLDV method for texture analysis in order to generate multi-spectral signature of various characteristics of cloud. Likewise, Welch et al. [164] used GLDV to extract 9 texture features in unsupervised classification. The analysis demonstrates decaying in classification accuracy while improvement is seen in run time storage capacity by 40% and 87%, respectively.
- Welch et al. [164] proposed a focus on gray level co-occurrence matrix (GLCM) [58] as texture analysis to classify different types of cloud to increase the accuracy of analysis over satellite images. It was investigated that GLCM saves 50% storage capacity and 30% in run time. Further, sum and difference histogram (SADH) and gray-level difference vector (GLDV) textural features which are developed from GLCM approach for classification purpose. Over comparison, it is observed that the accuracy using SADH and GLDV for cirrus cloud classification is not improved but the improvement in accuracy is found for the discrimination of stratocumulus and cumulus cloud.

2.3 Other parameters

The studies based on few other parameters are,

- Emissivity in NIR band between 0.1 to 0.95 is utilized to measure variation in cloud thickness [41, 66, 68, 101, 170]. Hunt [66] studied the emissivity of cloudy region and observed its sensitive in 8.5–13 μm range as compare to the other channels of SWIR band such as 2.3, 3.5, and 3.8 μm . For night time analysis, the difference of 12 μm and 3.7 μm ranges is used for mid-level cloud detection [11, 124, 137].
- Water cloud consists comparatively less water droplets than ice cloud. This information helps to distinguish water cloud against ice cloud using cloud scattering parameters ($\cos\theta$).
- Contrast parameter is calculated from blocked image to get the multiple optimal threshold values and remove the thin cloud [190].
- Shen et al. [144] observed that thin cloud consist of low frequency information and hence developed the frequency domain based algorithm using homomorphic filter. This algorithm is applied over cloudy region of Landsat ETM+ and GaoFen-1 images in order to reduce the false detection for clear pixel by removing thin clouds and recover the ground information.

3 Experimental analysis with existing techniques on cloud detection

This section presents the work where the focus is on the evaluation of compared applicable results between different techniques.

Table 1 demonstrates the experimental analysis of various cloud classification algorithms. The content in the column of the table shows: (i) name of the satellite/ imager used for the analysis, (ii) classification method (iii) publication year of the paper, (iv) average accuracy of the classification method of all classes, (v) maximum accuracy obtained, (vi)

Table 1 Experimental results based on maximum used classifier on cloud classification

Instrument	Classifier	Publish	Overall Accuracy	Best Accuracy	Least Accuracy	Classes	References
Landsat MSS	FFBP	1990	93%	96%, cloudy boundary layer	88%, stratocumulus	3	[91]
AVHRR LAC	FFBP	1992	87.6%	≥90%, cirrus/land/snow	75%, cloud/fog over ice	10	[165]
AVHRR	PNN	1994	79.8%	91.2%, low cloud/altostratus	–	10	[8]
GOES-west	threshold & NN	1994	94%	–	93%	2	[130]
Landsat TM	Kohonen SOFM	1994	85.53%	Farm	bare land	9	[177]
AVHRR	MLP NN	1995	>95%	97%, TLP NN	–	2	[175]
AVHRR	PNN	1996	86.8%	98.6%, clear	83.1%, midlevel cloud	10	[9]
Meteosat	ANN	1997	–	–	–	3	[93]
AVHRR	CANN	1997	82%	–	–	9	[112]
AVHRR	ANN	2000	91%	88%, thin cirrus over low cloud	76%, cirrus over water	13	[156]
GOES 8	PNN	2000	75.4%	95.2%, water	20.7%, altocumulus	10	[154]
GOES 8	SVM	2000	78.5%	90%, land/water	31%, stratus/cumulus	10	[5]
Landsat TM	LSMM	2000	–	–	–	3	[15]
NOAA-14/AVHRR	thresholding	2002	89%	–	–	5	[22]
AVHRR	MNN	2002	–	–	–	2	[103]
NOAA-16/AVHRR	MVC	2003	93%	–	–	3	[21]
RADAR	SOFM	2004	–	–	–	7	[59]
MODIS	Multicategory SVM	2004	82%	–	–	3	[90]
SPOT	MLP NN	2006	97.6%	97.2%	>50%, thick cloud	3	[70]
Enhanced ETM+	SVM	2006	–	–	–	4	[111]
MISR	SVM	2006	81%	80.1%, cloud	33%, aerosols	5	[110]
GOES	FCNN	2007	~90%	90.1%, day sea	77.9%, night sea	25	[10]
MODIS	spectral	2008	85%	–	–	3	[1]

Table 1 (continued)

Instrument	Classifier	Publish	Overall Accuracy	Best Accuracy	Least Accuracy	Classes	References
GOES	thresholding	2008	87.6%	90.9%, summer	81.8%, low cloud in winter	3	[71]
SPOT, PLEIADES_HR	SVM-BEE	2009	94.08%	98%, concrete	72%, roof	9	[39]
Kalpana-1	SVM	2009	78.05%	92.3%, LSM with TF	65%, using GRBF	2	[25]
–	Hybrid Threshold	2011	88.53%	stratiform	84.01%, cirriform	3	[96]
Landsat ETM+	ACCA	2011	90.12%	94.94%	82.12%	2	[160]
SPOT	region growing	2011	79.6%	95%, cloud size; 50 ha	54%, cloud size; 10ha	6	[140]
Landsat	Fmask	2012	96.4%	92.1%, cloud	70%, cloud shadow	2	[192]
LiDAR	SVM	2013	>92.34%	99.89%, ground	83.80%, buildings	6	[187]
Landsat 7 ETM+	NN	2014	98.8%	–	–	2	[76]
SPOT	threshold	2014	95%	–	85%, shadow	4	[40]
–	SVM	2015	90%	99.13%, ground	77.29%, thick cloud	4	[95]
Landsats 4?, 8, Sentinel 2	Fmask	2015	89%	100%, clear water	66.66%, cloud shadow	4	[191]
Gao Fen -2 (GF-2)	SVM	2016	91.45%	95.67%	81.76%, TF	4	[7]
–	Bayesian classifiers	2016	≈90%	97% clear sky	77% low cloud	4	[2]
Chinese Huan Jing 1 (HJ-1A/B), Gao Fen 1 (GF-1)	OCM	2017	89.6%	93.72% clear sky	85.26%, cloudy	2	[189]

minimum accuracy achieved and (viii) number of classes. All the information in the table are grouped based on classifier and its accuracy. Here accuracy is subjective to human factor because results are based on visual comparison. The parameter which affects the accuracy of classifier is the number of output classes varying from 2 to 10 distinguish regions. Also, the accessible channels and cost of equipments which tend to the use of final measurements can also affect the accuracy of classifier. It is observed that different strategies have been developed conceptually based on neural network (NN) method for cloud classification application. The frequency growth of NN method is demonstrated in Fig. 3 for cloud classification application. Therefore, next section incorporates detailed study on NN approach to improve the accuracy of cloud classification for multi-spectral satellite images.

4 Neural network approach

Neural network approach was utilized to improve classification accuracy in geophysical data [91]. It can put the solutions of complicated problems by training the network on the feature value. Therefore, NN is described as a possible alternative way to discriminate cloud using simple components which function in a parallel manner [128].

4.1 Basic concept

A NN consists neurons which are represented as nodes. Nodes are defined by real values which demonstrate the activity of neurons and are directly connected with the weighted paths. Therefore, neural network layer is defined as weighted sum of the input nodes which is passed through the activation function. The architecture of the neural network is demonstrated here in Fig. 4.

Generally, NN consists three or more layers i.e. an input layer, an output layer and one or more hidden layers between input and output layers. The nodes of the network consist adder and non-linear function. The input (x_l) and its output (a_l) are collectively written with l^{th} layers in (1)

$$[n_{1l} \ n_{2l} \ n_{3l} \ \dots \ n_{sl}]^T = \begin{bmatrix} w_{11l} & w_{12l} & w_{13l} & \dots & w_{1Rl} \\ w_{21l} & w_{22l} & w_{23l} & \dots & w_{2Rl} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{s1l} & w_{s2l} & w_{s3l} & \dots & w_{sRl} \end{bmatrix} \begin{bmatrix} x_{1l} \\ x_{2l} \\ x_{3l} \\ \vdots \\ x_{Rl} \end{bmatrix} + \begin{bmatrix} b_{1l} \\ b_{2l} \\ b_{3l} \\ \vdots \\ b_{sl} \end{bmatrix} \quad (1)$$

$$a_l = f_l(n_l) \quad \text{with} \quad n_l = W_l x_{l-1} + b_l$$

where, f_l demonstrates the activation function of l^{th} layer in NN to introduce the non-linearity and calculate nontrivial problems.

4.2 Advancement in cloud classification algorithms

Cloud classification application in remote sensing has recently found numerous work on NN. For the review, we found approximately 55 relevant articles which are published in last few 20 years (Fig. 2). This section summarizes NN advancements where papers are merely compared NN performance with other algorithms or discusses the specific applications.

