Insights on Deep Learning based Segmentation Schemes Towards Analyzing Satellite Imageries

Natya S[1], Ramya K[2], Dr. Seema Singh[3]

Research Scholar, Department of Electronics and Telecommunication Engineering, BMS Institute of Technology & Management, Bengaluru, Karnataka, India[1]

(Also, Assistant Professor, Department of Electronics and Communication Engineering, Presidency University, Bengaluru, Karnataka, India)[1]

Assistant Professor, Department of Electrical and Electronics Engineering, Presidency University, Bengaluru, Karnataka, India[2]

Professor and Dean (External Relations), Electronics and Telecommunication Engineering Department, BMS Institute of Technology & Management, Bengaluru, Karnataka, India[3]

Abstract—**Satellite imageries are essentially a complex form of an image when subjected to critical analytical operation. The analytical process applied on remotely sensed satellite imageries are utilized for generating the land cover map. With an abundance of traditional techniques evolved to date, deep learning-based schemes are progressively gaining pace for identifying and classifying a terrestrial object in satellite images. However, different variants of deep learning approaches have different operations, and so are the consequences. At the same time, there is no reported literature to highlight the issues, trends, and effectiveness much on a generalized scale concerning segmentation. Therefore, this paper reviews some of the recent segmentation approaches using deep learning to contribute towards review findings in the form of research trends, research gaps, and essential learning outcomes. The paper offers a compact and distinct picture of deep learning approaches used to boost segmentation for satellite images.**

Keywords—Deep learning; landcover; map generation; remotely sense image; satellite image; segmentation

I. INTRODUCTION

The satellite images of various resolutions are used for generating maps for land cover [1]. Analysis of satellite images potentially assists in developing multiple classes of landcover, e.g., water, impervious surface, vegetation surface, residential area, etc. [2]. To construct an accurate landcover map, it is essential to classify various elements, e.g., trees, individual buildings, roads, cars, etc. In this process of building a landcover map, spatial resolution is proven more efficient than spectral resolution [3][4]. It will also mean that image pixels with finer resolution are more beneficial than the maximized number of spectral bands [5]. This is the prime justification behind the increased usage of remote sensing satellite images to enhance the terrestrial object's visibility [6]. In this direction towards generating an accurate landcover map of the satellite image, most existing studies are now emphasized towards pixel-based analysis where segmentation plays a significant role. This is also boosted by applying a deep learning approach towards land cover map generation and classification of objects [7]-[9]. With the availability of various wireless technologies, satellite images are transmitted from a satellite to the earth receiving center, where processing is carried out [10][11]. However, the problem surfaces towards the quality of received signal where signal fluctuations are witnessed due to electrical signals in wireless transmission systems [12]-[14]. This finally leads to various errors and artifacts within the received satellite images [15][16]. Most of these artifacts are in the form of noises, while it is challenging to eliminate coupled noises. Existing filters are not capable of controlling these forms of noises [17]-[19].

The available processing algorithms for satellite images are required to be proven for their efficiencies concerning computational complexities (space and time) that are significantly missing from literature towards wireless image transmission. The pixel-based analysis approach also leads to issues with the increase of spatial resolution; the complex patterns start to surface for spectral response originated from multiple objects of much smaller dimension in an urban region. The prime reason is using a similar object to develop various structures in landcover while emission of similar spectral response is continued. In object-based analysis, multiple segments are generated from an image indexed with different attributes, followed by subjecting it to rules of classification operation (e.g., texture, size, area, length, etc.). However, both pixel and object-based analyses have inefficiencies towards classification, making way for the deep learning approach to contribute to the segmentation process. However, deep learning cannot be fully considered an end solution, although it shows potential progress in analyzing satellite images. Despite some dedicated research approaches towards segmentation problems, deep learning has shown promising results, as well as there are also pitfalls associated with this approach.

The paper presents a discussion of existing deep learning approaches towards improving segmentation mechanism of satellite images. With an increasing proliferation towards adopting deep learning techniques, it is essential to understand its rate of success for analyzing satellite imageries, which itself is one complex form of signal. The significant problem under consideration is that there is no standard reporting for implementation effectiveness of using deep learning methods towards addressing segmentation problem, which is one of the

essential steps in analyzing satellite images. Inspite of availability of different variants of deep learning, the research question arises are i) what are the identified advantages and limitation of different deep learning schemes towards segmentation problem in satellite images? ii) which is the most preferred dataset of satellite imageries considered for existing evaluation, iii) what is the existing direction of research trend of exploiting the potential of deep learning over analyzing satellite images?. Hence, based on the above stated research question.

Hence, this paper studies the effectiveness of existing deep learning schemes for the segmentation of satellite images. The contribution of this review paper are multi-fold viz. i) a compact briefing of satellite imageries concerning conventional segmentation and its associated challenges, ii) exhaustive recent reviews of frequently adopted deep learning segmentation schemes on satellite images, iii) compact briefing of the frequently adopted dataset by the existing researcher, iv) Discussion of research trends of existing deep learning segmentation methods, v) exclusive highlights of research gap and essential findings of this review study of recent papers. The paper's organization is: Section II discusses satellite imageries while existing literature of different deep learning approaches is discussed in Section III. Section IV discusses data adoption while Section V highlights research trends with various perspectives. Section VI discusses review contribution, while Section VII outlines the paper's conclusion.

II. SATELLITE IMAGERIES

Satellite imageries are considered one of the meteorologists' primary sets of information to predict the atmosphere's behavior. Satellite images are of three types, as shown in Fig. 1 viz. water-vapor images, infrared images, and visible images [20][21]. Some of the actual applications of remotely-sensed ideas are viz. i) tracking cloud for weather prediction, ii) monitoring growth of city area, iii) identifying changes in forest and farmland over some time, iv) Mapping and exploring the topography of ocean bed, v) forest fire, etc. Apart from this, satellite images have broader applications, e.g., anomaly hunting, regional planning, cartography, geology, oceanography, agriculture, forestry, etc. [22]. With the modernization, there has been a change in the forms and types of satellite images based on capturing it. Visual sensors play a significant role in this regard, integrated into modern-day satellites to generate remote sensing images [24]. This is a process of identifying different physical characteristics specific to a monitoring region based on emitted radiation from the air-borne vehicle or the satellite. This paper has discussed the case of satellite images in remote sensing, which is characterized by five types of resolution, i.e., geometric, radiometric, temporal, spectral, and spatial [25].

The resolution of satellite images largely depends upon the orbit altitude and types of the instrument being used. This manuscript will not illustrate fundamental theories of satellite images, as information can be easily accessed from various online articles, e.g. [26][27]. Instead, the proposed study will emphasize the challenges of processing it and understanding the existing literature's effectiveness. A series of image

processing is carried out to process satellite images, e.g., enhancement, feature extraction, segmentation, fusion, detection of changes, compression, classification, and feature detection [28]. Out of all these processes, the most critical function that significantly contributes to accuracy in prediction is the segmentation process. The process of segmenting satellite images targets to obtain a distinct segment of boundaries and their objects.

Some of the conventional methods of satellite image segmentation are briefed in Table I which exhibits conventional approaches of segmentation over satellite images e.g. Gabor filter and graph-based [29], firefly algorithm [30], deep learning [31], Markov Random Field [32], and Cuckoo Search [33]. From the preliminary outcomes, it can be seen that performance (especially with respect to accuracy) is higher for deep learning compared to other conventional methods irrespective of any dataset being used. The segmentation techniques mentioned above in Table I are considered as a baseline by various research work; however, the segmentation approaches are still found not to exhibit better predictive accuracy as they should be. There is a reason for this trade-off which is complexities associated with segmenting satellite images which are as follows: i) limitation of coverage are of satellite images, ii) availability of limited information from current satellite data [34], iii) higher possibilities of degradation of image quality during retrieval process [35], and iv) possibilities of generation of artifact-data when aggregated from multi-instrument data. Hence, there are significant challenges in performing proper segmentation for any complex remote sensing satellite images in all the cases mentioned above.

