Deep Learning in Remote Sensing Paper Summaries

Note: References here do not match reference numbers in the paper. The references are provided at the end.

Online content provided with paper “A Comprehensive Survey of Deep Learning in Remote Sensing: Theories, Tools and Challenges for the Community” by Ball, Anderson and Chan

Legend: Spec = Spectral, Spat = Spatial, Temp = Temporal/Multi-temporal, HS/MS = HSI/Multispectral, US = Ultraspectral, AP/AD = Aerial Photo/Aerial Data (includes satellite data), RGB = RGB color imagery.

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Alam et al. [1]	HSI Image Segmentation	X	X				X	X					CNN + Conditional Random field. CNN performs superpixel-level labeling.	HSI	Indian Pines
Alcantarilla et al. [2]	Street-view change detection		X						X	X			A multi-sensor fusion SLAM and fast dense 3D reconstruction pipeline gives coarsely registered image pairs to a deep deconvolutional network for pixel-wise change detection. An urban change detection dataset (order of magnitude larger than existing datasets) and contains challenging changes due to seasonal and lighting variations is also provided.	RGB	VL-CMU-CD
Alexandre [3]	3D Object recognition				X					X			Transfer learning.	RGBD	Multi-View RGBD
Alidoost et al. [4]	Urban building classification		X		X	X		X			X		Building roof classification using CNN for RGB from orthophoto data fused with roof DSM.	LiDAR, Aerial orthophotos	Custom Aerial imagery and LiDAR Stuttgart, Germany

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Audebert et al. [5]	Segmentation prior to deep learning processing then SVM		X										Used Deep learner to create features for SVM classifier.	IR-R-G orthoimages	ISPRS
Basaeed et al. [6]	Image segmentation	X	X				X	X					A boosted committee of CNNs coupled with inter-band and intra-band fusion. Region boundary probability maps are derived from multispectral bands.	Multispectral	Prague Texture Segmentation Benchmark
Basaeed et al. [7]	Multispectral segmentation	X	X				X	X					A supervised hierarchical segmentation of remote-sensing images using a committee of multi-scale convolutional neural networks.	Multispectral	Prague Texture Segmentation Benchmark
Basu et al. [8]	Deep belief network satellite image classification	X	X				X	X		X			Features extracted then classified by a deep belief network. SAT-4 and SAT-6 images are also contributions.	Multispectral satellite images (R,G,B,NIR)	SAT-4, SAT-6
Bazi et al. [9]	HSI classification	X	X				X	X					An evolutionary optimization algorithm uses cross-validation accuracy to determine the extreme learning machine parameters.	HSI images	Indian Pines, Kennedy Space Center, Washington DC Mall, Pavia University
Becker et al. [10]	Explosive hazard detection in ground penetrating radar		X									X	Use deep belief network to discriminate false alarms after CFAR prescreener.	GPR data	Custom
Bentes et al. [11]	Oceanographic Target Classification					X							SAR amplitude imagery processed via CFAR. Block features are processed via CNN.	SAR	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Bergado et al. [12]	Classifying very high resolution satellite imagery	X	X				X	X		X			Spatial multiresolutinal features are learned by a CNN. A sensitivity analysis of the network hyper-parameters is also provided.	Orthophoto with NIR + two optical bands.	Custom
Besaw [13]	Buried explosive hazard detection in GPR		X									X	The artificial neural network is used to discriminate clutter from buried explosive hazards. A custom ConvNet with 2X2 masks and dropout was utilized.	GPR	Custom
Besaw et al. [14]	Buried explosive hazard detection in GPR		X						X	X		X	A custom deep belief network detects buried explosive hazards.	GPR	Custom
Brust et al. [15]	Road detection and urban scene understanding		X						X	X			Convolutional path networks allow for scene understanding and pixel labeling.	RGB	KITTI
Cadena et al. [16]	Estimating depth from a single image		X						X	X			Fusion of RGB, depth images and semantic labels. Multi-modal Autoencoders can solve the depth estimation and the semantic segmentation problems simultaneously.	RGB	KITTI
Cao et al. [17]	HSI classification	X	X				X	X					Graph-based spatial fusion and Convolutional Neural Network (CNN).	AVIRIS	Indian Pines
Cao et al. [18]	HSI classification	X	X				X	X					Graph-based spatial/spectral fusion + CNN provides pixel-level classification.	AVIRIS	Indian Pines

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Castelluccio et al. [19]	Land Use Classification	X	X				X	X					Used CaffeNet and GoogleNet. Convolutions at different sizes allow processing across scales. Compared to a wide variety of other solutions. Transfer learning utilized.	Aerial Optical Images and R-G-NIR false-color images	Trained on ImageNet, tested on UC Merced, Brazilian Coffee Bean
Chen et al. [20]	Aircraft detection in hi-res satellite imagery		X					X					Preprocessing using gradients, gray thresholding images followed by DBN	Hi-res satellite imagery (Google Earth)	Custom
Chen et al. [21]	HSI classification	X	X					X					Spectral classification by stacked autoencoders. Spatial information fusion of PCA, deep learning and logistic regression.	HSI	Kennedy Space Center, Pavia City Center
Chen et al. [22]	HSI classification	X	X				X	X					A Restricted Boltzmann Machine (RBM) and DBN perform spectral information-based classification. A combined spectral–spatial feature extraction are utilized. The framework is a hybrid of PCA, hierarchical learning-based FE, and logistic regression (LR).	HSI	Indian Pines, Pavia City Center
Chen et al. [23]	Autonomous Driving								X	X			Input image is mapped to a small number of key perception indicators that directly relate to the affordance of a road/traffic state for driving. This method provides a set of compact yet complete descriptions of the scene to enable a simple controller to drive autonomously.	RGB video	KITTI

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Chen et al. [24]	Predicting drought index			X									A short-term drought prediction model based on deep belief networks (DBNs) predicts the time series of different time-scale standardized precipitation index (SPI).	Custom artificial data, Custom hydrologic datasets	Custom
Chen et al. [25]	Thematic Classification	X	X		X		X	X			X		Novel 3D CNN to extract spectral-spatial features from the HSI data plus a deep 2D CNN extracts elevation features of LiDAR data. A fully connected Deep NN fuses the data.	HSI + LiDAR	Custom
Chen et al. [26]	SAR ATR		X			X						X	A single layer convolutional neural network is used to automatically learn features from SAR images. Instead of using the classical backpropagation algorithm, the convolution kernel is trained on randomly sampled image patches using unsupervised sparse auto-encoder. After convolution and pooling, an input SAR image is then transformed into a series of feature maps. These feature maps are then used to train a final softmax classifier.	SAR	MSTAR