Probabilistic neural network (PNN) is a supervised classifier that classifies cloudy region by estimating probability function of the input information. Bankert [8] proposed a focused

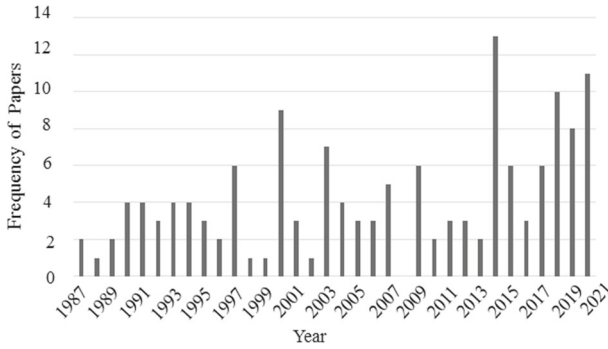


Fig. 3 Frequency growth of NN method in cloud classification

on two layer PNN with spectral, textural, and physical input features and hold-one-out training and testing method to differentiate the classes: cirrus, cirrocumulus, cirrostratus, altostratus, nimbostratus, stratocumulus, stratus, cumulus, cumulonimbus and clear [149]. Lee et al. [91] proposed a focus on feed-forward network trained by back propagation (BP) NN classifier with cloud textural input features. This algorithm worked for the discrimination of cirrus, stratocumulus and cumulus cloud as these clouds are important radiative cloud types. Miller et al. [113] proposed a different classification scheme which worked as an alternative to BP called CMAC algorithm. CMAC algorithm can improve interaction between cloud and climate using prior knowledge of contextual information. Xie et al. [168] proposed a focus on multilevel cloud detection algorithm using deep learning method. CNN approach was used to differentiate thick cloud, thin cloud, and clear pixel using simple linear iterative clustering (SLIC) method. This approach is further used in [104] where

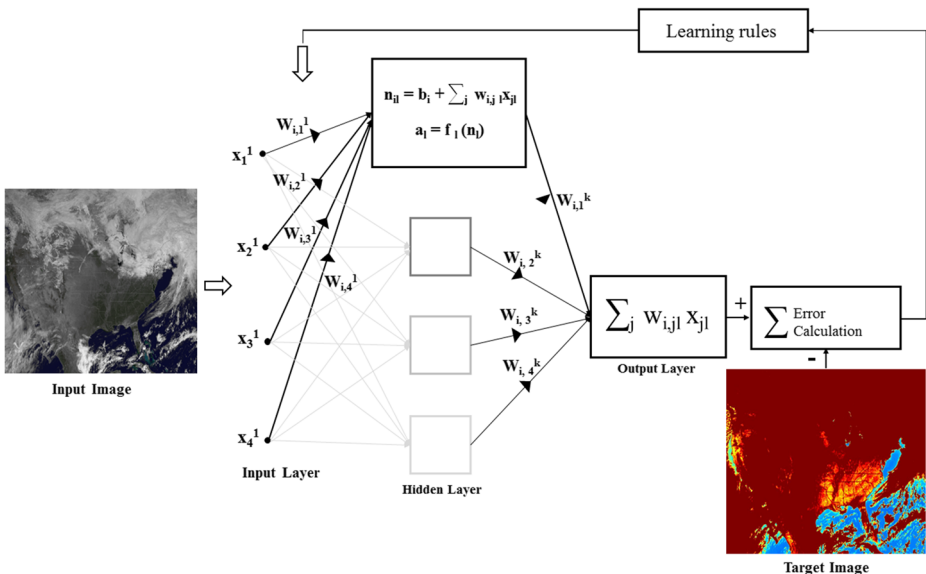


Fig. 4 Architecture of artificial neural networks to extract cloud contaminated pixels

combination of SLIC and SEEDS has been utilized to distinguish the entire image into super-pixels and classifies into thick cloud, thin cloud, building and other culture. Lee et al. [91] presented similar approach on three layer NN to detect cirrus cloud and obtained 85% accuracy. However, it has been observed that cirrus cloud was misclassified with cumulus cloud. Therefore, four layer NN with nine texture features in the input layer was proposed where cirrus cloud was detected with accuracy of 96%, stratocumulus cloud with 92% and cumulus cloud with 90%. In a different approach, Walder and MacLaren [156] studied ANN to design automatic cloud classification system by using textural and spectral features extracted from AVHRR data.

Lewis et al. [92] integrated hopfield NN method to track the cloud in the visible and infrared channels of Meteosat 5. Shape and spectral parameters have been used in this method. Shape parameter includes area, perimeter, area to perimeter ratio, length, width, and length to width ratio of cloudy region while visible spectral parameters include average and maximum intensities of cloud contaminated pixels in the satellite images.

Self-organizing map (SOM) is the only unsupervised NN method that reduces the dimensions of data by mapping it into two dimensional data space. Murao et al. [118] evaluated a classification scheme emphasizing both SOM and three-layered feed forward NN trained by back propagation method. SOM is used to extract texture features. This method is named as hybrid NN which was applied over GMS infrared temporal data to estimate rainfall rate. Peak and Tag [129] obtained the training data based on geographical information and processed for Kohonen's self-organizing feature map. Using Kohonen SOFM method, nine patterns were differentiated correspond to dark part of forest, bare land, road, grassy place, water region, farm, cloud, shadow of cloud and density inhabited district. Neural network classifier was then compared with bayesian method and it was found that average accuracy is almost same using both the methods but in case of overall accuracy, NN gave better results using nine neurons in the hidden layer.

Feijt et al. [38] designed neural network, named as mapping neural network (MNN) with two hidden layers to achieve cloud parameters: mean optical depth, effective radius, relative cloud inhomogeneity, and fractional cloud cover of multispectral reflectance data (e.g. MODIS) using NIR channels. Further, same authors have used MNN [36, 37] to measure radiant flux (i.e. reflectance, transmittance, and absorption) of inhomogeneous cloud using Monte Carlo method. Evaluation of method was done by using horizontal changes of optical depth and not of radius of cloud particles with the inclusion of two solar zenith angle (i.e. 0° and 60°). It was also observed that results gave better outcome for the reflectance than for the transmittance and absorption. The shadowing effect of neighboring pixels is found in the transmittance of 60° while transmittance of 0° reflects the more photon penetration into the optically thin cloudy pixels.

Jang et al. [70] observed difficulty in cloud detection using SPOT database because of the absence of thermal bands. Hence ANN method was used with single sigmoid layer by varying number of neurons and trained by Levenberg-Marquardt back-propagation (LM-BP) method. Good performance was obtained in thick and thin clouds. Bankert and Wade [10] improved neural network as cloud classifier by reducing the training databases using fast condensed nearest neighbor (FCNN) rule discussed by Angiulli [3].

4.3 Classification learning rules for cloud detection

At the heart of weight update method, an expression for partial derivative of the cost function with respect to the weight of the network is present. The expression demonstrates the change in cost function value correspond to change in weight and bias values. This method

is termed as back propagation (BP) method. Hence BP is not only fast learning method for the network but also detailed insight information of change in behavior of weights and bias of the network are also obtained. Therefore the discussion in this section is only an onset of thorough understanding of the various equations used by the researchers that are carried out in Table 2.

4.4 Performance comparison of learning rules in NN model

Neural network has attracted many researchers over the last couple of decades to tackle cloud-related problems. So this section provides a detailed discussion on binary cloud classification using NN approach with different texture features (or parameters). The core idea of the method is to compute GLCM matrix of each pixel of multispectral databases. After that texture feature is calculated from computed GLCM matrix and then input to the NN system.

NN approach includes single Layer Perceptron (SLP), two Layer Perceptron (TLP) and multi Layer Perceptron (MLP) schemes. MLP is computationally complex and is compared for binary cloud classification. MLP is expected to be used for deep learning applications and used to solve complex classification problems. But in case of binary classification, they need long training time and produced correct classification accuracy similar to TLP network. Hence, performance and execution time of SLP and TLP with different neurons are compared and not compared with MLP in order to investigate the effectiveness of NN learning rules. Number of neurons in hidden layers are selected based on the past work which is provided in Table 3.

Various feature selection methods were discussed in order to select subset of features from the textural and spectral feature bank. While evaluating methodology, principal component analysis (PCA) is used in this article to reduce the dimensionality of the input and to identify the most important features. These selected features aim to get more accurate day-time cloud detection as well as capable to adapt different satellites. The following are the important textural features which has been identified by PCA scheme to design input of the network:

$$Correlation = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (2)$$

$$Contrast = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - j)^2 p(i, j) \quad (3)$$

Equation (2) demonstrates the mathematical formulation of correlation that shows the joint probability occurrence of the function, $p(\dots)$ which returns the value of specified pair of GLCM matrix. The mean and standard deviation are denoted by μ_i and σ_i , respectively of i^{th} row of matrix. Similarly, μ_j and σ_j are the mean and standard deviation of its corresponding j^{th} column. Contrast in (3) shows the local variations in luminance that distinguish the object. In order to validate the performance of various learning rules (provided in Table 2), investigation is carried out on Landsat8 OLI/TIRS database. Table 4 enlisted with real-life multi-spectral databases with their detailed information of training and testing samples. All images of these databases are converted into 256 gray-level intensity value. The simulation of the classification learning algorithm is carried out on MATLAB version

Table 2 Mathematical model of learning rules in ANN model

Learning rule	Characteristics	Details	References
Convolutional NN (CNN)	Explanation Representation	Back propagation method is used to determine weight iteratively to minimize mean square error $\vec{\omega}_{n+1} = \vec{\omega}_n - \eta \vec{\nabla} E$ $\vec{\nabla} E = \text{gradient of error function,}$ $\eta = \text{learning rate}$	[135, 165]
Momentum decent algorithm	Explanation Representation	Back propagation method is used to determine weight iteratively to minimize mean square error $\vec{\omega}(n+1) = \vec{\omega}(n) + \eta \delta_{\omega_{ij}}(n) x_j + \alpha \delta_{\omega_{ij}}(n-1)$ $\delta_{\omega_{ij}} = \begin{cases} (1 - y_i) y_i (y_i - t_i), & \text{output layer} \\ (1 - y_i) y_i \sum \omega_{ki} \delta_{\omega_{ki}}, & \text{hidden layer} \end{cases}$ $n = \text{iteration step, } \alpha = \text{Momentum constant}$ $\delta_{\omega_{ij}} = \text{back propagation error}$	[77, 175]
Noarl-Neural	Explanation Representation	Learning trial is used to stabilize weight adjustment. $\Delta \omega_{ij}(n+1) = \eta (\delta_{p_i} a_{p_j}) + \alpha \delta \omega_{ij}(n)$ <p>where, n = iteration, δ_{p_j} = error, a_{p_j} = network output</p>	[127, 128]
Kohonen self-organizing feature map	Explanation Representation	weight vector adjustment to minimize Euclidean distance from neuron to input vector. $\omega_j(t+1) = \begin{cases} \omega_j(t) + \eta [x(t) - \omega_j(t)], & j \in N_c; \\ \omega_j(t), & \text{otherwise} \end{cases}$ <p>where, x = input vectors, N_c = neighborhood nodes</p>	[6, 80, 102, 128, 178]
Hopfield NN	Explanation Representation	To minimize energy of a network $E = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \omega_{ij} x_i x_j$ <p>where, x_i, x_j = binary state of two neurons.</p> $\omega_{ij} = \omega_{ji}; \omega_{ii} = 0$ <p>The weight matrix ω is symmetrical and null on self feedback connections</p>	[28, 60, 92, 93, 166]
Self-organizing feature map (SOFM)	Explanation	The training data is distributed into N clusters with ω_j center of j class data. The optimal centers of clusters are defined with minimum average reconstruction error	[59, 61–63, 118, 156]

Table 2 (continued)

Learning rule	Characteristics	Details	References
	Representation	Average Error is defined as, $E(\omega) = \sum_{j=1}^K \int_{C_j} \ x - \omega_j(x)\ P(x) dx$ Where, $P(x)$ = Probability density function of input data x , $\omega_j(x)$ = cluster center of C_j , $j = 1 \dots K$, Weight: $\omega_v(t + 1) = \omega_v(t) + d(u, v, t) \cdot \eta(t) (x - \omega_v(t))$ where, t = iteration, x = input sample, $d(u, v, t)$ = distance between u and v neurons, $\eta(t)$ = monotonically decreasing coefficient	

R2015a on Intel(R) core (TM) i5processor 2.30 GHz CPU along with 4GB of inbuilt RAM in Windows-7(64-bit).