Further, it is noticed that landcover classification is associated with various challenges viz. multitemporal images, presence of clouds, classification of the object, small scale benchmarks, etc. Landcover images of the satellite are considered to be massive, and hence the mining community has already started to utilize the Big Data concept towards mapping and classifying crops [36]. In this respect, deep learning is a potential player to contribute to classification. It is because. However, this is also another factor of motivation to take an interest in working on this topic in the current era. Hence, this paper explores the impact of existing deep learning on the segmentation of satellite images.
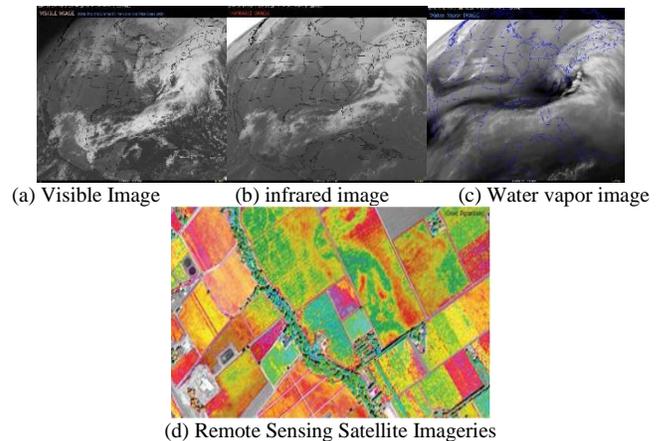

(a) Visible Image   (b) infrared image   (c) Water vapor image

(d) Remote Sensing Satellite Imageries

Fig. 1.   Types of Satellite Images [23].

Segmentation Method	Test Image	Outcome
Gabor filter, graph-based [29]	QuickBird	Quality=85%, Correctness=92% Completeness=93%
Firefly Algorithm [30]	NASA satellite image (multiband)	SSIM=0.92448 MSE=3376.16 PSNR=13.71
Deep Learning [31]	Infrared Image	Accuracy=90%
Markov Random Field [32]	QuickBird	Recall= 0.77, Precision=0.71, F-Measure=0.74
Cuckoo Search [33]	Pleiades Satellite Image	FSIM=0.96, MSE=411.86, PSNR=21.99

III. DEEP LEARNING SCHEMES

Deep learning schemes are constructed based on the neural network that consists of neurons in the form of multiple layers capable of transforming the input of satellite images to an outcome image in the form of identified land covers. This is achieved by progressively learning the high-level features. This section discusses different taxonomies and approaches of deep learning schemes that have been used for image segmentation towards the satellite data as follows:

A. Convolution Neural Network

The generalized framework of Convolution Neural Network (CNN) consists of harnessing the softmax layer using different blocks with distinct architecture and ensemble of its outcome. The conventional practices of CNN comprise different layers, e.g., input, convolution, pooling, completely connected layer, and outcome of the soft-max layer. The computation of the filter adopted in the convolution layer is mathematically represented as follow:

$$\phi_n^l = g\left(\sum_{m \subseteq M_n} \triangle a + b_m^l\right) \qquad (1)$$

In the above expression (1), the outcome of the filter ϕ_n^l is dependent on multiple variables: The variable g represents nonlinear activation function while the first component $\triangle a$ represents $(x_m^{l-1} . w_{m,n}^l)$ where variable x is input during the second component b_m^l Represents bias term considering weight w with m^{th} filter corresponding to l^{th} layer. Further feature maps are produced from progressive usage of pooling and convolution layer that is finally transformed into one-dimensional features leading to final prediction using soft-max layer. Fig. 2 highlights the sequence of the process mentioned above in CNN.
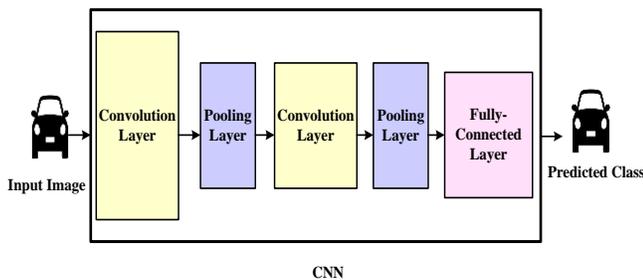

Fig. 2.   Process Flow of CNN.

Further, the process of training in CNN can be boosted by using Adam Optimizer [37], stochastic gradient descent, batch normalization [38], dropout [39], parametric rectified linear unit [40], etc. At present, there are various approaches of CNN being applied towards image segmentation of satellite data.

One of the significant beneficial characteristics of the CNN model is that it doesn't have any dependencies towards tuning the parameters. This fact was investigated in the study of Wurm et al. [41]. The study also analyzes the capabilities of transferring a trained network to a different type of dataset. The idea of this implementation is to carry out semantic segmentation of landcover using CNN. The presented study emphasizes the transfer learning operation of fully CNN.

The study of segmentation carried out by Zhang et al. [42] has extracted road area of land from satellite images using CNN. A raster map is constructed from satellite signal trajectories, where the outcome shows better accuracy. A recent work carried out by Li et al. [43] used CNN for extracting features considering the case study of developing footprint maps of buildings. This mechanism also uses a graph model to consider the spatial correlation of data to retain boundary-related information. Preprocessing is carried out using co-registration and truncated signed distance labels. The CNN inputs satellite image and ground truth, leading to extracted feature where a segmentation probability is computed and pairwise potential extraction. This combined yields a new graph model that finally generates multi-class results emphasizing object detection. A similar category of work has also been carried out by Saetchnikov et al. [44]. Multiple deep neural networks of CNN have been used for comparative analysis to perform segmentation leading to object detection and tracking.

Wang et al. [45] have developed a unique segmentation process that is iterative in its operation to preprocess remote sensing images for detecting ships. The region detection of the ship is carried out using multivariate Gaussian distribution. The training is carried out using optical panchromatic data during a hardware synthesis of the model over a Field Programmable Gate Array. The work carried out by Jiang et al. [46] has captured geographic information of road using remote sensing technology of satellite imageries. This technique uses CNN to classify satellite images into two categories, i.e., road and non-road sections. Further optimization is carried out to address the inclusion issues of non-road noises owing to natural scene factors using wavelet packets. Another unique work was carried out by Persello et al. [47], where the identification of information settlement over a landcover has been investigated. The study has emphasized over-extraction of spatial feature and texture information. The authors have used a complete convolution network where labeling towards the pixel of satellite images has been carried out. The higher representation of the data is autonomously subjected to a learning algorithm considering six layers of convolution network. Different from conventional studies towards land cover, Wang et al. [48] have carried out the study using CNN to find the ice concentration.