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Chen et al. [27]	3D object detection for autonomous driving		X		X				X	X	X		Sparse 3D LiDAR data is encoded with a compact multi-view representation. Two subnetworks process the data: one for 3D object proposal generation and another for multi-view feature fusion. The proposal network generates 3D candidate boxes efficiently from the bird's eye view representation of 3D point cloud. A deep fusion scheme to combine region-wise features from multiple views and enable inter-actions between intermediate layers of different paths is developed.	RGB, LiDAR	KITTI
Chen et al. [28]	Object class detection		X		X				X	X	X		Stereo imagery is analyzed to place 3D bounding box proposals. The solution is found by minimizing an energy function encoding object size priors, ground plane as well as several depth informed features that reason about free space, point cloud densities and distance to the ground.	RGB, LiDAR	KITTI
Chen et al. [29]	Vehicle detection in satellite imagery		X					X					A hybrid DNN (HDNN), divides the maps of the last convolutional layer and the max pooling layer of DNN into multiple blocks of variable receptive field sizes or max-pooling field sizes, to enable the HDNN to extract variable-scale features.	Custom Google Earth imagery of San Francisco, CA	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Chen et al. [30]	Thematic classification	X	X		X		X	X			X		A 3D convolutional neural network (CNN) to extract the spectral-spatial features of HSI data and a deep 2D CNN extracts the elevation features of LiDAR data. A fully connected deep neural network to fuses the features and the classification results are produced by a logistic regression method.	Custom HSI and LiDAR dataset over Houston, TX	Custom
Cheng et al. [31]	Land Use Classification		X					X					A library of pretrained part detectors used for midlevel visual elements discovery called “partlets” is utilized. To address computational cost, coarse-to-fine shared intermediate representations, which are termed “sparselets,” are created using a single-hidden-layer autoencoder and a single-hidden-layer neural network with an L0-norm sparsity constraint.	Orthoimagery	LULC
Cheng et al. [32]	Scene classification	X	X				X	X					Pre-trained CNN models (AlexNet, VGGNet, and GoogleNet) are used as universal feature extractors and then fine-tuning on our scene classification dataset. Scene classification is carried out by using simple classifiers such as linear SVM.	HSI	UC Merced

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Cheng et al. [33]	Scene classification in VHR images	X	X				X	X		X			A rotation-invariant CNN (RICNN) model is trained by optimizing a new objective function via imposing a regularization constraint, which explicitly enforces the feature representations of the training samples before and after rotating to be mapped close to each other.	Pansharpened images	NWPU VHR-10
Chigorin et al. [34]	Road sign detection		X							X			Adaboost first stage and color suppression. Custom CNN.	RGB imagery	Russian Traffic Signs Dataset (RTSD)
Cireşan et al. [35]	Road sign detection and character recognition		X						X	X			Multi-column DNN based on feline visual cortex	Video converted to still imagery	GTSRB Traffic Signs
Couprie et al. [36]	Indoor semantic segmentation		X		X					X			A multiscale convolutional network to learn features directly from the images and the RGBD depth information	Pairs of RGB and depth imagery	NYU-v2
Cui et al. [37]	Low-resolution image upscaling		X							X			Cascade of autoencoders used to upscale imagery	RGB	Custom
Dahmane et al. [38]	Object detection	X	X				X	X					Deep learning applied to VHR Pleaides imagery to detect cars and trees.	MSI (Pleaides)	Custom
De et al. [39]	Urban classification from PolSAR					X						X	A Deep learner learns features and statistics (alpha and gamma parameters) are also used. A multi-layer perceptron is used as a classifier.	SAR Custom UAVSAR MLC L-Band dataset over San Francisco	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Del Frate et al. [40]	Classification sing Multi-Layer Perceptron (MLP)		X					X		X			The generalization capabilities of the Multi-Layer Perceptron (MLP) on the generalization capabilities of this type of algorithms with the purpose of using them as a tool for fully automatic classification of collections of satellite images, either at very high or at high-resolution.	HR and VHR imagery (LANDSAT and Quickbird)	Custom
Diao et al. [41]	Object recognition		X					X		X			A Deep Belief Network (DBN) has two stages, the unsupervised pre-training stage and the supervised fine-tuning stage. A stacked set of restricted Boltzmann machines (RBMs) is used to build a deep generative model. The RBM is trained in an unsupervised manner, and fine tuning is performed by using supervised learning on a supervised layer at the end of the RBM chain. Finally, the deep model generates good joint distribution of images and their labels.	VHR (Quickbird)	Custom
Ding et al. [42]	Detecting objects via transfer learning		X										Deep features and classifier parameters are obtained simultaneously. A weighted scheme to couple source and target output by assigning pseudo labels to target data, therefore knowledge can be transferred knowledge from source (i.e., MWIR) to target (i.e., LWIR).	LWIR, MWIR	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Dong et al. [43]	Super-resolution		X							X			A sparse coding SR method is shown to be a deep convolutional network. The SR maps low-resolution imagery to high-resolution imagery. All layers are jointly optimized. The method achieves high speed. Different network structures and parameter settings to achieve tradeoffs between performance and speed.	RGB	ILSVRC 2013 ImageNet
Du et al. [44]	SAR automatic target recognition		X			X						X	To overcome rotation and displacement problems with a traditional CNN, a displacement- and rotation-insensitive deep CNN is utilized.	SAR	MSTAR
Ducournau et al. [45]	Super-resolution for ocean remote sensing sea surface temperature analysis		X	X		X		X				X	The SRCNN (Super Resolution CNN) has a considerable gain in terms of PSNR compared to classical downscaling techniques. Data uses AVHRR, AATSR, SEVERI, AMSRE, TMI and buoy data.	Satellite infrared and microwave data, Buoy data (OSTIA SST Time Series)	Custom
Donahue et al. [46]	Can CNN training cross to other domains		X							X			Deep convolution features and DeCAF code.	RGB	SUN-397 large scale object recognition
Elawady [47]	underwater coral classification		X							X			Coral Classification, underwater color adjustment algorithms.	RGB	Moorea, Atlantic Deep Sea
Fang et al. [48]	Remote Sensing Image Classification		X										Uses CaffeNet fine-tuning.	RGB	UC Merced, and 19-class satellite scene