To check the validity of different learning methods over different databases, demonstrated in Fig. 5 and discussed in Table 4, binary cloud classification is considered for comparison due to the differences in output classes, hypothesis and considered parameters. Table 5 demonstrates the summary of effect of different training samples over classification accuracy, performance and execution time of classifier. An empirical comparison of the Landsat8 OLI/TIRS data with different percentage of training samples with respect to different neurons is presented in Table 3. We have intentionally selected CNN learning rule with Levenberg-Marquardt weight updating technique because it is a most widely used technique to realize for practical application.

To validate the performance of NN as a classifier or predictor, it is found that use of cross-entropy is better than other classification error (e.g., mean square error (MSE)). Cross-entropy calculates a network performance using each pair of target (T) and output (y) elements. Minimization of cross-entropy leads to be a good classifier. Equation 4 demonstrates the calculation of cross-entropy for each t - y pair. $N=1$ is a special case where classifier behaves as a binary classifier having 0 or 1 value with the target element. The results show that NN approach with 10% training condition in case-e has achieved approximately 69% accuracy with less execution time which is significantly sufficient than all

Table 3 Selection of number of hidden layer neurons in NN architecture

No. of Neurons			No. of Neurons		
First Layer	Second Layer	References	First Layer	Second Layer	References
(SLP)	(TLP)		(SLP)	(TLP)	
4		[175]	6	6	[127]
11	–	[70]	10	10	[35]
22	–	[69]	40	20	[91]
41	–	[148]	55	55	[165]

Table 4 Multi-spectral data analysis with training and testing approaches in NN model

S. No.	Databases	Bands used	Samples	Train	Test
i	Landsat8 OLI/TIRS [85]	4	29480	2948	26532
ii	AVHRR [85]	4	201285	20128	181157
iii	GOES [42]	4	589824	58982	530842
iv	NOAA [120]	4	32041	3204	28837

other cases. Meanwhile, features extracted have positive influence on the performance (here cross-entropy) in case of TLP approach.

$$\begin{aligned} \text{Cross entropy} &= -T \log(y); K = 1 \\ &= -T \log(y) - (1 - T) \log(1 - y); K > 1 \end{aligned} \quad (4)$$

Two sets of comparison experiments with concluded TLP scheme are listed: one using the supervised NN rules and other using the unsupervised NN rule. Table 6 demonstrates the results of comparing the different supervised learning rules (discussed in Table 2) with TLP NN having (6,6) neurons in hidden layers and 10% training samples. It is found that the

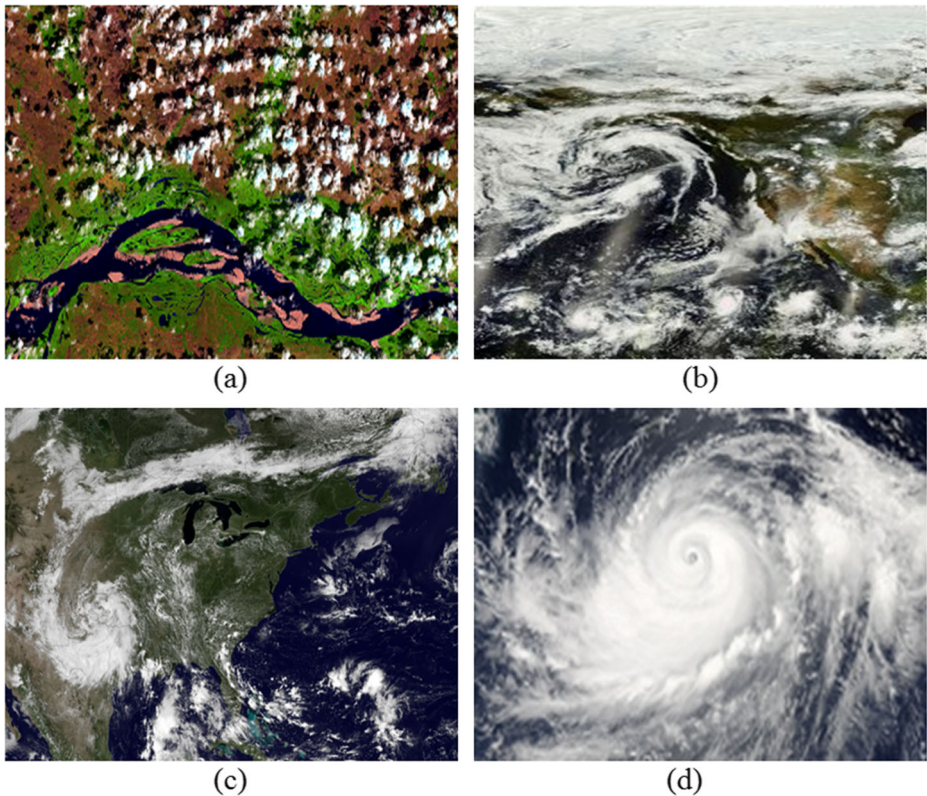


Fig. 5 Multi-spectral databases used to analyze the cloud detection performance with existing methods.: (a) Landsat8 OLI/TIRS (b) AVHRR (c) GOES (d) NOAA

Table 5 Arranged Levenberg-Marquardt training sets with inserted neurons in NN model

Case	FL	SL	No. of Neurons	Classification	Execution time	Performance	Epochs	Classification	Execution time	Performance	Epochs
				10% training samples				20% training samples			
a	4			67.22	1.1183	0.0066	50	67.22	1.117	0.0092	39
b	11			65.9	1.1816	0.0123	35	68.08	1.1282	0.0097	30
c	22			69.21	1.1171	0.0117	88	69.38	1.1146	0.0081	37
d	41			69.38	1.1538	0.0055	44	68.08	1.1642	0.0078	78
e	6	6		69.38	1.1369	0.0053	52	66.7	1.1471	0.0107	47
f	10	10		69.37	1.153	0.0097	46	69.38	1.15	0.0109	97
g	40	20		69.38	1.2055	0.0117	36	69.38	1.2067	0.0086	46
h	55	55		69.38	1.2826	0.0162	35	69.38	1.2883	0.0092	27
				30% training samples				40% training samples			
a	4			69.38	1.1503	0.0091	26	69.38	1.1406	0.008	39
b	11			67.22	1.1373	0.0096	20	69.38	1.1488	0.0085	41
c	22			69.38	1.1352	0.0087	30	66.56	1.1565	0.0086	37
d	41			69.38	1.1746	0.0066	55	65.81	1.1952	0.0097	61
e	6	6		69.38	1.1483	0.0066	50	69.39	1.1445	0.0091	70
f	10	10		69.38	1.15	0.0064	32	69.38	1.1756	0.0089	44
g	40	20		69.38	1.2232	0.0079	84	67.22	1.2222	0.0109	31
h	55	55		65.9	1.3149	0.0086	44	69.38	1.2782	0.0081	29

Table 5 (continued)

FL	SL	No. of Neurons	Classification	Execution time	Performance	Epochs	Classification	Execution time	Performance	Epochs
Case			50% training samples				60% training samples			
a	4	6	66.69	1.1521	0.0149	39	66.69	1.1137	0.0133	43
b	11	10	67.22	1.1309	0.0157	42	66.69	1.1207	0.016	43
c	22	20	67.22	1.1363	0.0159	58	67.051	1.1401	0.0143	42
d	41	55	67.22	1.17	0.0159	80	66.7	1.1758	0.0145	41
e	6	6	66.69	1.1404	0.0154	28	66.7	1.1406	0.0156	47
f	10	10	67.22	1.162	0.0172	49	67.22	1.1831	0.013	69
g	40	20	67.81	1.2468	0.0145	75	68.08	1.2171	0.0139	34
h	55	55	67.22	1.3031	0.0142	25	68.08	1.2843	0.0157	29
Case			70% training samples				80% training samples			
a	4	6	66.69	1.1162	0.0272	33	66.69	1.1119	0.031	21
b	11	10	67.22	1.1068	0.028	37	66.7	1.1281	0.0284	29
c	22	20	66.7	1.1381	0.0275	42	67.05	1.1575	0.0319	49
d	41	55	68.08	1.167	0.0259	104	66.29	1.1926	0.0342	27
e	6	6	69.38	1.1877	0.029	33	68.08	1.1721	0.0311	76
f	10	10	68.08	1.1613	0.028	60	67.22	1.1739	0.0299	51
g	40	20	66.7	1.2593	0.0311	17	67.81	1.2676	0.0287	40
h	55	55	67.22	1.2737	0.0261	37	69.37	1.3858	0.0309	47

Table 6 Comparison of supervised learning rules in NN model

Dataset	Learning rule	Accuracy	Error Rate	TPR	FPR	FNR	TNR	Prevalence
L-8	Gradient Decent	40.7	59.3	99.8	63.4	36.6	0.2	6.5
	Gradient Decent Momentum	42.9	57.1	99.9	62.5	37.5	0.1	8.7
	Levenberg-Marquardt	68.1	31.9	100	48.2	51.8	0	33.8
	Hopfield network	65.8	34.2	0	34.2	65.8	0	65.8
AVHRR	Gradient Decent	63	37	64	40.1	59.9	36	76.4
	Gradient Decent Momentum	64	36	64	35.9	64.1	36	80
	Levenberg-Marquardt	69.2	30.8	65.5	1.5	98.5	34.5	88.8
	Hopfield network	58.4	41.6	0	41.6	58.4	0	58.4
NOAA	Gradient Decent	55.2	44.8	80.1	62.9	37.1	19.9	42.2
	Gradient Decent Momentum	63.1	36.9	100	55.3	44.7	0	33.3
	Levenberg-Marquardt	54.7	45.3	79.9	63.3	36.7	20.1	41.7
	Hopfield network	70.2	27.8	0	29.8	70.2	0	70.2
GOES	Gradient Decent	88.1	11.9	88.9	14.3	85.7	11.1	74.8
	Gradient Decent Momentum	53.8	46.2	70.2	70.1	29.9	29.8	59.1
	Levenberg-Marquardt	80.3	19.7	78.1	0	100	21.9	89.9
	Hopfield network	70.2	29.8	0	29.8	70.2	0	70.2

calculated cloud detection accuracy tolerates minor degradation while the minimum error rate is noticed with the use of Levenberg-Marquardt learning rule.