TABLE II.   SUMMARY OF EXISTING TECHNIQUES ON CNN

Authors	Problems	Segmentation Technique	Dataset	Advantages	Limitation
Wurm et al. [41].	Slum area segmentation	Fully CNN, transfer function	QuickBird dataset	Results in high-resolution image	Narrowed test cases with low scope
Zhang et al. [42]	Road extraction	CNN, trajectories of GPS	Google Earth Satellite Imageries	Simplified process	No benchmarking
Li et al. [43]	Building footprint map	CNN, graph model	-Inria Aerial image dataset -Kaggle -ISPRS -Planetscope satellite images	Optimal training time	Increased processing time
Jiang et al. [46]	Identification of road	CNN, wavelet packet filter	GIS data	Achieves >4% accuracy	computational intensive
Persello et al. [47]	Identification of informal settlement	Fully convolution neural network	Dar es Salaam	Low computational cost	Higher training time
Saetchnikov et al. [44]	Object detection	Multiple variants of CNN	DSTL Dataset	Supportive of practical application	Computational complexity not addressed.
Wang et al. [48]	Identification of ice concentration	CNN	RADARSAT-2	Simpler process	The lower scale of analysis
Wang et al. [45]	Identification of ship	CNN, multivariate Gaussian distribution, FPGA	Panchromatic data	Usage of few parameters, robust detection	Extensive test environment not adopted

Observation: From this discussion, it can be seen that CNN is used as a standalone and in combination with other schemes to boost the segmentation performance. For the majority of the implementation scheme, the performance of the CNN remains nearly similar with respect to method simplification and accuracy. It is also seen that the usage of CNN is found highly efficient for extracting potential features, classification of the scene, and detection of a specific form of land in satellite images. However, it is observed that features often tend to diminish while using the pooling layer. This will potentially affect the computational performance. In contrast, the outcome of the feature map and predictive resultants is not much improved during satellite image segmentation using CNN.

Table II highlights the strength and weaknesses of the existing segmentation approaches facilitated by CNN to understand the contribution of existing literature.

B. Recurrent Neural Network

Recurrent Neural Network (RNN) formulates a directed / undirected graph obtained from node connection in neural network considering temporal sequence. Input with variable length sequence is processed using internal state of RNN which is a network class for infinite impulse response. Used over wide variety of application in current time, RNN is another frequently used supervised learning model deployed towards segmentation of satellite images. Essentially meant to carry out analysis of discrete sequences, it is found that RNN can generate deep feedforward networks. The conventional architecture of RNN is shown in Fig. 2.

According to Fig. 3, the conventional RNN model usually connects the outcome of all the neurons to the input to construct a topology of the network. RNN is of different types, i.e., one-to-one, one-to-many, many-to-one, and many-to-many. The common activation function in RNN is sigmoid, Tanh, and relu. In this case, the outcome of the previous step is considered an input for the existing step. One of the significant advantages of RNN in the segmentation process is its memory system, which retains information about all its calculations, thereby reducing the complexity of attributes not present in other neural networks. The discussion presented by Taberner et al. [49] gave a good insight into using deep learning over time-series datasets of satellite imageries. The most recent work carried out by Turkoglu et al. [50] has used RNN with multiple layers where gated cells were drawn. The most significant contribution of this study is findings that state a change in gradient magnitude when it moves through the cell over a deep lattice of RNN. The study has used MNIST dataset which is benchmarked dataset consisting of satellite imageries. The work carried out by Ienco et al. [51] has harnessed the potential of RNN to carry out the classification of landcover from satellite images. This technique has carried out segmentation using a multi-temporal stack to obtain information associated with multi-temporal layers of an object. The researcher has used a multiresolution segmentation approach followed by applying statistical evaluation over its features. Maggiori et al. [52] present a similar direction of work, where RNN has been used as an iterative and semantic segmentation process. Sun et al. [53] have implemented Long Short Term Memory RNN, which harnesses temporal factors of crops captured from satellite images over a time series.
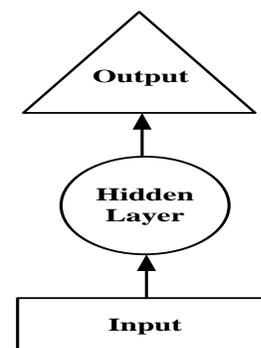

Fig. 3.   Conventional Architecture of RNN.

TABLE III. SUMMARY OF EXISTING TECHNIQUES ON RNN

| Authors	Problems	Segmentation Technique	Dataset	Advantages	Limitation
Turkoglu et al. [50]	Improving the performance of RNN	Gated RNN	MNIST dataset	Computationally improved performance, higher accuracy	Non-scalable performance
Ienco et al. [51]	Classification of landcover	RNN, multi-temporal stack	Pleiades VHSR, Corine Land Cover.	Very simplified approach	Don't consider complexities within the data.
Maggiori et al. [52]	Map classification	RNN, semantic segmentation	OpenStreetMap	Higher accuracy	Higher processing time
Sun et al. [53]	Land cover classification	Semantic segmentation	Cropland Data Layer	Higher accuracy	Resource intensive operation
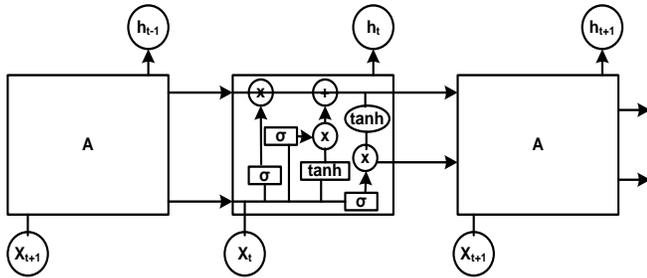
Fig. 4. Conventional Architecture of LSTM.

Observation: RNN is a good option for boosting segmentation performance; however, not much recent research has been carried out in this perspective. Table III highlights the summary.

C. Long Short Term Memory

Long Short Term Memory, also called LSTM, is a typical case of RNN to offer extensive dependencies of long-term learning. It is capable of recording information over a more extended period without much effort. It is noticed that RNN consists of iterative modules formed in the chain for the neural network. In conventional RNN, a single tanh layer is used as a simplified structure for the iterative module. A similar chain-based system also exists in LSTM; however, the structure slightly differs concerning iterative modules. LSTM offers four layers of the interactive structure instead of a single layer. Fig. 4 highlights the conventional architecture of LSTM with multiple blocks of operation, i.e., neural network layer, pointwise operator, vector transfer, concatenation, and copy function.

The prime notion of LSTM is basically about a cell state that linearly runs over the entire chain letting the network simplify the flow of information. A specialized structure is used in LSTM called gates, which can add or eliminate information over the cell state. Usually, conventional gates in LSTM consist of pointwise multiplication operations and use a sigmoid neural network. From the image processing viewpoint, LSTM can be applied to images only after anticipated features have been extracted from them. At present, the applicability of LSTM towards improving the segmentation process of satellite images has been researched upon by various investigators. It has been noticed that although deep learning is preferred for classifying remote sensing images, simultaneous recognition of multiple objects and extracting their spatial relationship is yet a bigger problem in deep learning. This problem is addressed in Cui et al. [54], where a novel deep learning model is constructed by combining LSTM with fully CNN. The model has used natural language to define the spatial relationship of the remote objects using attention-based LSTM. The model has carried out semantic segmentation of multi-scale using CNN and U-Net entirely for object recognition of remotely sensed images. Nearly similar work is also carried out by Ghosh et al. [55]. A bidirectional LSTM is used along with UNet to extract temporal and spatial features of the satellite images for landcover mapping. The presented study has also used attention-based aggregation for all hidden representations of Spatio-temporal factors using a feedforward neural network followed by softmax normalization and spatial averaging. One of the significant advantages of this model is that it can perform segmentation even if different forms of atmospheric disturbances cover the images. Another recent work carried out by Kalinicheva et al. [56] has carried out a similar direction of study towards analyzing satellite images considering dynamic land cover changes. However, the authors have developed a better version of LSTM by introducing a unique unsupervised approach with an evolution graph. This technique makes use of image segmentation over the changed areas using the graph-based tree-merging method. Combined usage of LSTM and fully CNN has been reported in Sefrin et al. [57], which offers the benefit of using multi-temporal information. Apart from this, a better classification is presented due to adopted preprocessing techniques. Another work carried out by Zhu et al. [58] has developed a hybrid model where the semantic segmentation process is integrated with relearning of post classification. After feature extraction, the model performs an object-based voting system for controlling fluctuation in different classes. Table IV briefs of comparison of existing LSTM based approaches.