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Farias et al. [49]	Automated feature extraction in large thermonuclear fusion databases		X							X			Sparse autoencoders for feature reduction.	Scattering images	TJ-II
Feng et al. [50]	3D shape retrieval		X		X					X			An ensemble of autoencoders in which each autoencoder is trained to learn a compressed representation of depth views synthesized from each database objects is proposed. Each autoencoder as a probabilistic model that generates a likelihood score. Weakly supervised learning is used for training.	Low-cost 3D sensor	Multi-View RGBD, NYU-v2
Firth [51]	Recurrent CNN for ocean and weather forecasting		X	X									A large number (186) of RNN networks are used to simulate solutions to numerical weather prediction problems using temporal atmospheric data.	Various atmospheric data from NOAAPORT (wind speed, temp. relative humidity, specific humidity)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Geng et al. [52]	SAR classification					X						X	A deep convolutional autoencoder (DCAE) extracts features and conduct classification. It is composed of eight layers: a convolutional layer to extract texture features, a scale transformation layer to aggregate neighbor information, four layers based on sparse autoencoders to optimize features and classify, and the last two layers for post processing.	SAR (HH polarizatio n TerraSAR-X over Lillestroem, Norway)	Custom
Georgakis et al. [53]	multi-view RGBD object recognition		X		X					X			A new multi-view 3D proposal generation method and several recognition baselines using AlexNet are given. A RGBD dataset is also provided.	RGBD	Custom
Ghazi et al. [54]	Plant identification using transfer learning		X							X			Transfer learning is utilized. To decrease the chance of overfitting, data augmentation techniques are applied based on image transforms such as rotation, translation, reflection, and scaling. Furthermore, the networks' parameters are adjusted and different classifiers are fused to improve overall performance. An in-depth performance evaluation of the critical factors affecting the fine-tuning of pre-trained models; specifically iteration size, batch size, and data augmentation is provided.	RGB Images	LifeCLEF 2015 Plant Task Dataset

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Goel et al. [55]	Detecting weed stress and corn nitrogen status	X					X						A decision tree and ANN approach was used to study weed stress and nitrogen status in corn.	HSI (CASI data over Ste-Anne-de-Bellevue, Quebec, Canada)	Custom
Gong et al. [56]	Multi-label image annotation		X							X			Convolutional architectures are combined with approximate top-k ranking objectives for multi-label annotation.	RGB	NUS-WIDE
Fu et al. [57]	HSI classification	X	X					X					Stacked autoencoders use softmax classifier for unsupervised pre-training. Fine tuning is done with a small number of training samples.	HSI	Pavia University
Garcia-Gutiérrez et al. [58]	LiDAR and imagery data fusion						X	X		X	X		A contextual classifier based on a Support Vector Machine (SVM) and an Evolutionary Majority Voting (SVM–EMV) are used to create thematic maps from LiDAR and imagery data.	LiDAR, Airborne Thematic Mapper multispectral (Spain)	Custom
Geng et al. [59]	High-resolution SAR Image Classification		X			X							Deep convolutional autoencoder, convolutional filters for Gray-level co-occurrence matrix and Gabor features, scale transformations reduce noise.	SAR (TerraSAR-X data imagery of Lillestroem, Norway)	Custom
Ghamisi et al. [60]	Land cover classification	X					X						Band selection using fractional order Darwinian particle swarm optimization.	HSI	Indian Pines and Pavia University
Gong et al. [61]	Change Detection in SAR Imagery		X			X							Deep Neural Net to do change detection without a difference image.	SAR	Ottawa and Yellow River

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Goyal et al. [62]	Vehicle type classification									X			Nonnegative Matrix Factorization preprocessing followed by feature extraction and hybrid deep neural network.	RGB (Overhead picture from Stanford University)	Custom
Guan et al. [63]	Tree Classification				X						X		Tree segmentation (individual trees) and deep Boltzmann machine classification using waveform modeling of tree geometric structure	LiDAR	Custom
Hadsell et al. [64]	Robotic scene understanding and path planning	X							X	X			A deep hierarchical network is trained to extract informative and meaningful features from an input image, and the features are used to train a real-time classifier to predict traversability.	Two stereo cameras and GPS	Custom
Han et al. [65]	High spatial resolution imagery scene classification							X		X			A patch-based spatial-spectral hierarchical convolutional sparse auto-encoder for unsupervised training.	RGB	UC Merced, Google Earth
Haque et al. [66]	Depth-based person identification		X	X	X					X			An attention-based model that reasons on human body shape and motion dynamics to identify individuals in the dark. A combination of convolutional and recurrent neural networks with the goal of identifying small, discriminative regions indicative of human identity using 4D spatio-temporal signatures.	RGBD	BIWI, IIT PAVIS, IAS-LAB

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
He et al. [67]	Classification	X	X				X	X					A deep stacking network (DSN) is used, which does not require stochastic gradient descent training. Feature extraction is obtained by a nonlinear activation function on the hidden layer nodes of each module (DSNs usually use linear weights).	HSI	Kennedy Space Center
Hedge et al. [68]	3D Object classification		X		X					X			Fusion of pixel-based and voxel-based CNNs.	CAD data and labels	Princeton ModelNet
Hou et al. [69]	Polarimetric SAR classification		X			X		X				X	The data coherency matrix are converted to a 9-dimentional data. A Restricted Boltzmann Machine (RBM) is trained by using these patches and the learned features and the elements of coherent matrix are combined to train a 3-layers DBNs for PolSAR image classification.	PolSAR imagery (Two custom AIRSAR imagery over Flevoland, Netherlands, and a custom ESAE multi-look images.)	Custom
Hu et al. [70]	Vehicle recognition		X							X			A Deep Boltzmann Machine (DBM) is trained on both image data and Log-Gabor, HOG and GIST features. The pre-feature processing significantly improves performance.	RGB	PASCAL VOC2012