In addition, it is observed that combination of TLP NN with Levenberg-Marquardt learning rule gives maximum value in true positive rate (TPR) and false negative rate (FNR) while minimum value in true negative rate (TNR) and false positive rate (FPR). Hence TLP with Levenberg-Marquardt learning rule is selected for comparison with unsupervised SOFM learning rule (refer Table 2) for all the four databases to check the comparative performance of each algorithm for binary cloud classification. The results in Table 7 are found in favor of SOFM in terms of accuracy while it favors TLP in terms of ellapse time which shows the time complexity of training and testing computation of the algorithm.

In a different approach, double-branch PCA Network (PCANet) architecture is designed in combination with support vector machine (linear function of NN) classifier where bright and thick clouds are separated by threshold method and then this information is combined with SVM classifier to extract high-level of cloud information from multi-spectral satellite images [23, 193]. Further, the neural network approach has been utilized to identify classification membership of cloud, cloud shadow, water, snow/ice and clear pixels in a Landsat image [64]. With the similar approach, cloud detection algorithm in optical satellite imagery enhanced towards deep learning model. Few works have been studied like model based on U-net architecture which discriminate cloud against snow region in the visible range of electromagnetic spectrum with minimum multi-spectral training samples [73, 136].

A deep learning model for detection of clouds in optical satellite imagery has been observed, named as remote sensing network (RS-Net) which is based on the concept of U-net architecture. The model is trained and validated on Landsat 8 and SPARCS images. The

Table 7 Comparison of supervised and unsupervised learning rules in NN model

Classifier	Unsupervised Classifier			Classifier			Supervised Classifier		
	Dataset	Accuracy	MSE	Elapse Time	Classifier	Dataset	Accuracy	MSE	Elapse Time
SOFM	L-8	93.47	0.049	7.81	TLP	L-8	69.37	0.0036	5.18
	AVHRR	93.48	0.049	8		AVHRR	66.48	0.006	38.23
	GOES	67.1	0.048	428.03		GOES	80.29	0.0057	5.43
	NOAA	92.81	0.0494	11.68		NOAA	70.18	0.0042	9.75
	Mean Accuracy=	86.72 %		SD = 13.08		Mean Accuracy=	71.58 %		SD = 6.02

evaluated results indicate the better discrimination performance on hardly distinguishable surfaces such as cloud over snowy and icy regions. In particular, the performance of deep learning model with visible bands demonstrates promising outcomes of cloud detection with multi-spectral satellite images [73, 141, 152]. Xie et al. Further, [169] designed a two layer deep convolution neural network (CNN) to discriminate clouds into thick, thin, and non-cloudy regions, but the network finds difficulty in distinguishing clouds in the presence of snow region.

Further, Cerdena et al. [17] used ANN algorithm with MLP network having back-propagation to get various cloud parameters like temperature of cloud, its optical depth and its effective droplet radius. However, ANN has some important disadvantages like network dependency on their architecture, characteristic parameters and training algorithms due to which the solution stuck to the local solution of the error surface. Therefore, genetic algorithm (GA) has been used to overcome this problem and get the global optimal solution. This indicates cloud detection implementations using artificial evolutionary algorithms (EA) which is a promising machine learning algorithm. EA is particularly appealing in cloud detection application because of their ability to reach the optimal solution in least computation time and limited training samples. Therefore, Kaminsky et al. [77] suggested an algorithm named optimized cloud detection index (OCDI) using GA. The analysis used PCA method initially over MODIS database to get the optimal channel and then GA is applied to optimize the OCDI parameters. Thus the probability of getting the global solution over error surface increases by overcoming the limitation of ANN approach. Earlier, Lisens et al. [103] described an algorithm using GA to optimize the NN parameters in order to design cloud mask. Therefore, it has been observed that amalgamation of NN and GA decrease the computation time and increases the NN ability. The cloud detection study has been further enhanced towards multi-dimensional objective space. Recently, Gupta et al. developed multi-objective social spider optimization technique to obtain label data for single band using multi-spectral Landsat 8 satellite images. This label data is further utilized to train the NN to get label data for rest bands comes under visible range. By combining the decision of bands of visible range, cloud contaminated region is extracted against various surface of earth [51].

5 Deep neural networks for cloud detection

The rapid growth of machine learning technology to solve complex problems acquires attention towards deep learning methods in the past few years. The conventional machine learning approaches require a domain expert to extract appropriate features for a targeted problem. Deep learning approaches overcome this problem by extracting significant features [122, 123] from the raw data [88]. It acquires a body of processing layers where it can learn distinguish features of data by abstracting at multiple levels [87]. Therefore, deep learning algorithm has emerged in various application such as face recognition [86], speech recognition [57], image segmentation [147], signal processing [181], bio-activity prediction [157] and so on.

It has been observed that deep learning algorithm is modeled with different types of architectures like convolutional neural networks (CNN), deep belief networks and recurrent neural networks. The CNN, usually named as ConvNet, has gained wide attention because of its astonishing ability in classification application based on contextual information [107].

5.1 Convolutional neural network (CNN)

The CNN structure is described in Fig. 6 in which the functionality of four processing layers are discussed such as convolutional layer, pooling layer, fully connected layer and activation function.

Convolution layer correlates the neurons of previous layer in the next layer which is technically termed as receptive field [121]. The functionality of receptive field of neurons is to extract local features that is associated with a particular location of input image [88] and forms a weight vector [86]. The receptive field of neurons share same weight to the next layer due to which same features associated at different location of an image can be detected which is demonstrated in Fig. 7.

Here, the weight vector acts as a kernel or filter which slides over targeted image in order to map its significant features. This process is termed as convolution operation due to which number of filters are mapped together to extract their corresponding features from the input image [125] and reduces training parameters [183]. Thereafter, output in next layer at particular location is computed as discussed in (1).

Pooling layer summarizes the extracted features in patched form. The main issue with output feature map of convolution layer is that it reduces the importance of features location once it has been detected [88]. One process to handle this sensitivity is to down sample the feature map in order to detect the robust changes of location of features in the image which is usually known as local translation invariance [183]. Therefore, Pooling layer is incorporated to down samples the feature maps. The average pooling or max pooling are the most common approaches that are used to significantly reduce the map-size [89].

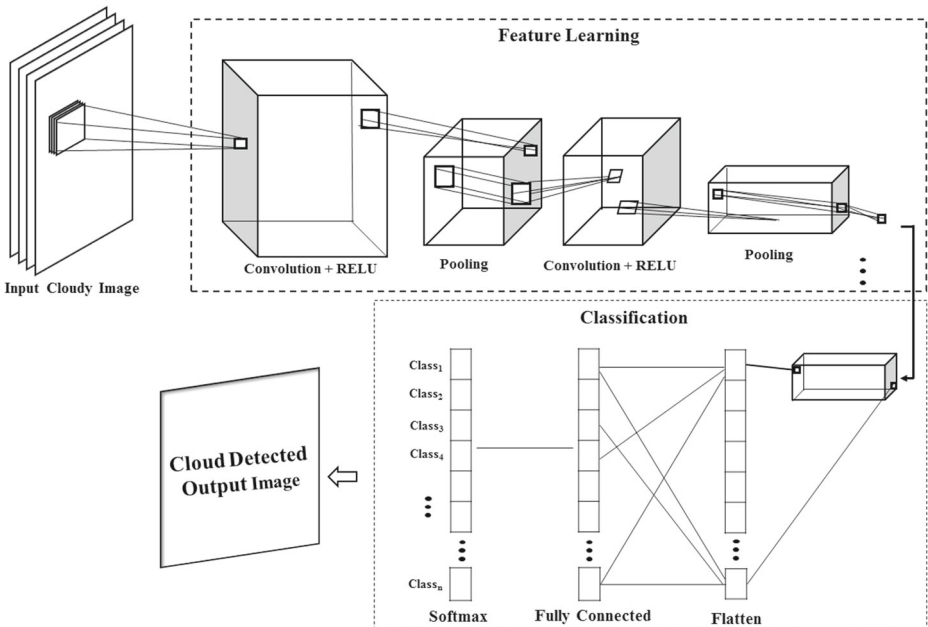


Fig. 6 Architecture of convolutional neural network for cloud detection

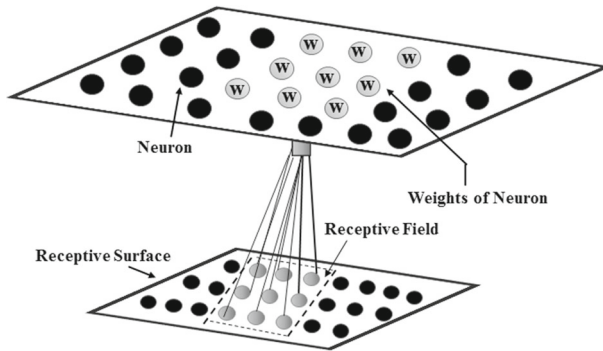


Fig. 7 Neuron receptive field towards next layer in NN architecture

Fully connected layer is equivalent to the concept that is used to describe the artificial neural network (ANN). The output of convolution and then pooling layer is forwarded to the fully connected layer. The output of the fully connected layer is computed as formulated in Eq. 1. The limitation gap of this layer is high computational time complexity to train the big data [183].