Observation: LSTM has been proven for better classification and segmentation performance for satellite images associated with landcover. However, the models don't emphasize its practicality as not many computing units can be used over a system with limited memory or channel capacity. This challenge remains unattended.

TABLE IV.   SUMMARY OF EXISTING TECHNIQUES ON LSTM

Authors	Problems	Segmentation Technique	Dataset	Advantages	Limitation
Cui et al. [54]	Object recognition	Semantic segmentation, multi-scale	RSICD	The spatial relation of remote sensing image	Higher processing time
Ghosh et al. [55]	Mapping landcover	Spatiotemporal segmentation, bidirectional LSTM	-Sentinel-based crop mapping. -Planet-based Cashew tree mapping	Effectively mitigate noise	Doesn't consider correlation among a large number of land cover
Kalinicheva et al. [56]	Change detection in landcover	Graph-based tree-merging segmentation	SPOT-5, Sentinel-2	Simplified clustering	Complexity increases with hierarchies of graphs
Sefrin et al. [57]	Change detection in landcover	LSTM, semantic segmentation	Federal State of Saxony	Better classification performance	Studies are not emphasized, unknown classes
Zhu et al. [58]	Change detection in landcover	LSTM, semantic segmentation	QuickBird, WorldView-2	Higher accuracy of classification	Time consumption is still more

D. Staked Auto Encoders

An autoencoder is an unsupervised learning structure characterized by input, hidden, and output layers, while the training operation in autoencoders consists of encoding and decoding. The encoder carries out the mapping of the input data into hidden representation, while the decoder carries out the reconstruction of input data from the hidden representation. The dependable parameters for the encoding process are encoding function, weight matrix, and bias vector. In contrast, dependable parameters for the decoding process are still the same, i.e., decoding function, weight matrix of decoder, and bias vector. To control the reconstruction error, an objective function exists for optimizing it considering the loss function. So, stacked autoencoders are basically about stacking a specific number of n autoencoders into the same n number of hidden layers using a supervised learning approach, as shown in Fig. 5.

Further supervised learning scheme is used for fine-tuning it. At present, there is the evolution of specific schemes that use autoencoders to analyze satellite images. In most existing approaches, the first line of action is to train the initial auto-encoders to extract the trained feature vector. The second step is to use that feature vector for the next layer as an input, and it's iterated till the completion of training. Finally, after all the hidden layers are trained, cost minimization is carried out, followed by updating weights. Existing approaches have reported the use of stacked auto encoders for change detection in landcover. The work carried out by Jing et al. [59] has used stacked auto encoders where multi-scale image segmentation is deployed over temporal images followed by the adoption of CNN to obtain a change map. A stacked autoencoder is used for classification. Usage of a similar principle was also reported in the work of Protopapadakis et al. [60] to evaluate targets over massive unlabeled data. Further denoising auto encoders have been reported in the work of Zhang et al. [61], where a spanning tree has been used for segmentation. The model can extract texture, spatial, and spectral features for all the identified objects, contributing to higher accuracy. Table V summarizes the existing contribution of stacked autoencoders towards segmentation process.
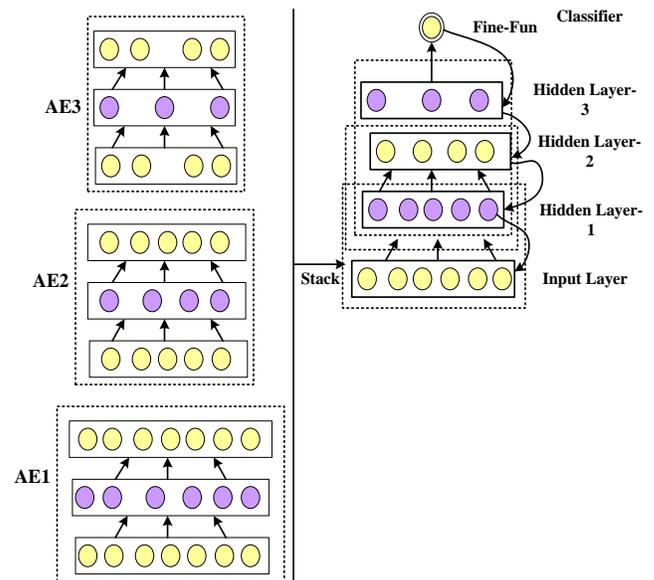

Fig. 5.   Conventional Architecture of Stacked Autoencoders.

TABLE V.   SUMMARY OF EXISTING TECHNIQUES ON STACKED AUTOENCODERS

Authors	Problems	Segmentation Technique	Dataset	Advantages	Limitation
Jin et al. [59]	Change detection	Multi-scale segmentation	ImageNet	Effective feature extraction	Induce computational complexity
Protopapadakis et al. [60]	Building extraction	Semantic segmentation	Vaihingen city in Germany	Redundancy reduction, Higher accuracy	Increased number of stacks
Zhang et al. [61]	Object classification	Spanning tree	UAV image of Anhui Province, China SPOT5	Higher accuracy	No assessment of computational complexity.

Observation: Deployment of stacked autoencoders is relatively a new scheme used for the segmentation of satellite images. Hence, there are few implementation studies towards stacked auto encoders to improve segmentation performance. Unfortunately, all the analyses using this approach don't find the appropriate number of stacks sufficient for encoding. The computational complexity raised towards maintaining the accumulation of information is not yet resolved in the existing system.

IV. ADOPTION OF DATASET

There are various types of a dataset of satellite images/remote sensing images used for analysis for landcover analysis. One of the essential datasets is the QuickBird dataset [62], which provides massive images of 0.60m Ground Sample Distance resolution and consists of four further multispectral bands with very high resolution ranging from 0.30-0.60m. The images were captured from 2001 to 2015 with an orbit height of 450 km. The following important dataset is Google Earth Satellite Imageries [63] which includes a raster dataset of satellites easier to be processed by any scripting environment. This is a massive dataset of satellite imageries consisting of a different crop type digital map, vegetation map, oil map, terrestrial field plots, elevation, human population, forest, water cover, etc. It also consists of Landsat satellite images with 30 meters resolution considered highly updated thermal and multispectral data [64]. MODIS [65] and Sentinel [66] is another dataset developed in collaboration with the Google dataset itself. MODIS dataset consists of satellite images ranging from 250m-1000m of snow cover, surface temperature, surface reflectance, leaf area index, and thermal anomalies, usually retrieved on 16days duration from Aqua and Terra spacecraft.