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Hu et al. [71]	HSI classification	X	X				X	X					Deep convolutional neural networks are employed to classify HSI images directly in spectral domain. The architecture contains five layers with weights which are the input layer, the convolutional layer, the max pooling layer, the full connection layer, and the output layer.	HSI	Pavia University, Salinas
Huang et al. [72]	Vehicle Detection in urban point clouds		X		X						X		After curb detection and removal in the segmentation stage, the algorithm estimates the orientation of the candidates and uses it to handle the difficult cases such as the vehicles in the parking lot. An orthogonal-view CNN, which are based on the orthogonal view projections of the candidates, is used to detect vehicles.	LiDAR	Ottawa
Huang et al. [73]	Pansharpening	X	X				X			X			A modified sparse denoising autoencoder (MSDA) algorithm trains the relationship between high-resolution (HR) and low-resolution (LR) image patches. A stacked MSDA (S-MSDA) pretrains the Deep Neural Network (DNN). The entire DNN is then trained by a back-propagation after pretraining. The HRMS image will be reconstructed from the observed LR MS image using the trained DNN.	Panchromatic and MS (IKONOS and Quickbird datasets)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Huang et al. [74]	Traffic flow prediction		X	X									First paper that applies deep learning transportation research. A deep belief network (DBN) performs unsupervised feature learning. To incorporate multitask learning (MTL) in the deep architecture, a multitask regression layer is used above the DBN for supervised prediction.	Inductive loop traffic data	Caltrans Performance Measurement System database, Custom
Huang et al. [75]	Building extraction		X					X		X			A deep deconvolution network is trained on a public large-scale building dataset. The output saliency maps of the fine-tuned models are fused to produce the final building extraction result.	Panchromatic and MS	IEEE GRSS 2016 Data Fusion Contest
Huang et al. [76]	Point cloud labeling		X		X						X		3D point cloud labeling using 3D CNN. Data are voxelized and voxel keypoints are generated. Labels assigned by dominating category around each keypoint.	LiDAR	Ottawa
Huval et al. [77]	Lane and vehicle detection		X	X					X	X	X		An extensive empirical study of DL on lane and vehicle detection.	RGB video, LiDAR, GPS (Custom dataset from vehicle sensors driven in San Francisco Bay area)	Custom
Iftene et al. [78]	Very high spatial resolution imagery scene classification		X					X		X			Transfer learning. Fine-tune CNN on one dataset and apply to another.	VHR imagery	UC Merced, WHU-RS

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Ishii et al. [79]	Satellite Image object recognition	X	X				X	X					A cuda-convnet CNN was modified to use 14 X 14 input image sizes. Results show CNN outperformed SVM.	Satellite imagery (Custom LANSAT 8 imagery over Kanto and Kagoshima in Japan)	Custom
Jia et al. [80]	Classification	X	X				X	X					The key of this method is to restructure spectral feature images and choose convolution filters with a reasonable size, so that the spectral features of different land coverings in high dimensions can be extracted properly. Filter sizes of 3, 5, 7 and 9 (in each dimension) are investigated.	HSI	Kennedy Space Center
Jiang et al. [81]	Vehicle detection in satellite imagery		X								X		Graph-based initial vehicle localization. CNN then classifies into vehicle / non-vehicle.	Satellite imagery	Custom
Kampffmeyer et al. [82]	Semantic segmentation and uncertainty modeling in urban scenes		X		X			X		X			Using recent advances in measuring uncertainty for CNNs, their quality is evaluated both qualitatively and quantitatively in a remote sensing context.	Orthophotos and DSM	ISPRS
Kaiser [83]	Uses remote-sensing datasets to generate large ground-truth datasets for cities		X							X			Semi-automated FCN approach to semantic segmentation of public very high resolution satellite imagery	Google Maps (Images from Google Maps and street maps from OpenStreetMap)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Kehl et al. [84]	3D object detection and 6D pose estimation		X		X					X			3D object detection uses regressed descriptors of locally-sampled RGBD patches for 6D vote casting. A convolutional auto-encoder is employed for regression. During testing, scene patch descriptors are matched against a database of synthetic model view patches and cast 6D object votes.	RGBD	LineMOD
Kim et al. [85]	Human detection and activity classification		X									X	CNN on micro-Doppler radar signatures.	Custom 7.25 GHz Doppler radar (microDoppler)	Custom
Kira et al. [86]	Long-range pedestrian detection		X	X	X				X	X			A combination of stereo-based detection, classification using deep learning, and a cascade of specialized classifiers is designed for long-range pedestrian detection. Stereo images are used to perform detection of vertical structures which are further filtered based on edge responses. A convolutional neural network (CNN) is used for classification of pedestrians using both appearance and stereo disparity-based features. A second CNN classifier was trained specifically for the case of long-range detections using appearance only.	RGB Stereo imagery	Custom
Konoplich et al. [87]	UAV image vehicle detection		X					X		X			Adapted hybrid Neural Network (AHNN) – handles scale	Aerial photos	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Kontschieder et al. [88]	Deep neural decision forests for classification		X							X			Deep Neural Decision Forests unifies classification trees with the representation learning functionality known from deep convolutional networks, by training them in an end-to-end manner. A stochastic and differentiable decision tree model, which steers the representation learning usually conducted in the initial layers of a deep CNN. The decision forest provides the final predictions and it differs from conventional decision forests due to a principled, joint and global optimization of split and leaf node parameters.	RGB	MNIST, ImageNet
Kovordanyi et al [89]	Cyclone track forecasting		X					X		X			A multi-layer neural network, resembling the human visual system, was trained to forecast the movement of cyclones based on satellite images. The results indicate that multi-layer neural networks could be further developed into an effective tool for cyclone track forecasting using various types of remote sensing data.	Satellite imagery (NOAA AVHRR satellite imagery)	Custom
Krishnan et al. [90]	Vehicle detection and road scene segmentation		X	X					X	X			A spatial grid of classifications is generated and then regressing bounding-boxes for pixels with a high object confidence score refine the classifications.	RGB	KITTI, MNIST