Activation function is introduced to incorporate the non-linearity in the deep learning model. The rectified linear unit (RELU) is the most frequently used activation function which returns 0 with negative input value whereas positive input is returned as it is in the output side. This simple approach benefits the model to operate in the non-linear environment [82].

5.2 Deep learning models

Cloud detection is a challenging pre-processing task in any applications of remote sensing images with limited number of spectral bands. Therefore, deep learning algorithm has acquired an immense breakthrough in the field of computer vision as an algorithm in artificial intelligence [30]. The convolutional neural network (CNN) as a deep learning algorithm has raised the performance of applications such as image segmentation [75, 82, 184], object detection, semantic segmentation and so on. [106, 172] proposed intensive predictor with fully CNN (FCNN) without fully connecting the entire layers. Moreover, this structure maps the segmented image of required size as well as enhances the processing speed over conventional CNN method. The FCNN approach is observed further to discriminate cloud and snow in the satellite image because of their similar low-level features such as color distribution and textural pattern [106]. The low-level feature based conventional methods [188] overlook the local textural pattern [185] and high semantic information which are the important element to discriminate cloud over snow region [13, 158]. Moreover, conventional local feature method difficult to model the semantic information such as cloud is enclosed with its shadow or snow is extended along the hills. However, FCNN deep learning algorithm has been used to learn their abstract pattern from its underneath layers [186]. Thereafter, bidirectional independent recurrent neural network has been introduced to extract deep features of images and perform the semantic segmentation; however the approach requires to be modified for multi-modal scenario [179].

Based on the principle of FCNN, the structure is modified, namely, multi-scale features convolutional neural network (MF-CNN) to get the multi-scale global features to discriminate the thick and thin clouds in Landsat 8 satellite images. This is due to the fact that MF-CNN integrates the spatial information at the low-level whereas it provides semantic information at high-level of its structure. This increases the accuracy to detect individual clouds in the combination image [81, 117, 143]. Normally, the analysis of thin cloud is in contrast over thick clouds such as in cloud removal and target detection application for which multi-level cloud detection is required to be addressed [19, 167]. Therefore, a deep CNN with two branches has been designed to get the multiscale features and to predict one of the three classes in which thick clouds, thin clouds and clear pixels has been presented [33, 169]. However, [182] proposed the dual branch CNN, named as multi-scale fusion gated network (MFGNet) to fuse the features at various depths and scales for cloud detection in Gaofen-5 (GF-5) database. [155] used deep neural network approach with feature fusing approach to extract cloud information by using temporal data of Meteosat second generation (MSG) satellite. With the similar approach, [18] has used SegNet with 13 convolutional layers and 13 deconvolution layers to distinguish thick cloud, thin cloud, cloud shadow and clear pixels in Landsat satellite images with multi-level spatial and spectral features. Further, multilevel feature fused segmentation network (MFFSNet) [171] has been proposed to discriminate cloud and its shadow information with the approach of pyramid pooling module. Similarly, [97, 98, 173] has developed multi-scale convolutional feature fusion (MSCFF) to map both local and global features of cloud and cloud shadow by incorporating encoder-decoder module. The identical encoder-decoder structure has been utilized with CNN architecture in [47], named as Cloud-AttU model to fuse the multi-scale features for cloud detection. Similarly, encoder-decoder architecture has been proposed in [48], named as, CDnetV2 by incorporating adaptive feature fusing model (AFFM) at the encoder side whereas decoder structure consists high-level semantic information guidance flows (HSIGFs) to get information of cloud location.

Further, cloud-net algorithm has been designed and trained fully convolutional neural network with local and global features of various blocks of Landsat 8 database [74, 109, 116, 176]. With the similar approach, [146] modified CNN model where input image is divided into super-pixel [161] as sub-region and then features are extracted from each sub-region with four convolutional layer and two fully-connected layers in CNN model to extract deep features of thick and thin cloudy pixels. Identically, [94] proposes a weekly supervised deep learning method for cloud detection (WDCD) on each block of images by incorporating modified global convolutional pooling (GCP) operation. It is observed that most of the algorithms are designed to handle multi-spectral satellite images. However, a few deep learning algorithm extended the methodology to handle hyper-spectral satellite images. Triple-attention Guided Residual Dense and BiLSTM networks(TARDB-Net) has been introduced to select essential spatial and spectral features for classification in hyper-spectral satellite images [16, 49, 50]. Further, CloudScout algorithm based on CNN has been deployed for nano satellite payload where the approach detect the cloudy image directly in on-board satellite and send only the cloud free images to ground station to develop the low power embedded application [44].

The major advantage of deep CNN approach as compare to conventional machine learning algorithms is that it can automatically detect the significant features for high dimensional data. This study has discussed the basic model of deep convolutional neural network and its extensive ability to use as a classifier to detect clouds in satellite images. Moreover,

it is observed that the immense use of CNN approach in various applications such as high-resolution data, medical images, speech recognition and so on. Therefore, present study is expected to add a value in extending the knowledge of deep learning and provide a broad understanding who like to venture more in these applications.

6 Conclusion

The effects of cloud on weather prediction made its detection an important pre-processing task of remote sensing. This article discusses the important contribution of spectral and textural parameters to detect cloud, its shadow, cloud removal and classification of different types of cloud from multi-spectral satellite images. The major issue is to optimize the enhanced observational capacity of different radiometers using retrieved cloud parameters. To investigate this problem, parameters associated with cloud are sorted into spectral, texture, albedo and other features. From the algorithmic perspective there is a significant discussion is done on feature selection and weight optimization in neural network (NN) architecture. In NN based techniques, Levenberg-Marquardt, supervised learning rule is found better in terms of ellapse time and SOFM, an unsupervised learning algorithm is found better in terms of accuracy. Night time and day time simultaneous cloud detection analysis are important challenging issue using NN based techniques. Genetic algorithm is observed to fill the limitation gap of this techniques. Hence artificial neural network trends with evolutionary algorithm was chosen to increase its ability of generalization with low computation time and better accuracy. Further, the discussion on cloud detection algorithm has been extended with deep neural network approach. This study has emerged as a prominent approach towards classification method with contextual features. It is expected that it will provide a broad understanding for various other classification based application to various venture of this field.

Declarations

Research involving Human Participants and/or Animal This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

Conflict of Interests The authors do not have any conflicts of interest.

References

1. Ackerman SA, Holz RE, Frey R, Eloranta EW, Maddux BC, McGill M (2008) Cloud detection with modis. part ii: validation. *J Atmos Ocean Technol* 25(7):1073–1086
2. Alonso-Montesinos J, Martinez-Durban M, del Sagrado J, del Aguila IM, Batlles FJ (2016) The application of bayesian network classifiers to cloud classification in satellite images. *Renew Energy* 97:155–161
3. Angiulli F (2005) Fast condensed nearest neighbor rule. In: *Proceedings of the 22nd international conference on Machine learning. ACM*, pp 25–32
4. Arking A (1964) Latitudinal distribution of cloud cover from tiros iii photographs. *Science* 143(3606):569–572

5. Azimi-Sadjadi MR, Zekavat SA (2000) Cloud classification using support vector machines. In: Geoscience and remote sensing symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 international, vol 2. IEEE, pp 669–671
6. Azimi-sadjadi MR, Shaikh MA, Tian B, Eis KE, Reinke D (1996) Neural network-based cloud detection/classification using textural and spectral features. In: Geoscience and remote sensing symposium, 1996. IGARSS'96.'remote sensing for a sustainable future.', international, vol 2. IEEE, pp 1105–1107
7. Bai T, Li D, Sun K, Chen Y, Li W (2016) Cloud detection for high-resolution satellite imagery using machine learning and multi-feature fusion. *Remote Sens* 8(9):715
8. Bankert RL (1994) Cloud classification of avhrr imagery in maritime regions using a probabilistic neural network. *J Appl Meteorol* 33(8):909–918
9. Bankert RL, Aha DW (1996) Improvement to a neural network cloud classifier. *J Appl Meteorol* 35(11):2036–2039
10. Bankert RL, Wade RH (2007) Optimization of an instance-based goes cloud classification algorithm. *J Appl Meteorol Climatol* 46(1):36–49
11. Bell GJ, Wong MC (1981) The near-infrared radiation received by satellites from clouds. *Mon Weather Rev* 109(10):2158–2163
12. Berendes T, Kuo KS, Logar AM, Corwin EM, Welch RM, Baum BA, Pretre A, Weger RC (1999) A comparison of paired histogram, maximum likelihood, class elimination, and neural network approaches for daylight global cloud classification using avhrr imagery. *J Geophys Res* 104(6):6199–6213
13. Bloschinskiy VD, Kuchma MO, Andreev AI, Sorokin AA (2020) Snow and cloud detection using a convolutional neural network and low-resolution data from the electro-l no. 2 satellite. *J Appl Remote Sens* 14(3):034506
14. Bowker DE, Davis RE, Myrick DL, Stacy K, Jones WT (1985) Spectral reflectances of natural targets for use in remote sensing studies. National Aeronautics and Space Administration Scientific and Technical Information Branch
15. Brown M, Lewis HG, Gunn SR (2000) Linear spectral mixture models and support vector machines for remote sensing. *IEEE Trans Geosci Remote Sens* 38(5):2346–2360
16. Cai W, Liu B, Wei Z, Li M, Kan J (2021) Tardb-net: triple-attention guided residual dense and bilstm networks for hyperspectral image classification. *Multimed Tools Appl*, 1–22
17. Cerdena A, Gonzalez A, Perez JC (2007) Remote sensing of water cloud parameters using neural networks. *J Atmos Ocean Technol* 24(1):52–63
18. Chai D, Newsam S, Zhang HK, Qiu Y, Huang J (2019) Cloud and cloud shadow detection in landsat imagery based on deep convolutional neural networks. *Remote Sens Environ* 225:307–316
19. Chen Y, Fan R, Bilal M, Yang X, Wang J, Li W (2018) Multilevel cloud detection for high-resolution remote sensing imagery using multiple convolutional neural networks. *ISPRS Int J Geo-Inform* 7(5):181
20. Chen PC, Pavlidis T (1979) Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm. *Comput Graph Image Process* 10(2):172–182
21. Chen P-Y, Srinivasan R, Fedosejevs G (2003) An automated cloud detection method for daily noaa 16 advanced very high resolution radiometer data over texas and mexico. *Journal of Geophysical Research: Atmospheres* 108(D23)
22. Chen PY, Srinivasan R, Fedosejevs G, Narasimhan B (2002) an automated cloud detection method for daily noaa-14 avhrr data for texas, usa. *Int J Remote Sens* 23(15):2939–2950
23. Cheng H-Y, Lin CL (2017) Cloud detection in all-sky images via multi-scale neighborhood features and multiple supervised learning techniques. *Atmospheric Measurement Techniques*, 10(1)
24. Cheng Q, Shen H, Zhang L, Yuan Q, Zeng C (2014) Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal mrf model. *ISPRS J Photogramm Remote Sens* 92:54–68
25. Chethan HK, Hemantha Kumar G, Raghavendra R (2009) Texture based approach for cloud classification using svm. In: Advances in recent technologies in communication and computing, 2009. ARTCom'09. International conference on. IEEE, pp 688–690
26. Coakley JA Jr (1914) Reflectance, albedo, surface. encyclopedia of the atmosphere. Holton Jr and Curry Ja (Eds)
27. Coakley JA, Bretherton FP (1982) Cloud cover from high-resolution scanner data- detecting and allowing for partially filled fields of view. *J Geophys Res* 87(C7):4917–4932
28. Cote S, Tatnall ARL (1995) A neural network-based method for tracking features from satellitesensor images. *Int J Remote Sens* 16(18):3695–3701
29. Crane RG, Anderson MR (1984) Satellite discrimination of snow/cloud surfaces. *Int J Remote Sens* 5(1):213–223