Further, the Sentinel dataset is a part of the European Space Agency consists of optimal high-resolution images (from Sentinel 1A/1B), land-ocean-climate images (Sentinel-3), and air quality images (from Sentinel-5P). They are frequently used in current research to analyze climatic change, emergency management, atmospheric monitoring, Marine monitoring, land monitoring. SPOT-5 dataset is another contribution for the European Space Agency [67], where the images were collected between 2002 and 2015 with an orbit height of 832 km and an orbit duration of 101 min. Similar organization of European Space Agency also offers RADARSAT-2 dataset whose resolution ranges from 9.0-160m [68]. This is the most updated dataset captured between 2008 and 2021, with both medium and very high resolution of wavelength between 5.2-7.7 cm. Inria Aerial image dataset consists of labeled remotely sensed images with 810 square kilometers [69]. With a spatial resolution of 0.3m, this dataset has color imageries with ground truth data and two semantics classes. This dataset consists of alpine towers, densely populated areas, and irregular urban settlements. A sample dataset for Inria is shown in Fig. 6, which exhibits sample Chicago landcover (Fig. 6(a)) and its reference as ground truth (Fig. 6(b)). The presence of reference/ground truth image assists in evaluating the correctness of analysis models, and hence this dataset is widely adopted.
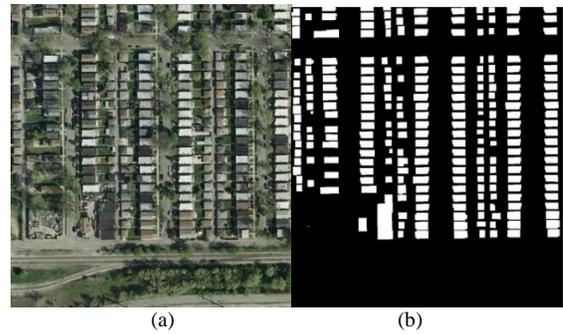

| (a) | (b) |

Fig. 6. Sample INRIA Dataset.

Existing studies have also been carried out considering the MNIST dataset, a benchmarked dataset for satellite images as a part of Kaggle [70]. This labeled dataset is managed in the form of images and CSV files. Kaggle dataset also consists of a DSTL dataset consisting of image identity with class type over its labeled area [71]. The dataset is maintained in 3/16-band satellite images with a resolution range of 0.31-7.5 m. The next dataset found in the current study is the ISPRS dataset, consisting of an indoor scene, old buildings, aerial images of specific locations, and satellite images of different parts of the earth [72]. Another dataset adopted in the current study is remote Sensing Image Captioning Dataset (RSICD) collected from the different applications of Google with all images with 224x224 pixels and 10921 remote sensing images [73]. Apart from the above-mentioned standard dataset, existing literature has also reported usage of another dataset too viz. Planetscope satellite images [74], GIS data [75], Pleiades VHSR, Corine Land Cover [76], OpenStreetMap [77], Cropland Data Layer [78], WorldView-2 [79], ImageNet [80]. It should be noted that all the dataset has the different characteristic of data of satellite images.

Fig. 7 showcases the adoption of different datasets towards the study of analyzing satellite images. The graphical trends shown in the above figure are obtained from filtering relevant research papers published between 2011 and 2021. It showcases that MODIS, Sentinel, and QuickBird are the most frequently adopted dataset by the researchers. Table VI highlights the comparative characteristic of different satellite image dataset with respect to spatial resolution. It should be noted that different dataset has different types of characteristic which is mainly based on the process of acquisition of signal. Hence, the proposed study considers highlighting spatial resolution of the images being captured to be mentioned in Table VI.
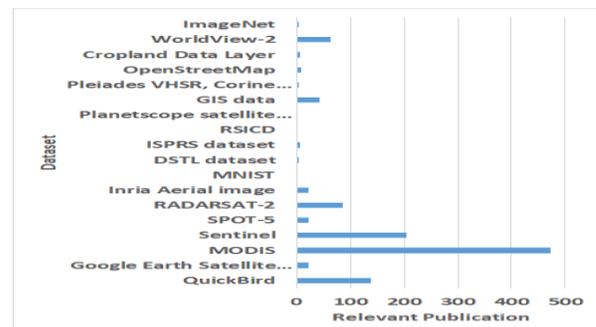

Fig. 7. Adoption of Dataset in Research Work.

TABLE VI.   CHARACTERISTICS OF SATELLITE IMAGE DATASET

Satellite Image dataset	Resolution
QuickBird dataset [62]	0.30-0.60 m
Google Earth Satellite Imageries [63]	30 m
MODIS [65]	250-1000m
Sentinel [66]	10-60 m
SPOT-5	10 m-600 KM
RADARSAT-2	10-100m
Inria Aerial image dataset [69]	0.3 m
MNIST dataset [70]	0.3-0.7 m per pixel
DSTL dataset [71]	0.31-7.5 m
Planetscope satellite images [74]	3.75m-50 cm
GIS data [75]	125 m
Pleiades VHSR, Corine Land Cover [76]	20 m
OpenStreetMap[77]	100 dpi
Cropland Data Layer [78]	30-56m
WorldView-2 [79]	0.46 m
ImageNet [80]	64x64 pixels

V.   RESEARCH TRENDS

At present, various Deep Learning (DL) approaches are being implemented towards analyzing satellite images / remotely sensed images. Last five years, data from IEEE Xplore digital archives have been studied to arrive at the inevitable conclusion of research trends.

From Fig. 8, it can be seen that the number of adoption of deep learning has been spontaneously increasing in the last five years. There are more probabilities towards the continuation of similar trends in upcoming years.

From Fig. 9, it is noticed that CNN is the dominant DL approach compared to other DL variants, i.e., RNN, LSTM, Stacked Autoencoders, Fully Convolution Network (FCN), and Deep Belief Network (DBN). Also, there is increasing adoption of LSTM and FCN approaches; however, they are significantly less in numbers. Hence, chances are more for CNN to be dominant in coming years too.

From Fig. 10, it is noticed that urban-based application is more investigated, followed by a water-based application using DL methods. The urban-based application will include identifying land covers, mainly exhibiting that it will focus on the research area.
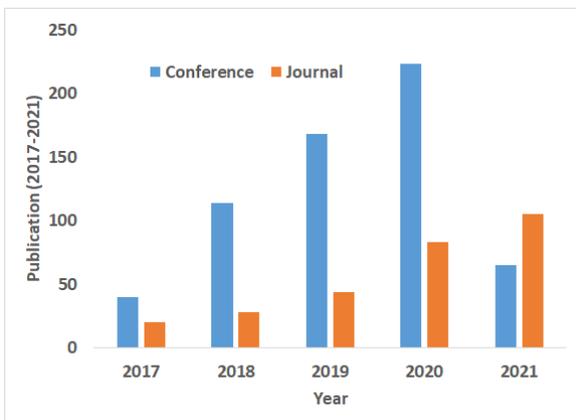

Fig. 8.   Year-Wise Trends of DL-Approaches on Satellite Images.
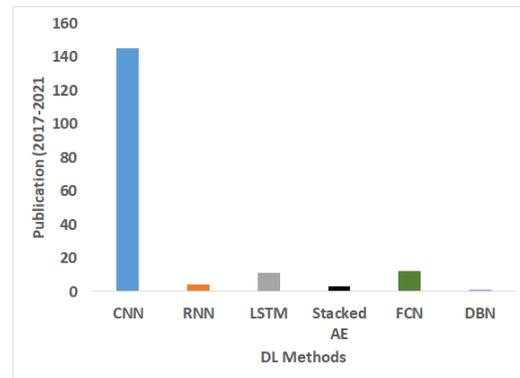

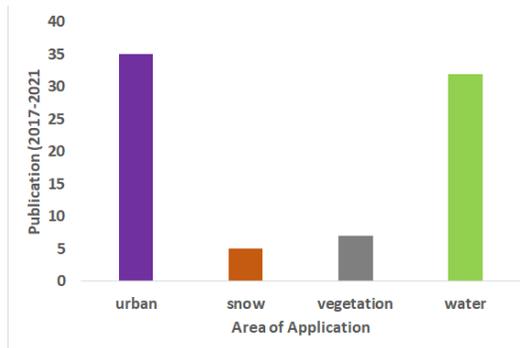Fig. 9.   Trends towards Taxonomies of DL-Approaches on Satellite Images.