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Kussul et al. [91]	Large-scale land cover mapping		X	X				X		X			Hierarchical model that includes self-organizing maps (SOM) for data preprocessing and segmentation (clustering), ensemble of multi-layer perceptrons (MLP) for data classification and heterogeneous data fusion and geospatial analysis for post-processing	SAR imagery (Custom Sentinel 1 SAR Imagery over Ukraine)	Custom
Kuwatta et al. [92]	Estimating corn crop yields	X	X	X			X	X					Several custom architectures using Caffe were investigated. A RBF SVM was used for regression estimation.	MODIS Enhanced Vegetation Index (EVI) data	Custom
Lagrange et al. [93]	Semantic labeling multisource Earth observation data	X	X					X		X	X		Multi-sensor fusion improves results, CNNs outperform traditional methods, transfer learning helps classifier results	RBG orthophoto, DSM, LiDAR	IEEE GRSS 2015 Data Fusion Contest
Landschützer et al. [94]	Estimating the inter-annual Atlantic ocean carbon sink variability		X	X									A two-step neural network approach to reconstruct basin-wide monthly maps of the sea surface partial pressure of CO_2 (pCO_2) at a resolution of $1^\circ \times 1^\circ$. The air-sea CO_2 flux maps are computed using a standard gas exchange parameterization and high-resolution wind speeds.	Surface ocean fugacity of CO_2 data	SOCAT
Lange et al. [95]	Online vehicle detection		X		X					X	X		This vehicle detection system uses convolutional neural networks on 2D image data. Calculation time of the algorithm can be significantly reduced by taking advantage LiDAR depth information.	RGB, LiDAR	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Längkvist et al. [96]	Classification and segmentation of MSI orthoimagery	X	X		X			X		X			A convolutional neural network (CNN) is applied to MSI orthoimagery and a digital surface model (DSM) of a small city for a full, fast and accurate per-pixel classification. The predicted low-level pixel classes are then used to improve the high-level segmentation. Various design choices of the CNN architecture are evaluated and analyzed.	MSI, DSM (over Sweden)	Custom
Lee et al. [97]	Estimate geo-information of photo (population density, demographics, etc.) via CNN									X			CNN to estimate geo-information of certain attributes.	RGB (Custom Flickr dataset - 40M images)	Custom
Lee et al. [98]	HSI classification	X	X				X	X					The contextual deep CNN jointly exploit spatial and spectral features for HSI classification. It concurrently applies multiple 3D local convolutional filters of different sizes jointly exploiting spatial and spectral features. The spatial and spectral feature maps are then combined to form a joint spatio-spectral feature map, which is the input to a set of fully convolutional layers that assign class labels.	HSI	Indian Pines, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Lee et al. [99]	Predicting geo-informative attributes		X							X			To recognize general geo-informative attributes of a photo, e.g. the elevation gradient, population density, demographics, etc., these attributes are estimated using a large (noisy) set of geo-tagged images from Flickr by training deep convolutional neural networks (CNNs). A large-scale dataset is also provided.	RGB (Flickr)	Custom
Levi et al. [100]	Obstacle detection and road segmentation		X	X					X	X	X		To detect obstacles, a single color camera is used and the task reduces to a column-wise regression problem, which is solved using a deep convolutional neural network (CNN).	RGB, LiDAR	KITTI, Custom
Li [101]	HSI classification	X	X				X						Stacked autoencoders train a DL network. Iterative procedure reduces uncertainty and allows better generalization capability.	HSI	Pavia University
Li [102]	Vehicle detection in point clouds		X		X						X		The fully convolutional network based detection techniques are extended from 2D to 3D and applied it to point cloud data. The CNN creates "objectness" and bounding box maps.	RGB	KITTI

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Li et al. [103]	Spatial and color pooling for 3D object recognition		X		X					X			To make object representations that are more robust to viewpoint changes, multiple scales of filters coupled with different pooling granularities are applied in a CNN and color is also used as an additional pooling domain. A new dataset for industrial objects is also provided (JHUIT-50).	RGBD	Multi-View RGBD, BigBIRD, JHUIT
Li et al. [104]	Classification	X	X				X	X					Principle components of the HSI image are filtered by three dimensional Gabor wavelets, followed by stacked autoencoders are trained on the outputs of the previous step through unsupervised pre-training, finally deep neural network is trained on those stacked autoencoders.	HSI	Indian Pines
Li et al. [105]	HSI image classification	X	X				X	X	X				Restricted Boltzmann machine and Deep Belief Network. DBN using 7x7 spatial neighbors and 144 HSI bands.	HSI	IEEE GRSS 2013 Data Fusion Contest
Li et al. [106]	HSI image classification	X	X				X	X					Deep CNN using pixel-pair features and voting classification. Achieves high results with small amount of training data.	HSI	Indian Pines, Salinas, Pavia University
Li et al. [107]	HSI image classification	X	X				X	X					Spatial features-based strategy for band selection, CNN optimized parameter model, then image is classified by the efficient extreme learning machine (ELM). Fast computation.	HSI	Indian Pines, Salinas, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Li et al. [108]	Land-cover mapping	X	X				X	X					Hand-tuned stacked auto-encoder.	MSI (Custom LANDSAT and time-series MODIS)	Custom
Li et al. [109]	Road network extraction		X					X					A convolutional neural network (CNN) is are applied to predict the probability of a pixel belonging to road regions, and assign labels to each pixel to describe whether it is road. A line integral convolution based algorithm is developed to smooth the rough map to connect small gaps. Finally, using common image processing operators, road centerlines are delineated.	Satellite images (Geoeye and Pleaides datasets)	Custom
Li et al. [110]	HSI classification	X	X				X	X					A restricted Boltzmann machine (RBM) model and its deep structure deep belief networks (DBN), are introduced in HSI image processing as the feature extraction and classification approach utilizing spatial-spectral classification.	HSI	IEEE GRSS 2013 Data Fusion Contest

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Li et al. [111]	HSI Anomaly detection		X				X	X					This approach uses transfer learning. A multilayer CNN is trained by using difference between pixel pairs generated from the reference image scene. Then, for each pixel in the image for anomaly detection, difference between pixel pairs, constructed by combining the center pixel and its surrounding pixels, is classified by the trained CNN with the result of similarity measurement. The detection output is simply generated by averaging these similarity scores.	HSI	Custom, Cuprite
Liebel et al. [112]	Super-resolution HSI image processing	X	X				X	X					Custom CNN with super-resolution output.	HSI	Copernicus SENTINEL
Lin et al. [113]	HSI classification	X	X				X	X					Extract features via autoencoder. PCA (spectral) + autoencoder on spatial dimensions.	HSI	Kennedy Space Center, Pavia University
Liu et al. [114]	SAR terrain classification					X						X	Spatial information between pixels on PolSAR image is combined into the input data. The proposed deep network only needs to tune a few parameters during pre-training and fine-tuning.	PolSAR (L-band PolSAR data over Flevoland, Netherlands)	Custom
Liu et al. [115]	HSI classification	X	X				X	X					Spectral used stacked denoising autoencoders, superpixels used for spatial constraints.	HSI	Indian Pines, Salinas, Pavia Center, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Liu et al. [116]	HSI classification	X	X				X	X					Spectral used stacked denoising autoencoders, superpixels used for spatial constraints.	HSI	Indian Pines, Salinas, Pavia Center, Pavia University
Liu et al. [117]	HSI classification	X	X				X	X					An active learning algorithm based on a weighted incremental dictionary learning is proposed for such applications. The proposed algorithm selects training samples that maximize two selection criteria, namely representative and uncertainty. This algorithm trains a deep network efficiently by actively selecting training samples at each iteration.	HSI	Pavia City Center, Pavia University, Botswana
Liu et al. [118]	Geological disaster recognition		X					X		X			A deep learning based landslide recognition method for optical remote sensing images. Preprocessing is performed using a wavelet transformation. A denoising method enhances the robustness of the model in recognize landslide features. Then, a deep autoencoder network with multiple hidden layers is proposed to learn high-level features. A softmax classifier is used for class prediction.	Satellite imagery (Google Earth)	Custom
Luus et al. [119]	Multispectral land use classification	X	X				X	X					DL learns hierarchical feature representation.	aerial orthoimagery	UC Merced