30. Cromwell E, Flynn D (2019) Lidar cloud detection with fully convolutional networks. In: 2019 IEEE Winter conference on applications of computer vision (WACV). IEEE, pp 619–627
31. Deng C, Li Z, Wang W, Wang S, Tang L, Bovik AC (2018) Cloud detection in satellite images based on natural scene statistics and gabor features. *IEEE Geosci Remote Sens Lett* 16(4):608–612
32. Desbois M, Seze G (1984) Use of space and time sampling to produce representative satellite cloud classifications. In: *Annales geophysicae*, vol 2. Gauthier-villars, pp 599–605
33. Dev S, Manandhar S, Lee YH, Winkler S (2019) Multi-label cloud segmentation using a deep network. In: 2019 USNC-URSI Radio science meeting (joint with AP-s symposium). IEEE, pp 113–114
34. Ebert E (1987) A pattern recognition technique for distinguishing surface and cloud types in the polar regions. *J Clim Appl Meteorol* 26(10):1412–1427
35. Faure T, Isaka H, Guillemet B (2001) Neural network retrieval of cloud parameters of inhomogeneous and fractional clouds: Feasibility study. *Remote Sens Environ* 77(2):123–138
36. Faure T, Isaka H, Guillemet B (2001) Mapping neural network computation of high-resolution radiant fluxes of inhomogeneous clouds. *J Geophys Res Atmospheres* 106(D14):14961–14973
37. Faure T, Isaka H, Guillemet B (2001) Neural network analysis of the radiative interaction between neighboring pixels in inhomogeneous clouds. *J Geophys Res Atmospheres* 106(D13):14465–14484
38. Feijt A, De Valk P, Van Der Veen S (2000) Cloud detection using meteosat imagery and numerical weather prediction model data. *J Appl Meteorol* 39(7):1017–1030
39. Filippi AM, Archibald R (2009) Support vector machine-based endmember extraction. *IEEE Trans Geosci Remote Sens* 47(3):771–791
40. Fisher A (2014) Cloud and cloud-shadow detection in spot5 hrg imagery with automated morphological feature extraction. *Remote Sens* 6(1):776–800
41. Fleming JR, Cox SK (1974) Radiative effects of cirrus clouds. *J Atmos Sci* 31(8):2182–2188
42. GOES (2020) image data sets <https://www.ncdc.noaa.gov/data-access/satellite-data/goes-r-series-satellites>
43. Garand L (1988) Automated recognition of oceanic cloud patterns. part i: Methodology and application to cloud climatology. *J Clim* 1(1):20–39
44. Giuffrida G, Diana L, de Gioia F, Benelli G, Meoni G, Donati M, Fanucci L (2020) Cloudscout: A deep neural network for on-board cloud detection on hyperspectral images. *Remote Sens* 12(14):2205
45. Grant IP, Hunt GE (1969) Discrete space theory of radiative transfer. ii. stability and non-negativity. In: *Proceedings of the royal society of london a: Mathematical, physical and engineering sciences*, vol 313. The Royal Society, pp 199–216
46. Greaves JR, Chang DT (1970) Technique Development to permit optimum use of satellite radiation data. Final Report, NASA Goddard Space Flight Center. Greenbelt, Maryland
47. Guo Y, Cao X, Liu B, Gao M (2020) Cloud detection for satellite imagery using attention-based u-net convolutional neural network. *Symmetry* 12(6):1056
48. Guo J, Yang J, Yue H, Tan H, Hou C, Li K (2020) Cdnets2: Cnn-based cloud detection for remote sensing imagery with cloud-snow coexistence. *IEEE Transactions on Geoscience and Remote Sensing*
49. Gupta R, Nanda SJ (2019) A binary nsga-iii for unsupervised band selection in hyper-spectral satellite images. In: 2019 IEEE Congress on evolutionary computation (CEC). IEEE, pp 522–529
50. Gupta R, Nanda SJ (2019) Many-objective b/nsga-iii for band selection in cloud contaminated hyper-spectral images. In: 2019 International conference on information technology (ICIT). IEEE, pp 348–352
51. Gupta R, Nanda SJ, Shukla UP (2019) Cloud detection in satellite images using multi-objective social spider optimization. *Appl Soft Comput* 79:203–226
52. Gupta R, Panchal P (2014) Performance analysis and verification of cloud detection algorithm in the presence of various ecological types. In: *Emerging technology trends in electronics, communication and networking (ETTECN)*, 2014 2nd international conference on. IEEE, pp 1–6
53. Gutman G, Tarpley D, Ohring G (1987) Cloud screening for determination of land surface characteristics in a reduced resolution satellite data set. *Int J Remote Sens* 8(6):859–870
54. Hagolle O, Huc M, Pascual DV, Dedieu G (2010) A multi-temporal method for cloud detection, applied to formosat-2, ven μ s, landsat and sentinel-2 images. *Remote Sens Environ* 114(8):1747–1755
55. He Quan-Jun (2011) A daytime cloud detection algorithm for fy-3a/virr data. *Int J Remote Sens* 32(21):6811–6822
56. He Q (2013) Night-time cloud detection for fy-3a/virr using multispectral thresholds. *Int J Remote Sens* 34(8):2876–2887
57. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR (2012) Improving neural networks by preventing co-adaptation of feature detectors. arXiv:1207.0580
58. Holyer RJ, Peckinpaugh SH (1989) Edge detection applied to satellite imagery of the oceans. *IEEE Trans Geosci Remote Sens* 27(1):46–56

59. Hong Y, Hsu K-L, Sorooshian S, Gao X (2004) Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J Appl Meteorol* 43(12):1834–1853
60. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Nat Acad Sci* 79(8):2554–2558
61. Hsu K-L, Gao X, Sorooshian S, Gupta HV (1997) Precipitation estimation from remotely sensed information using artificial neural networks. *J Appl Meteorol* 36(9):1176–1190
62. Hsu K-L, Gupta HV, Gao X, Sorooshian S (1999) Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resour Res* 35(5):1605–1618
63. Hsu K-L, Sorooshian S, Gao X, Gupta HV (1997) Rainfall estimation from satellite infrared imagery using artificial neural networks. Technical report, Department of Hydrology and Water Resources, University of arizona (tucson AZ)
64. Hughes M, Hayes D (2014) Automated detection of cloud and cloud shadow in single-date landsat imagery using neural networks and spatial post-processing. *Remote Sens* 6(6):4907–4926
65. Hulley GC, Hook SJ (2008) A new methodology for cloud detection and classification with aster data. *Geophys Res Lett*, 35(16)
66. Hunt GE (1973) Radiative properties of terrestrial clouds at visible and infra-red thermal window wavelengths. *R Meteorologic Soc Quarter J* 99:346–369
67. Islam J, Li X, Mei Y (2017) A time-varying transfer function for balancing the exploration and exploitation ability of a binary pso. *Appl Soft Comput* 59:182–196
68. Jacobowitz H (1970) EM [SSION Scattering And Absorption Of Radiation In Cirrus Cloud Layers. PhD thesis, Massachusetts Institute of Technology
69. Jang J-D, Viau AA, Anctil F (2004) Neural network estimation of air temperatures from avhrr data. *Int J Remote Sens* 25(21):4541–4554
70. Jang J-D, Viau AA, Anctil F, Bartholome E (2006) Neural network application for cloud detection in spot vegetation images. *Int J Remote Sens* 27(4):719–736
71. Jedlovec GJ, Haines SL, La Fontaine FJ (2008) Spatial and temporal varying thresholds for cloud detection in goes imagery. *IEEE Trans Geosci Remote Sens* 46(6):1705–1717
72. Jensen JR (2009) Remote sensing of the environment: An earth resource perspective 2/e. Pearson Education India, Chapter 2
73. Jeppesen JH, Jacobsen RH, Inceoglu F, Toftegaard TS (2019) A cloud detection algorithm for satellite imagery based on deep learning. *Remote Sens Environ* 229:247–259
74. Ji S, Dai P, Lu M, Zhang Y (2020) Simultaneous cloud detection and removal from bitemporal remote sensing images using cascade convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*
75. Johnston T, Young SR, Hughes D, Patton RM, White D (2017) Optimizing convolutional neural networks for cloud detection. In: *Proceedings of the machine learning on HPC environments*, pp 1–9
76. Joseph Hughes M, Hayes DJ (2014) Automated detection of cloud and cloud shadow in single-date landsat imagery using neural networks and spatial post-processing. *Remote Sens* 6(6):4907–4926
77. Kaminsky EJ, Barad H, Brown W (1997) Textural neural network and version space classifiers for remote sensing. *Int J Remote Sens* 18(4):741–762
78. Kazantzidis A, Eleftheratos K, Zerefos CS (2011) Effects of cirrus cloudiness on solar irradiance in four spectral bands. *Atmos Res* 102(4):452–459
79. Knottenberg H, Raschke E (1982) On the discrimination of water and ice clouds in multispectral avhrr-data. *Annalen der Meteorologie*, (18):145–147
80. Kohonen T (2012) Self-organization and associative memory, vol 8. Springer Science & Business Media, Berlin
81. Kristollari V, Karathanassi V (2020) Convolutional neural networks for detecting challenging cases in cloud masking using sentinel-2 imagery. In: *Eighth international conference on remote sensing and geoinformation of the environment (RSCy2020)*, vol 11524. International Society for Optics and Photonics, p 115240k
82. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp 1097–1105
83. Kuenning JA, McKee TB, Cox SK (1978) A laboratory investigation of radiative transfer in cloud fields. Department of atmospheric science colorado state university
84. Kuhn PM (1963) Measured effective long-wave emissivity of clouds. *Mon Weather Rev* 91(10):635–640
85. Landsat (2020) image data sets <https://earthexplorer.usgs.gov>
86. Lawrence S, Giles CL, Tsoi AC, Back AD (1997) Face srecognition: a convolutional neural-network approach. *IEEE Trans Neural Netw* 8(1):98–113

87. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
88. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
89. Lee KB, Cheon S, Kim CO (2017) A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes. *IEEE Trans Semicond Manuf* 30(2):135–142
90. Lee Y, Wahba G, Ackerman SA (2004) Cloud classification of satellite radiance data by multicategory support vector machines. *J Atmos Ocean Technol* 21(2):159–169
91. Lee J, Weger RC, Sengupta SK, Welch RM (1990) A neural network approach to cloud classification. *IEEE Trans Geosci Remote Sens* 28(5):846–855
92. Lewis HG, Cote S (1995) ARL Tatnall A Neural network approach to cloud classification from multi-temporal satellite imagery
93. Lewis HG, Cote S, Tatnall ARL (1997) Determination of spatial and temporal characteristics as an aid to neural network cloud classification. *Int J Remote Sens* 18(4):899–915
94. Li Y, Chen W, Zhang Y, Tao C, Xiao R, Tan Y (2020) Accurate cloud detection in high-resolution remote sensing imagery by weakly supervised deep learning. *Remote Sens Environ* 250:112045
95. Li P, Dong L, Xiao H, Xu M (2015) A cloud image detection method based on svm vector machine. *Neurocomputing* 169:34–42
96. Li Q, Lu W, Yang J (2011) A hybrid thresholding algorithm for cloud detection on ground-based color images. *J Atmospher Oceanic Technol* 28(10):1286–1296
97. Li Z, Shen H, Cheng Q, Liu Y, You S, He Z (2019) Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors. *ISPRS J Photogramm Remote Sens* 150:197–212
98. Li Z, Shen H, Wei Y, Cheng Q, Yuan Q (2018) Cloud detection by fusing multi-scale convolutional features. *ISPRS Annals of the photogrammetry, Remote Sensing and Spatial Information Sciences* 4(3):149–152
99. Li H, Zheng H, Han C, Wang H, Miao M (2018) Onboard spectral and spatial cloud detection for hyperspectral remote sensing images. *Remote Sens* 10(1):152
100. Liljas E (1986) Use of the AVHRR 3.7 micrometer channel in multispectral cloud classification. Swedish Meteorological and Hydrological Institute, Swedish
101. Liou K-N (1973) Transfer of solar irradiance through cirrus cloud layers. *J Geophys Res* 78(9):1409–1418
102. Liou R-J, Azimi-sadjadi MR, Reinke DL, Vonder-Haar TH, Eis KE (1994) Detection and classification of cloud data from geostationary satellite using artificial neural networks. In: *Neural networks, 1994. IEEE world congress on computational intelligence., 1994 IEEE international conference on*, vol 7. IEEE, pp 4327–4332
103. Lisens G, Kempencers P, Fierens F, Van Rensbergen J (2000) Development of cloud, snow, and shadow masking algorithms for vegetation imagery. In: *Geoscience and remote sensing symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 international*, vol 2. IEEE, pp 834–836
104. Liu H, Zeng D, Tian Q (2018) Super-pixel cloud detection using hierarchical fusion cnn. In: *2018 IEEE Fourth international conference on multimedia big data (bigMM)*. IEEE, pp 1–6
105. Lo RC, Johnson DR (1971) An investigation of cloud distribution from satellite infrared radiation data. *Mon Weather Rev* 99:599–605
106. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 3431–3440
107. Maggiori E, Tarabalka Y, Charpiat G, Alliez P (2016) Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Trans Geosci Remote Sens* 55(2):645–657
108. Maktav D (2009) Remote Sensing for a Changing Europe: Proceedings of the 28th Symposium of the European Association of Remote Sensing Laboratories, Istanbul, Turkey, 2-5 June 2008. IOS Press, Amsterdam
109. Mateo-García G, Laparra V, López-Puigdollers D, Gómez-Chova L (2020) Transferring deep learning models for cloud detection between landsat-8 and proba-v. *ISPRS J Photogramm Remote Sens* 160:1–17
110. Mazzoni D, Garay MJ, Davies R, Nelson D (2007) An operational misr pixel classifier using support vector machines. *Remote Sens Environ* 107(1):149–158
111. Melgani F (2006) Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Trans Geosci Remote Sens* 44(2):442–455
112. Miller SW, Emery WJ (1997) An automated neural network cloud classifier for use over land and ocean surfaces. *J Appl Meteorol* 36(10):1346–1362

113. Miller WT, Glanz FH, Kraft LG (1990) Cmas: An associative neural network alternative to backpropagation. *Proc IEEE* 78(10):1561–1567
114. Minnis P, Harrison EF (1984) Diurnal variability of regional cloud and clear-sky radiative parameters derived from goes data. part iii: November 1978 radiative parameters. *J Climate Appl Meteorol* 23(7):1032–1051
115. Minnis P, Harrison EF, Gibson GG (1987) Cloud cover over the equatorial eastern pacific derived from july 1983 international satellite cloud climatology project data using a hybrid bispectral threshold method. *J Geophys Res Atmospheres* 92(D4):4051–4073
116. Mohajerani S, Saeedi P (2019) Cloud-net: An end-to-end cloud detection algorithm for landsat 8 imagery. In: *IGARSS 2019-2019 IEEE International geoscience and remote sensing symposium*. IEEE, pp 1029–1032
117. Morales G, Huamán SG, Telles J (2018) Cloud detection in high-resolution multispectral satellite imagery using deep learning. In: *International conference on artificial neural networks*. Springer, pp 280–288
118. Murao H, Nishikawa I, Kitamura S, Yamada M, Xie P (1993) A hybrid neural network system for the rainfall estimation using satellite imagery. In: *Neural Networks, 1993. IJCNN'93-Nagoya. Proceedings of 1993 International Joint Conference on*, vol 2. IEEE, pp 1211–1214
119. NASA (2016) <https://isccp.giss.nasa.gov/products/onlinedata.html>
120. NOAA (2020) Environmental visualization laboratory <http://www.nnvl.noaa.gov/imagegallery.php>
121. Nebauer C (1998) Evaluation of convolutional neural networks for visual recognition. *IEEE Trans Neural Netw* 9(4):685–696
122. Ning X, Duan P, Li W, Zhang S (2020) Real-time 3d face alignment using an encoder-decoder network with an efficient deconvolution layer. *IEEE Signal Process Lett* 27:1944–1948
123. Ning X, Gong K, Li W, Zhang L, Bai X, Tian S (2020) Feature refinement and filter network for person re-identification. *IEEE Transactions on Circuits and Systems for Video Technology*
124. OF TIROS-N (1979) Data extraction and calibration of tiros-n/noaa radiometers
125. Palsson F, Sveinsson JR, Ulfarsson MO (2017) Multispectral and hyperspectral image fusion using a 3-d-convolutional neural network. *IEEE Geosci Remote Sens Lett* 14(5):639–643
126. Parikh J (1977) A comparative study of cloud classification techniques. *Remote Sens Environ* 6(2):67–81
127. Peak JE (1991) Application of neural networks to large-scale cloud pattern recognition. Technical report, DTIC Document
128. Peak JE (1991) Neural network methodologies and their potential application to cloud pattern recognition. Technical report, DTIC Document
129. Peak JE, Tag PM (1994) Segmentation of satellite imagery using hierarchical thresholding and neural networks. *J Appl Meteorol* 33(5):605–616
130. Peak JE, Tag PM (1994) Segmentation of satellite imagery using hierarchical thresholding and neural networks. *J Appl Meteorol* 33(5):605–616
131. Pickett RM (1970) Visual analysis of texture in the detection and recognition of objects. *Picture Processing and Psychopictorics*, 289–308
132. Raschke E, Rossow W, Schiffer R (1987) The international satellite cloud climatology project: preliminary results and its potential aspects. *Adv Space Res* 7(3):137–145
133. Rossow WB (1989) Measuring cloud properties from space: A review. *J Clim* 2(3):201–213
134. Rossow WB, Garder LC, Laciis AA (1989) Global, seasonal cloud variations from satellite radiance measurements. part i: Sensitivity of analysis. *J Clim* 2(5):419–458
135. Rumelhart DE, Hinton GE, Williams RJ (1985) Learning internal representations by error propagation. Technical report, DTIC Document
136. Sam Navin M, Agilandeswari L (2020) Multispectral and hyperspectral images based land use/land cover change prediction analysis: an extensive review. *Multimed Tools Appl* 79(39):29751–29774
137. Saunders RWL, Kriebel KTS (1988) An improved method for detecting clear sky and cloudy radiances from avhrr data. *Int J Remote Sens* 9(1):123–150
138. Savory TH (1928) *The biology of spiders*. Sidgwick And Jackson Ltd; London
139. Schiffer RA, Rossow WB (1983) The international satellite cloud climatology project(isccp)- the first project of the world climate research programme. *Amer Meteorologic Soc Bull* 64:779–784
140. Sedano F, Kempeneers P, Strobl P, Kucera J, Vogt P, Seebach L, San-Miguel-Ayanz J (2011) A cloud mask methodology for high resolution remote sensing data combining information from high and medium resolution optical sensors. *ISPRS J Photogramm Remote Sens* 66(5):588–596
141. Segal-Rozenhaimer M, Li A, Das K, Chirayath V (2020) Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (cnn). *Remote Sens Environ* 237:111446