Fig. 10. Trends of Application Area Year-Wise Trends of DL-Approaches on Satellite Images.

This outcome eventually shows a higher scope of continued research work using DL methods towards satellite images in more progressive order. These findings of the research trend are one of the essential contributions of this manuscript.

VI. REVIEW CONTRIBUTION

From the prior section, it can be seen that various degree of research work is being carried out towards the segmentation of satellite images. It should be noted that not all deep learning mechanisms have implemented segmentation towards the input image. This paper has discussed only the research work where segmentation has been applied. The *scope* of this paper is i) the paper considers discussion of approaches published in last five years, ii) the paper emphasizes assessing the impact of different deep learning models towards segmentation. After reviewing different taxonomies of deep learning methods towards segmentation, different points of research gap are concluded that are briefed as follows:

A. Research Gap

The essential research gap explored after reviewing existing approaches are as follows:

- Unattended Computational Problems in Deep Learning: Despite the frequently adopted technique, the studies using deep learning have witnessed an evident trade-off between achieving simplification in learning (positive point) and poor computational performance (negative point). Almost all the CNN

approaches studied in this paper are witnessed with lower performance scalability or increased computational complexities. Unless lower computational complexities characterize the framework; it is eventually not feasible to prove its application over real-life resourced defined computing devices. This fact is also witnessed for almost all deep learning techniques.

- Complexities of Satellite Images: Irrespective of any dataset of satellite imageries, it is now known that such images are accompanied by various challenges, from analyzing the raw sensed image to obtaining higher accuracy by extracting essential features. Although, deep learning methods can only be applied if the preprocessing is effectively carried out, which is missing in the most implementations. Hence, it is essential to preprocess the satellite image before feature extraction because it consists of a large amount of information. The deep learning method can automate the process for better classification; however, it will still depend on practical preprocessing input images before training.

- Biased Focused on Segmentation Approach: A closer look into existing approaches shows a majority of semantic segmentation methods used over satellite images. Such a technique labels each pixel concerning the associated class of satellite images, further subjected to dense prediction. One of the pitfalls of such an approach is that instances of the same classes are not segmented, potentially affecting landcover applications. Apart from this, various methods reported in this paper using deep learning don't include object irregularity, illumination factor, poor contrast, noise, etc. The exclusion of these points will eventually affect the accuracy of classification using deep learning.

- Less Emphasis towards Computational Performance: Although there are more comprehensive deep learning approaches to analyze satellite images, it is essential to identify the proper model. Different models have different performance patterns, and there is no full-proof deep learning model generalized towards analyzing satellite images. Inappropriate selection of deep learning model towards segmentation is a complex task as segmentation operation to be applied wholly subjected and application-oriented. This eventually leads to computational complexities, evident from limiting points found in existing studies discussed in this paper. Without addressing computational performance, it is quite questionable to understand their applicability.

B. Discussion

A closer look into the deep learning approaches shows many methods for analyzing satellite images, but not much work is reported towards segmentation approaches. One interesting observation noted in all segmentation-based approaches is that almost every research work has adopted different datasets with different properties and implemented them. Although Sentinel, MODIS, and QuickBird are frequently adopted datasets, they differ in addressing different problems. A better form of progressive work towards a deep learning approach is required considering a large scale similar dataset first, which can be compared with other existing datasets later. But this is not the case with existing methods. Another observation is that existing approaches are not witnessed much with scalable and consistent performance. LSTM, which is considered a better variant of RNN, is seen with time consumption and increased complexities in many cases. This is because architecture towards an extensive memory system is theoretically proven, and its performance doesn't scale up when exposed to a complex and challenging environment. Apart from this, CNN has implemented either individually or in combination with other training models. The standalone implementation of CNN towards the segmentation process is witnessed with various challenges that are not attended. The first challenge in standalone CNN implementation is associated with the drastically slower operation that consumes enough training time for the larger size of the satellite image. None of the existing studies has reported overcoming the dependencies of a resourceful graphical processing unit for supporting extensive layers of training in CNN. Apart from this, after the object is identified from satellite images using the CNN technique, it must be encoded for better accuracy. However, it is not feasible for CNN to encode pixel position and identify changes in object orientation. This will potentially affect the accuracy. Hence, there is a need to address the inherent issues of deep learning and attend to other matters.

VII. CONCLUSION

Remote sensing and satellite images have become essential applications for change detection and classification. With an increasing rise of deep learning-based approaches for analyzing satellite images, the idea of this paper is to review the existing approaches associated with segmentation. The novelty points of this paper are i) existing review papers has reviewed semantic segmentation, segmentation with specific application, whereas the proposed review paper has explicitly discussed all the recent segmentation approaches using deep learning with various application grounds offering more technical insights, ii) proposed review contributes to updated research trends to understand most dominant deep learning-based technique suitable for segmentation, iii) each study has been discussed concerning good points and limiting factors for offering more granularity in review findings, iv) the proposed review work contributes towards research gap followed by a discussion to know less spoken information about strength and weakness of existing schemes.

Future work will be further carried out to address the research gap identified in this paper. A computational framework can be designed to consider various artifact inclusion combined that has never been done before. This will offer a scope to introduce a novel preprocessing approach that can potentially contribute to the enriched feature extraction process. Further, a novel deep learning model can be framed, emphasizing reduced training demands, reduced processing time, and optimal computational performance.

REFERENCES

[1] I.L.Turnera, M. D. Harleya, R. Almarb, E.W.J.Bergsma, "Satellite optical imagery in Coastal Engineering", Elsevier-ScienceDirect, Coastal Engineering, Vol.167, August 2021.

[2] H. Zhang, H. Liu, Y. Wang, "A new scheme for urban impervious surface classification from SAR images", ISPRS Journal of Photogrammetry and Remote Sensing, vol.139, 2018.

[3] K. Maurya, S. Mahajan, & N. Chaube, "Remote sensing techniques: mapping and monitoring of mangrove ecosystem—a review" Springer-Open Access-Journal for Complex and Intelligent System, 2021.

[4] R. Pazur, B. Price, & P. M. Atkinson, "Fine temporal resolution satellite sensors with global coverage: an opportunity for landscape ecologists", Springer-Open Access, Research Article for Landscape Ecology, vol.36, pp.2199-2213, 2021

[5] V. Yu. Ignatiev, I. A. Matveev, A. B. Murynin, A. A. Usmanova & V. I. Tsurkov, "Increasing the Spatial Resolution of Panchromatic Satellite Images Based on Generative Neural Networks", Springer, Journal of Computer and Systems Sciences International, Vol.60, pp.239-247, 2021.

[6] C. Vallentin, K. Harfenmeister, S. Itzerott, B. Kleinschmit, C. Conrad & D. Spengler, "Suitability of satellite remote sensing data for yield estimation in northeast Germany", Springer-Open Access Journal for Precision Agriculture, 2021.

[7] P. Gudžius, O. Kurasova, V. Darulis & E. Filatovas, "Deep learning-based object recognition in multispectral satellite imagery for real-time applications", Springer-Machine Vision and Application, vol.32, Article No.98, 2021.