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Lv et al. [120]	Traffic Flow Prediction	X											Stacked AE.	Traffic data every 30sec from 15,000 sensors	Caltrans Performance Measurement System database
Lv et al. [121]	Landcover classification	X	X	X		X		X				X	Multi-temporal PolSAR data is analyzed using a pixel window where the data is stacked and processed via a multi-layer RBM.	PolSAR multi-temporal data (RADARSAT-2 C-band data over Toronto, Ontario, Canada)	Custom
Lv et al. [122]	Urban land use and land cover	X	X	X		X		X				X	A DBN model extracts contextual mapping features from the PolSAR data to improve the classification performance.	PolSAR multi-temporal data (RADARSAT-2 C-band data over Toronto, Ontario, Canada)	Custom
Ma et al. [123]	Land Classification	X	X				X	X					Pre-label unlabeled using a local decision based on weighted neighborhood information, and a global decision based on deep learning is performed by the most similar training samples. Unlabeled ones with high confidence are selected to extend the training set.	HSI	Indian Pines, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Ma et al. [124]	Classification	X	X				X	X					A regularization term in the energy function encodes sample similarity and features are updated by integrating contextual information. A collaborative representation-based classification helps deal with small training datasets. To suppress salt-and-pepper noise, a graph-cut-based spatial regularization is performed.	HSI	Indian Pines, Pavia Center, Botswana
Ma et al. [125]	Classification	X	X				X	X					A prior is imposed on the deep network to deal with the instability of parameter estimation. The proposed method adjusts parameters of the whole network to minimize the classification error as all supervised deep learning algorithm and minimizes the discrepancy within each class and maximize the difference between different classes.	HSI	Washington DC Mall, Pavia City Center
Makantasis et al. [126]	Land Classification	X	X				X	X					Deep learning based classification method that hierarchically constructs high-level features. CNN to encode pixels' spectral and spatial information, Multi-Layer Perceptron classifier	HSI	Indian Pines, Salinas, Pavia Centre, Pavia University
Marmanis et al. [127]	Urban object classification							X					The MLP detects patterns in DEM that characterize buildings.	Digital Elevation Maps (VHSR DEM Munich, Dongying City)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Marmanis et al. [128]	Semantic segmentation of aerial imagery		X		X			X		X			A CNN with added deconvolutional network layers to undo the spatial downsampling and Fully Convolution Networks (FCNs) are used to perform pixel-based classification at full resolution.	Aerial orthoimagery, DSM	ISPRS
Marmanis et al. [129]	Classification		X					X		X			A pre-trained CNN designed for tackling an entirely different classification problem (the ImageNet challenge) is used to extract an initial set of representations. The derived representations are then transferred into a supervised CNN classifier, along with their class labels, effectively training the system. This two-stage framework successfully deals with the limited-data problem commonly encountered in HSI processing.	Orthoimagery	ImageNet, LULC
Masi et al. [130]	Pansharpening	X	X				X	X		X			Use CNN for pansharpening. Uses Wald protocol to downsample high-res image to MS image reference	Aerial Photos, multispectral (IKONOS, GeoEye-1, WorldView-2)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Maturana et al. [131]	LiDAR landing zone detection		X		X						X		A volumetric occupancy map is paired with a 3D Convolutional Neural Network (CNN) is applied to detecting safe landing zones for autonomous helicopters from LiDAR point clouds.	LiDAR (Synthetic dataset, semi-synthetic datasets - simulated solid objects with real point cloud data for vegetation)	Custom
Mei et al. [132]	Land Classification	X	X				X	X					5-layer CNN and Parametric ReLU.	HSI	Indian Pines, Pavia University, Salinas
Mei et al. [133]	Infrared ultraspectral image classification	X					X	X					An arctan-like term is added to the objective function as a sparse constraint to improve classification accuracy. A Gaussian prior helps avoid overfitting. A multi-layer Restricted Boltzmann Machine model, a deep belief network provides ultraspectral signature classification.	USI (ASTER and Environmental Protection Agency data)	Custom
Merentis et al. [134]	Classification	X	X		X		X	X			X		To handle data fusion of high orders, outputs of DL first-layer are used to perform HSI and LiDAR data fusion.	HSI, LiDAR	Indian Pines, IEEE GRSS 2013 Data Fusion Contest
Midhun et al. [135]	Land Classification	X	X				X	X					Non-linear band-by-band diffusion preprocessing is followed by a Restricted Boltzmann Machine (RBM) that encodes spectral features. Regenerative RBM generates features for classifier.	HSI	Indian Pines

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Mnih et al. [136]	Road Detection		X					X		X			Two robust loss functions handle incomplete and poorly labeled training imagery.	High-resolution aerial imagery	URBAN1, URBAN2
Mou et al. [137]	Space video scene interpretation	X	X	X			X	X		X			Fusion of spaceborne video and MSI is performed. A deep neural network uses unpooling to create coarse probability maps which are refined into average superpixel maps. Traffic activities are analyzed using tracklets produced by a Kanade-Lucas-Tomasi keypoint tracker. The labeled ground truth and visualization code are available at http://www.sipeo.bgu.tum.de/downloads/gt4dfc16video.rar	MSI, satellite video	IEEE GRSS 2016 Data Fusion Contest
Morgan [138]	SAR ATR		X			X		X					A custom CNN is designed. It can detect targets not in training set.	SAR	MSTAR
Nogueira et al. [139]	Land Classification	X					X	X		X			Investigate fine tuning, full training, and feature extractor from CNN. SVM classifier.	HR aerial photos and MSI	UC Merced, RS19, Brazilian Coffee
Ni et al. [140]	SAR ATR		X			X						X	This algorithm contains three stages: (1) Image preprocessing where a Kuan filter provided image enhancement and an adaptive Intersecting Cortical Model (ICM) to do the segmentation, (2) feature extraction using a sparse autoencoder, and (3) classification using a softmax regression classifier	SAR	MSTAR