142. Seze G, Rossow WB (1991) Time-cumulated visible and infrared radiance histograms used as descriptors of surface and cloud variations. *Int J Remote Sens* 12(5):877–920
143. Shao Z, Pan Y, Diao C, Cai J (2019) Cloud detection in remote sensing images based on multiscale features-convolutional neural network. *IEEE Trans Geosci Remote Sens* 57(6):4062–4076
144. Shen H, Li H, Qian Y, Zhang L, Yuan Q (2014) An effective thin cloud removal procedure for visible remote sensing images. *ISPRS J Photogramm Remote Sens* 96:224–235
145. Shenk WE, Holub RJ, Neff RA (1976) A multispectral cloud type identification method developed for tropical ocean areas with nimbus-3 mrir measurements. *Mon Weather Rev* 104(3):284–291
146. Shi M, Xie F, Zi Y, Yin J (2016) Cloud detection of remote sensing images by deep learning. In: 2016 IEEE International geoscience and remote sensing symposium (IGARSS). IEEE, pp 701–704
147. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
148. Slawinski O, Kowalski JG, Cornillon PC (1991) A neural network approach to cloud detection in avhrr images. In: *Neural networks, 1991., IJCNN-91-seattle international joint conference on, vol 1*. IEEE, pp 283–288
149. Specht DF (1990) Probabilistic neural networks. *Neural Netw* 3(1):109–118
150. Stowe LL, Wellemeyer CG, Yeh HYM, Eck TF (1988) Nimbus-7 global cloud climatology. part i: Algorithms and validation. *J Clim* 1(5):445–470
151. Strahler AH, Muller JP, Lucht W, Schaaf C, Tsang T, Gao F, Li X, Lewis P, Barnsley MJ (1999) Modis brdf/albedo product: algorithm theoretical basis document version 5.0. MODIS Documentation 23(4):42–47
152. Sun L, Yang X, Jia S, Jia C, Wang Q, Liu X, Wei J, Zhou X (2020) Satellite data cloud detection using deep learning supported by hyperspectral data. *Int J Remote Sens* 41(4):1349–1371
153. Szejwach G (1982) Determination of semi-transparent cirrus cloud temperature from infrared radiances: Application to meteosat. *J Appl Meteorol* 21(3):384–393
154. Tian B, Azimi-Sadjadi MR, Haar THV, Reinke D (2000) Temporal updating scheme for probabilistic neural network with application to satellite cloud classification. *IEEE Trans Neural Netw* 11(4):903–920
155. Tuia D, Kellenberger B, Pérez-Suey A, Camps-Valls G (2018) A deep network approach to multi-temporal cloud detection. In: *IGARSS 2018-2018 IEEE International geoscience and remote sensing symposium*. IEEE, pp 4351–4354
156. Walder P, MacLaren I (2000) Neural network based methods for cloud classification on avhrr images. *Int J Remote Sens* 21(8):1693–1708
157. Wallach I, Dzamba M, Heifets A (2015) Atomnet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv:1510.02855
158. Wang L, Chen Y, Tang L, Fan R, Yao Y (2018) Object-based convolutional neural networks for cloud and snow detection in high-resolution multispectral imagers. *Water* 10(11):1666
159. Watmough GR, Atkinson PM, Hutton CW (2011) A combined spectral and object-based approach to transparent cloud removal in an operational setting for landsat etm+. *Int J Appl Earth Observ Geoinform* 13(2):220–227
160. Watmough GR, Atkinson PM, Hutton CW (2011) A combined spectral and object-based approach to transparent cloud removal in an operational setting for landsat etm+. *Int J Appl Earth Observ Geoinform* 13(2):220–227
161. Wei J, Huang W, Li Z, Sun L, Zhu X, Yuan Q, Liu L, Cribb M (2020) Cloud detection for landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches. *Remote Sens Environ* 248:112005
162. Welch RM, Kuo KS, Wielicki BA, Sengupta SK, Parker L (1988) Marine stratocumulus cloud fields off the coast of southern california observed using landsat imagery. part i: Structural characteristics. *J Appl Meteorol* 27(4):341–362
163. Welch RM, Navar MS, Sengupta SK (1989) The effect of spatial resolution upon texture-based cloud field classifications. *J Geophys Res Atmospheres* 94(D12):14767–14781
164. Welch RM, Sengupta SK, Chen DW (1988) Cloud field classification based upon high spatial resolution textural features: 1. gray level co-occurrence matrix approach. *J Geophys Res Atmospheres* 93(D10):12663–12681
165. Welch RM, Sengupta SK, Goroch AK, Rabindra P, Rangaraj N, Navar MS (1992) Polar cloud and surface classification using avhrr imagery: An intercomparison of methods. *J Appl Meteorol* 31(5):405–420

166. Widrow B, Hoff ME et al (1960) Adaptive switching circuits. In: IRE WESCON Convention record, vol 4. New York, pp 96–104
167. Wu X, Shi Z (2018) Utilizing multilevel features for cloud detection on satellite imagery. *Remote Sens* 10(11):1853
168. Xie F, Shi M, Shi Z, Yin J, Zhao D (2017) Multilevel cloud detection in remote sensing images based on deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*
169. Xie F, Shi M, Shi Z, Yin J, Zhao D (2017) Multilevel cloud detection in remote sensing images based on deep learning. *IEEE J Select Topics Appl Earth Observ Remote Sens* 10(8):3631–3640
170. Yamamoto G, Tanaka M, Asano S (1970) Radiative transfer in water clouds in the infrared region. *J Atmos Sci* 27(2):282–292
171. Yan Z, Yan M, Sun H, Fu K, Hong J, Sun J, Zhang Y, Sun X (2018) Cloud and cloud shadow detection using multilevel feature fused segmentation network. *IEEE Geosci Remote Sens Lett* 15(10):1600–1604
172. Yanan G, Xiaoqun C, Bainian L, Kecheng P (2020) Cloud detection for satellite imagery using deep learning. In: *Journal of physics: Conference series*, vol 1617. IOP Publishing, p 012089
173. Yang J, Guo J, Yue H, Liu Z, Hu H, Li K (2019) Cndet: Cnn-based cloud detection for remote sensing imagery. *IEEE Trans Geosci Remote Sens* 57(8):6195–6211
174. Yang J, Lu W, Ma Y, Yao W (2012) An automated cirrus cloud detection method for a ground-based cloud image. *J Atmos Ocean Technol* 29(4):527–537
175. Yhann SR, Simpson JJ (1995) Application of neural networks to avhrr cloud segmentation. *IEEE Trans Geosci Remote Sens* 33(3):590–604
176. Yin Z, Ling F, Foody GM (2020) Yun Du. Cloud detection in landsat-8 imagery in google earth engine based on a deep neural network. [arXiv:2006.10358](https://arxiv.org/abs/2006.10358)
177. Yoshida T, Omatu S (1994) Neural network approach to land cover mapping. *IEEE Trans Geosci Remote Sens* 32(5):1103–1109
178. Yoshida T, Omatu S (1994) Neural network approach to land cover mapping. *IEEE Trans Geosci Remote Sens* 32(5):1103–1109
179. You H, Tian S, Yu L, Lv Y (2019) Pixel-level remote sensing image recognition based on bidirectional word vectors. *IEEE Trans Geosci Remote Sens* 58(2):1281–1293
180. Young MJ (1967) Variability in estimating total cloud cover from satellite pictures. *J Appl Meteorol* 6(3):573–579
181. Yu D, Deng L (2010) Deep learning and its applications to signal and information processing [exploratory dsp]. *IEEE Signal Proc Mag* 28(1):145–154
182. Yu J, Li Y, Zheng X, Zhong Y, He P (2020) An effective cloud detection method for gaofen-5 images via deep learning. *Remote Sens* 12(13):2106
183. Yu Z, Wang H, Xu F, Jin Y-Q (2016) Polarimetric sar image classification using deep convolutional neural networks. *IEEE Geosci Remote Sens Lett* 13(12):1935–1939
184. Yu W, Yang X, Celik T, Sushkova O, Albertini MK (2020) Multifocus image fusion using convolutional neural network. *Multimedia Tools And Applications*
185. Yuan K, Meng G, Cheng D, Bai J, Xiang S, Pan C (2017) Efficient cloud detection in remote sensing images using edge-aware segmentation network and easy-to-hard training strategy. In: *2017 IEEE International conference on image processing (ICIP)*. IEEE, pp 61–65
186. Zhan Y, Wang J, Shi J, Cheng G, Yao L, Sun W (2017) Distinguishing cloud and snow in satellite images via deep convolutional network. *IEEE Geosci Remote Sens Lett* 14(10):1785–1789
187. Zhang J, Lin X, Ning X (2013) Svm-based classification of segmented airborne lidar point clouds in urban areas. *Remote Sens* 5(8):3749–3775
188. Zhang J, Zhou Q, Shen X, Li Y (2019) Cloud detection in high-resolution remote sensing images using multi-features of ground objects. *J Geovisualiz Spat Anal* 3(2):14
189. Zhong B, Chen W, Wu S, Hu L, Luo X, Liu Q (2017) A cloud detection method based on relationship between objects of cloud and cloud-shadow for chinese moderate to high resolution satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*
190. Zhu H, Wan G (2016) Local contrast preserving technique for the removal of thin cloud in aerial image. *Optik-International Journal for Light and Electron Optics* 127(2):742–747
191. Zhu Z, Wang S, Woodcock CE (2015) Improvement and expansion of the fmask algorithm: cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images. *Remote Sens Environ* 159:269–277
192. Zhu Z, Woodcock CE (2012) Object-based cloud and cloud shadow detection in landsat imagery. *Remote Sens Environ* 118:83–94

193. Zi Y, Xie F, Jiang Z (2018) A cloud detection method for landsat 8 images based on pcanet. *Remote Sens* 10(6):877
194. Zitzler E, Deb K, Thiele L (2000) Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolution Comput* 8(2):173–195

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.