[8] S. Srivastava, A. V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni & V. Pattabiraman, "Comparative analysis of deep learning image detection algorithms" Springer-Open Access Journal of Big Data, Vol.8, Article No.66, 2021.

[9] J. Nalepa, M. Myller, M. Cwiek, L. Zak, T. Lakota, L. Tulczyjew, and M. Kawulo, "Towards On-Board Hyperspectral Satellite Image Segmentation: Understanding Robustness of Deep Learning through Simulating Acquisition Conditions", MDPI Journal of Remote Sensing, vol.13, Iss.1532, 2021.

[10] El-Habib Bensikaddour, Y. Bentoutou, N. Taleb, "Embedded implementation of multispectral satellite image encryption using a chaos-based block cipher", ScienceDirect-Journal of King Saud University - Computer and Information Sciences, Vol.32, Iss.1, pp.50-56, 2020.

[11] A. Hagag, I. Omara, S. Chaib, X. Fan, F. E. Abd El-Samie, "Distributed Coding and Transmission Scheme for Wireless Communication of Satellite Images", Springer- International Conference on Computer and Information Science, Computer and Information Science, pp 29-42, 2017.

[12] D. Yan, H. Yi, D. He, K. Guan, B. Ai, Z. Zhong, J. Kim, H.Chung, "Channel Characterization for Satellite Link and Terrestrial Link of Vehicular Communication in the mmWave Band", IEEE-Open Access, vol.7, 2019.

[13] A. M. Crisan, A. Martian, R. Cacoveanu, D. Coltuc, "Angle-of-Arrival Estimation in Formation Flying Satellites: Concept and Demonstration", IEEE-Access, vol.7, 2019.

[14] I. F, Akyildiz, A. Kak, S.Nie, "6G and Beyond: The Future of Wireless Communications Systems", IEEE Access, vol.8, 2020.

[15] J. Thrane, D. Zibar, H. L. Christiansen, "Model-Aided Deep Learning Method for Path Loss Prediction in Mobile Communication Systems at 2.6 GHz", IEEE Access, vol.8, 2020.

[16] M. Ju, C. Ding, Y.J Guo, D. Zhang, "Remote Sensing Image Haze Removal Using Gamma-Correction-Based Dehazing Model", IEEE-Access, vol.7, 2019.

[17] N. A. Golilarz, H. Gao, S. Pirasteh, M. Yazdi, J. Zhou, and Y. Fu, "Satellite Multispectral and Hyperspectral Image De-Noising with Enhanced Adaptive Generalized Gaussian Distribution Threshold in the Wavelet Domain", MDPI-Remote Sensing, vol.13, Iss.101, 2021.

[18] B. Dhruv, N. Mittal and M. Modi, "Analysis of different filters for noise reduction in images," 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE), pp. 410-415, doi: 10.1109/RDCAPE.2017.8358306, 2017.

[19] L. Fan, F. Zhang, H. Fan & C. Zhang, "Brief review of image denoising techniques", Springer-Open Access-Visual Computing for Industry, Biomedicine, and Art, Vol.2, Article No.7, 2019.

[20] S. Mahajan & B. Fataniya, "Cloud detection methodologies: variants and development—a review", Springer-Open Access, Survey and State of the Art, vol.6, pp.251-261, 2020.

[21] J-R Lee, C-Y Chung. & M-L Ou, "Fog detection using geostationary satellite data: Temporally continuous algorithm", Springer-Asia-Pacific Journal of Atmospheric Sciences, vol.47, pp.113-122, 2011.

[22] G. P. Obi ReddyS. K. Singh, "Geospatial Technologies in Land Resources Mapping, Monitoring and Management", Springer-Geotechnologies and the Environment book series, vol.21, 2018.

[23] https://cimss.ssec.wisc.edu/satmet/modules/5_sat_images/si-1.html. Retrieved on 17-November, 2021.

[24] L. Zhu, J. Suomalainen, J. Liu, J. Hyyppa, H. Kaartinen and H. Haggren, "A Review: Remote Sensing Sensors", IEEE-Open access peer-reviewed chapter, IntechOpen, 2017.

[25] N. Verde, G. Mallinis, M.T. Strati, C. Georgiadis, and P. Patias, "Assessment of Radiometric Resolution Impact on Remote Sensing Data Classification Accuracy", MDPI-Remote Sensing, vol.10, Iss.8, 2018.

[26] Unsalan, Cem, Boyer, Kim L. "Multispectral Satellite Image Understanding", Springer-Advances in Computer Vision and Pattern Recognition, 2011.

[27] Hemanth, D. Jude, "Artificial Intelligence Techniques for Satellite Image Analysis", Springer-Remote Sensing and Digital Image Processing, 2020.

[28] A. Asokan, J. Anitha, M. Ciobanu, A. Gabor, A. Naaji, and D. Jude Hemanth, "Image Processing Techniques for Analysis of Satellite Images for Historical Maps Classification—An Overview", MDPI-Applied Science, vol.10, Iss.12, 2020.

[29] Alshehhi, R.; Marpu, P.R. Hierarchical graph-based segmentation for extracting road networks from high-resolution satellite images. ISPRS J. Photogramm. Remote Sens. 126, 245–260, 2017.

[30] S. Pare, A.K. Bhandari, A. Kumar, G.K. Singh, "A new technique for multilevel color image thresholding based on modified fuzzy entropy and Lévy flight firefly algorithm". ScienceDirect-Computers & Electrical Engineering, vol.70, pp.476–495, 2018.

[31] D. Marmanis, K. Schindler, J.D. Wegner, S. Galliani, M. Datcu, U. Stilla, "Classification with an edge: Improving semantic image segmentation with boundary detection", ISPRS Journal of Photogramm. Remote Sensing, 135, 158–172, 2018.

[32] I. Grinias, C. Panagiotakis, G. Tziritas, "MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images", ISPRS Journal of Photogramm. Remote Sens, 122, 145–166, 2016.

[33] S. Suresh, S. Lal, "An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions". Expert System Application, 58, 184–209, 2016.

[34] A. Sayer, Y. Govaerts, P. Kolmonen, A. Lipponen, M. Luffarelli, T. Mielonen, "A review and framework for the evaluation of pixel- level uncertainty estimates in satellite aerosol remote sensing". *Atmos. Meas. Tech.* 13, 373–404, 2020. doi:10.5194/amt-13-373-2020.

[35] L. Liang, and L. Di Girolamo, L., "A global analysis on the view-angle dependence of plane-parallel oceanic liquid water cloud optical thickness using data synergy from MISR and MODIS". *Journal of Geophysical Research Atmospheresm* vol.118, pp.2389–2403, 2013 doi:10.1029/2012JD018201.

[36] A. Shelestov et al., "Cloud Approach to Automated Crop Classification Using Sentinel-1 Imagery," in IEEE Transactions on Big Data, vol. 6, no. 3, pp. 572-582, 1 Sept. 2020, doi: 10.1109/TBDATA.2019.2940237.

[37] D.P. Kingma,; J.A. Ba, "A method for stochastic optimization". arXiv 2014, arXiv:1412.6980.

[38] S. Ioffe, S. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", arXiv 2015, arXiv:1502.03167.

[39] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A simple way to prevent neural networksfrom overfitting". Journal of Machine Learning Research, vol. 15, pp. 1929–1958, 2014.

[40] K. He, X. Zhang, S. Ren, J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification". In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1026–1034.

[41] M. Wurma, T. Stark, X. X. Zhu, M. Weigand, H. Taubenböck, "Semantic segmentation of slums in satellite images using transfer learningon fully convolutional neural networks", Elsevier-ISPRS Journal of Photogrammetry and Remote Sensing, vol.150, pp.59-69, 2019.