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Ondruska et al. [141]	Multi-object tracking		X	X							X		A recurrent neural network (RNN) learns mapping from sensor measurements to object tracks which handles occluded sensor data. The system can track a large number of dynamic objects with occlusion.	2D laser scan (2D laser data)	Custom
Othman et al. [142]	Land Use Classification	X	X				X	X		X			An initial feature representation is generated by a CNN pre-learned on a large amount of labeled data from an auxiliary domain. These are inputs to a sparse autoencoder for learning a new suitable representation in an unsupervised manner.	High resolution aerial photos and MSI	UC Merced, Banja-Luka LU Public
Ouwang et al. [143]	Pedestrian detection with occlusion handling		X						X	X			A deformable part-based model is used to obtain the scores of part detectors and the visibilities of parts are modeled as hidden variables. A discriminative deep model is used for learning the visibility relationship among overlapping parts at multiple layers. A new dataset (CUHK) is provided for testing occlusion handling in pedestrian detection.	RGB	Caltech Pedestrian, ETH, Daimler
Pacifici et al. [144]	Change detection in HR satellite imagery		X	X				X		X			This change detection algorithm is based on neural networks, and it is able to exploit both the multi-band and the multi-temporal data to discriminate between real changes and false alarms. In general, the classification errors are reduced by a factor of 2–3.	HR satellite imagery (Landsat data over Tor Vergata University in Rome, Italy, and Rock Creek in Superior, CO, USA.)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Paisitkriangkrai et al. [145]	Semantic pixel labelling		X					X		X			Both CNN and hand-crafted features are applied to dense image patches to produce per-pixel class probabilities. A CNN creates probability maps and predicted labels. The conditional random field processing infers a labeling that smoothes regions while preserving edges.	Aerial orthoimagery	ISPRS
Palafox et al. [146]	Detection of impact craters and volcanic rootless cones in HR Martian imagery		X					X		X			An AE is used to determine descriptive features. A two convolutional layer CNN is then used to combine spatial features.	HR imagery (HiRISE data over the Elysium Planitia, Mars)	Custom
Pal et al. [147]	Urban growth prediction		X				X			X			A CNN is proposed with no pooling and two layers of convolution. The class of a given pixel is related to the classes around it.	Satellite imagery (Landsat 7 ETM+ imagery over Mumbai, India and Rajarhat, Kolkata)	Custom
Pan et al. [148]	Classification		X				X	X		X			The vertex component analysis network (R-VCANet) achieves higher accuracy with a small number of training samples. The inherent properties of HSI data, spatial information and spectral characteristics, are utilized to construct the network. Spectral and spatial information are combined via the rolling guidance filter, which examines the structure features and removes small details from HSI.	HSI	Indian Pines, Pavia University, Kennedy Space Center

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Papadomanolaki et al. [149]	VHR MSI Classification	X	X				X	X					AlexNet FC7 layer + SVM classifier.	VHR MSI (R,G,B,NIR) (NAIP)	Custom
Penatti et al. [150]	Land Classification	X	X				X	X		X			Transfer learning for ConvNets trained for everyday objects to classify aerial and remote sensing imagery.	High resolution aerial photos and MSI	UC Merced, Brazilian Coffee
Piramanayagam et al. [151]	Classification	X	X			X		X				X	A CNN and random forest with decision trees based on features obtained from image patches and labels are compared.	Satellite imagery (AIRSAR L-band imagery over Flevoland, Netherlands)	ISPRS, Custom
Qin et al. [152]	Object-oriented PolSAR classification		X			X						X	A RBM is used with an AdaBoost instead of a stacked deep model, for object-oriented classification in PolSAR imagery. The experimental results demonstrate that the proposed model is superior to the stacked RBM model.	PolSAR (PolSAR L-band imagery over Flevoland, Netherlands)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Qin et al. [153]	Underwater live fish recognition		X	X					X	X			Sparse and low-rank matrix decomposition is used to extract foreground information. Then, a deep architecture is used to extract features of the foreground fish images. Principal component analysis (PCA) is used in two convolutional layers, followed by binary hashing in the non-linear layer and block-wise histograms in the feature pooling layer. Spatial pyramid pooling (SPP) is used to extract information invariant to large poses. Classification uses a linear SVM.	RGB video	Fish Recognition Ground-Truth dataset
Qin et al. [154]	Underwater live fish recognition		X	X					X	X			Sparse and low-rank matrix decomposition is used to extract foreground information. Then, a deep architecture is used to extract features of the foreground fish images. Principal component analysis (PCA) is used in two convolutional layers, followed by binary hashing in the non-linear layer and block-wise histograms in the feature pooling layer. Spatial pyramid pooling (SPP) is used to extract information invariant to large poses. Classification uses a linear SVM.	RGB video	Fish Recognition Ground-Truth dataset
Qu et al. [155]	Semantic understanding of high-res imagery		X					X		X			Deep multimodal NN analyzes image and exports sentences with content	HR aerial photos	UC Merced, Sydney

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Quan et al. [156]	SAR image registration		X	X		X		X				X	The deep learning module learns essential features of the images and a new algorithm to help remove incorrect matching points using RANSAC are implemented.	SAR imagery (June 2008 and 2009)	Yellow River
Rajan et al. [157]	Classification	X	X				X	X					An active learning technique that efficiently updates existing classifiers by using fewer labeled data points than semisupervised methods. This method is well suited for learning or adapting classifiers when there is substantial change in the spectral signatures between labeled and unlabeled data.	HSI	Kennedy Space Center, Botswana
Ran et al. [158]	Classification	X	X				X	X					A deep convolutional network splits the spectrum bands into groups based on their correlation relationships. A band variant CNN sub-model is built, where each group is modeled by one of those sub-models. A conventional CNN model is also learned globally on the spatial-spectral space, to maintain robustness of sub-model changes. The global CNN model and band-specific CNN sub-models are fused into to one unique model.	HSI	Indian Pines, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Rebetez et al. [159]	Crop classification in HR UAV imagery		X					X		X			A deep neural network which consists of a convolutional side (CNN) which uses the raw pixel values and a dense side which uses RGB histograms (HistNN). The output of both networks was merged by a final layer which predicts the class of each pixel.	HR imagery (from UAV)	Custom
Romero et al. [160]	Land use and land cover classification	X	X				X	X		X			Greedy layer-wise unsupervised pre-training with an efficient algorithm for unsupervised sparse feature learning. Algorithm simultaneously enforces population and lifetime sparsity.	land use: VHR imagery land cover: HSI (Custom Quickbird II multispectral)	UC Merced, Custom
Saito et al. [161]	Building and road detection for large aerial imagery		X					X		X			Convolutional Neural Networks (CNN) learn the mapping from raw pixel values in aerial imagery to three object labels (buildings, roads, and others).	high spatial resolution aerial imagery	Mass. Building, Mass. Roads