[42] J. Zhang, Q. Hu, J. Li and M. Ai, "Learning From GPS Trajectories of Floating Car for CNN-Based Urban Road Extraction With High-Resolution Satellite Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 3, pp. 1836-1847, March 2021, doi: 10.1109/TGRS.2020.3003425.

[43] Q. Li, Y. Shi, X. Huang and X. X. Zhu, "Building Footprint Generation by Integrating Convolution Neural Network With Feature Pairwise Conditional Random Field (FPCRF)," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 11, pp. 7502-7519, Nov. 2020, doi: 10.1109/TGRS.2020.2973720.

[44] I. V. Saetchnikov, E. A. Tcherniavskaia and V. V. Skakun, "Object Detection for Unmanned Aerial Vehicle Camera via Convolutional Neural Networks," in IEEE Journal on Miniaturization for Air and Space Systems, vol. 2, no. 2, pp. 98-103, June 2021, doi: 10.1109/JMASS.2020.3040976.

[45] N. Wang, B. Li, X. Wei, Y. Wang and H. Yan, "Ship Detection in Spaceborne Infrared Image Based on Lightweight CNN and Multisource Feature Cascade Decision," in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 5, pp. 4324-4339, May 2021, doi: 10.1109/TGRS.2020.3008993.

[46] Y. Jiang, "Research on road extraction of remote sensing image based on convolutional neural network", Eurasip Journal on Image and Video processing, 2019.

[47] C. Persello and A. Stein, "Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images," in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 12, pp. 2325-2329, Dec. 2017, doi: 10.1109/LGRS.2017.2763738.

[48] L. Wang, K. A. Scott, L. Xu and D. A. Clausi, "Sea Ice Concentration Estimation During Melt From Dual-Pol SAR Scenes Using Deep Convolutional Neural Networks: A Case Study," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4524-4533, Aug. 2016, doi: 10.1109/TGRS.2016.2543660.

[49] M. C. Taberner, F. J. García-Haro, B. Martínez, E. I. Verdiguier, C. Atzberger, G. C. Valls, & M. A. Gilabert, "Understanding deep learning in land use classification based on Sentinel-2 time series", Open-Access-Scientific Reports, vol.10, 2020.

[50] M. O. Turkoglu, S. D'Aronco, J. D. Wegner, and K. Schindler, "Gating Revisited: Deep Multi-layer RNNs That Can Be Trained", arXiv:1911.11033v4 [cs.CV] 6 Mar 2021.

[51] D. Ienco, R. Gaetano, C. Dupaquier and P. Maurel, "Land Cover Classification via Multi-temporal Spatial Data by Recurrent Neural Networks", arXiv:1704.04055v1 [cs.CV] 13 Apr 2017.

[52] E. Maggiori, G. Charpiat, Y. Tarabalka and P. Alliez, "Recurrent Neural Networks to Correct Satellite Image Classification Maps," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 4962-4971, Sept. 2017, doi: 10.1109/TGRS.2017.2697453.

[53] Z. Sun, L.Di, "Using long short-term memory recurrent neural network in land cover classification on Landsat and Cropland data layer time series", Taylor and Francis, International Journal of Remote Sensing, vol.40, Iss.2, 2019.

[54] W. Cui, F. Wang, X. He, D. Zhang, X. Xu, M. Yao, Z. Wangand J. Huang, "Multi-Scale Semantic Segmentation and SpatialRelationship Recognition of Remote Sensing Images Based on an Attention Model", MDPI-Remote Sensing, vol.11, 2019.

[55] R. Ghosh, P. Ravirathinam, X. Jia, C. Lin, Z. Jin, V. Kumar, "Attention-augmented Spatio-Temporal Segmentation for Land Cover Mapping", arXiv:2105.02963v2 [cs.CV] 14 Sep 2021.

[56] E. Kalinicheva, D. Ienco, J. Sublime, M. Trocan, "Unsupervised Change Detection Analysis in Satellite Image Time Series using Deep Learning Combined with Graph-Based Approaches". IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE, 2020, 13, pp.1450-1466.

[57] O. Sefrin, F. M. Riese, and S. Keller, "Deep Learning for Land Cover Change Detection", MDPI-Remote Sensing, Vol.13, 2021.

[58] Y. Zhu, C. Geiß, E. So and Y. Jin, "Multitemporal Relearning With Convolutional LSTM Models for Land Use Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 3251-3265, 2021, doi: 10.1109/JSTARS.2021.3055784.

[59] R. Jing, Z. Gong and H. Guan, "Land Cover Change Detection With VHR Satellite Imagery Based on Multi-Scale SLIC-CNN and SCAE Features," in IEEE Access, vol. 8, pp. 228070-228087, 2020, doi: 10.1109/ACCESS.2020.3045740.

[60] E. Protopapadakis, A. Doulamis, N. Doulamis, and E. Maltezos, "Stacked Autoencoders Driven by Semi-Supervised Learning for Building Extraction from near Infrared Remote Sensing Imagery", MDPI-Remote Sensing, vol.13, 2021.

[61] X. Zhang, G. Chen, W. Wang, Q. Wang and F. Dai, "Object-Based Land-Cover Supervised Classification for Very-High-Resolution UAV Images Using Stacked Denoising Autoencoders," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 7, pp. 3373-3385, July 2017, doi: 10.1109/JSTARS.2017.2672736.

[62] https://earth.esa.int/eogateway/catalog/quickbird-full-archive, Retrieved on 19 Sep-2021.

[63] https://developers.google.com/earth-engine/datasets/catalog, Retrieved on 19 Sep-2021.

[64] https://www.usgs.gov/core-science-systems/nli/landsat/landsat-data-access?qt-science_support_page_related_con=0#qt-science_support_page_related_con, Retrieved on 19 Sep-2021.

[65] https://modis.gsfc.nasa.gov/data/, Retrieved on 19 Sep-2021.

[66] https://sentinel.esa.int/web/sentinel/sentinel-data-access, Retrieved on 19 Sep-2021.

[67] https://earth.esa.int/eogateway/missions/spot-5, Retrieved on 19 Sep-2021.

[68] https://earth.esa.int/eogateway/catalog/radarsat-2-esa-archive, Retrieved on 19 Sep-2021.

[69] https://project.inria.fr/aerialimagelabeling/, Retrieved on 19 Sep-2021.

[70] https://www.kaggle.com/datamunge/overheadmnist, Retrieved on 19 Sep-2021.

[71] https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection, Retrieved on 19 Sep-2021.

[72] https://www.isprs.org/data/, Retrieved on 19 Sep-2021.

[73] https://github.com/201528014227051/RSICD_optimal, Retrieved on 19 Sep-2021.

[74] https://www.planet.com/products/planet-imagery/, Retrieved on 19 Sep-2021.

[75] https://www.usgs.gov/products/data-and-tools/gis-data, Retrieved on 19 Sep-2021.

[76] https://land.copernicus.eu/pan-european/corine-land-cover, Retrieved on 19 Sep-2021.

[77] https://www.openstreetmap.org/#map=5/51.495/-0.088, Retrieved on 19 Sep-2021.

[78] https://catalog.data.gov/dataset/cropscape-cropland-data-layer, Retrieved on 19 Sep-2021.

[79] https://www.satimagingcorp.com/satellite-sensors/worldview-2/, Retrieved on 19 Sep-2021.

[80] https://www.image-net.org/update-mar-11-2021.php, Retrieved on 19 Sep-2021.