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Salberg [162]	Seal detection using transfer learning	X	X					X		X			A CNN pre-trained on ImageNet is used perform object recognition in remote sensing data. The method consists of three stages: (i) Detection of potential objects, (ii) feature extraction and (iii) classification of potential objects. The first stage is application dependent, with the aim of detecting all seal pups in the image, with the expense of detecting a large amount of false objects. The second stage extracts generic image features from a local image corresponding to each potential seal detected in the first stage using a CNN trained on the ImageNet database. The third stage uses a linear SVM for classification.	MSI: R,G,B,NIR (MSI image with 123 adult and 84 pup harp seals)	Custom
Shwegman et al. [163]	Ship discrimination in SAR imagery		X			X		X				X	Highway Networks allow for very deep networks that can be trained using the smaller datasets typical in SAR-based ship detection. A very deep Highway Network perform ship discrimination stage for SAR ship detection. The paper also presents a three-class SAR dataset that allows for more meaningful analysis of ship discrimination performances.	SAR imagery (Custom dataset using 22 Sentinel-1 and 3 Radarsat 2 images)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Sedaghat et al. [164]	3D object recognition		X		X					X	X	X	The system predicts the object pose and the class label. This yields significant improvements in the classification results. An orientation-boosting voxel net was proposed and LiDAR data, CAD models and RGBD images were analyzed.	LiDAR, RGBD	Sydney, KITTI, NYU-v2, Princeton ModelNet
Sherrah et al. [165]	Dense semantic labelling of HR imagery		X					X		X			Deep CNNs are applied to semantic labelling of HR remote sensing data. A full-resolution labelling is inferred using a deep FCN with no downsampling (requires no deconvolution or interpolation).	Orthoimagery (R,G,IR), (R,G,B,IR) and DSM	ISPRS
Shi et al. [166]	Cloud detection		X					X		X			The deep CNNs consists of four convolutional layers and two fully-connected layers. The image is clustered into superpixels as sub-region through simple linear iterative clustering, and the probability of each superpixel that belongs to cloud region is generated. The cloud region is obtained according to the gradient of the cloud map.	RGB	Custom Quickbird, Google and Flickr imagery

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Shi et al. [167]	Precipitation nowcasting		X	X								X	Precipitation nowcasting as a spatiotemporal sequence forecasting problem in which both the input and the prediction target are spatiotemporal sequences. By extending the FC LSTM to have convolutional structures in both the input-to-state and state-to-state transitions, the convolutional LSTM (ConvLSTM) is used for precipitation nowcasting.	Weather radar intensity data (Custom radar weather intensity)	Custom
Sladojevic et al. [168]	Plant disease recognition									X			Plant disease recognition based on CNN leaf image classification using Caffe.	RGB imagery	Custom
Slavkovikj et al. [169]	Classification	X	X				X	X					The CNN is able to learn spectral band-pass filters during image classification.	HSI	Indian Pines
Stent [170]	Change detection in tunnel imagery		X							X			A two-channel CNN implements change detection using pairs of approximately registered image patches taken at different times and classifies them to detect anomalous changes.	RGB (Custom dataset with artificial cracks, leaks and rust applied to tunnels)	Custom
Sun et al. [171]	Land classification	X	X				X	X					Autoencoder classifier using RBM pretraining and multinomial logistic regression to generate outputs.	HSI	Kennedy Space Center, Indian Pines

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Sun et al. [172]	ATR in SAR		X		X		X				X		The method makes use of a probabilistic NN, RBM, modeling probability distribution of environment. The model learns a shared representation of the target and its shadow to reflect the target shape.	SAR	MSTAR
Tang et al. [173]	Ship detection satellite imagery		X					X		X			A ship detection approach to uses wavelet coefficients extracted from JPEG2000 compressed domain combined with DNN and ELM. Compressed domain is adopted for fast ship candidate extraction, DNN is exploited for high-level feature representation and classification, and ELM is used for efficient feature pooling and decision making	Satellite imagery (SPOT 5 panchromatic images)	Custom
Tome et al. [174]	Pedestrian detection		X						X	X			Fast CNN running on Jetson TK1. Optimized detection pipeline.	RGB	Caltech Pedestrian
Uba [175]	Land use and land cover classification	X	X				X	X		X			Uses transfer learning with ConvNets and dataset augmentation with rotated imagery	High resolution aerial photos and MSI	UC Merced, Brazilian Coffee

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Vaduva et al. [176]	Semantic labelling of VHR imagery		X					X		X			A DL approach depicts the lowest level by the primitive feature vectors color, texture and shape. At the next level of representation, single objects like building, highway, forest, boat or lake are described by unique combinations of the primitive feature vectors. Object spatial interactions are learned at a hierarchical level. Using Latent Dirichlet Analysis, three-level hierarchy, in which documents of a corpus are represented as random mixtures over latent topics and each topic is characterized by a distribution over words.	Satellite imagery (Custom WorldView 2 image of Bucharest, Romania)	Custom
Vakalopoulou et al. [177]	Building detection in VHR MSI	X	X				X	X		X			Automated building detection using deep convolutional NNs with a MRF model to label detected buildings.	VHR MSI imagery (Quickbird and WorldView-2 imagery)	Custom
Volpi et al. [178]	Dense semantic labeling of sub-decimeter images	X	X					X		X			A CNN-based system relying on a downsample-then-upsample architecture learns a rough spatial map of high-level representations by means of convolutions and then learns to upsample them back to the original resolution by deconvolutions. By doing so, the CNN learns to densely label every pixel at the original image resolution.	VHR imagery (R,G,NIR)	ISPRS

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Wang et al. [179]	Object detection		X					X		X			THSI elementary idea of using zeroth order (depth), first-order (surface normal) and second-order (surface curvature) features are also supplied to a CNN hat has been pre-trained on a color image database.	RGBD	NYU-v2, Multi-View RGBD
Wang et al. [180]	Road detection in very high resolution satellite imagery		X					X		X			CNN plus state machine for road extraction.	VHR satellite imagery (Google Earth imagery)	Custom
Wang et al. [181]	HSI classification	X	X				X	X					Uses semisupervised approach and utilizes unlabeled image samples by evaluating the confidence probability of the predicted labels. Classifier parameters and dictionary atoms jointly optimized. Spatial information used via Laplacian smoothness regularization to the output of the classifier making the spatial constraint more flexible	HSI	Indian Pines, Salinas, Pavia University

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Wang et al. [182]	Vehicle detection		X							X			A deep learning based vehicle detection algorithm with 2D deep belief network (2D-DBN) is proposed. In the algorithm, the proposed 2D-DBN architecture uses second-order planes instead of first-order vector as input and uses bilinear projection for retaining discriminative information so as to determine the size of the deep architecture which enhances the success rate of vehicle detection.	RGB	Caltech1999
Wang et al. [183]	Vehicle brake-light detection		X		X				X	X	X		On a large database, brake lights patterns are learned by a multi-layer perception NN. Given an image, the vehicles can be classified as “brake” or “normal” using the deep classifier. The vehicle can be detected quickly and robustly by combining multi-layer LiDAR and fusing with mono-camera imagery. Road segmentation and a novel vanishing point region of interest determination method are explored to further speed up the detection and improve the robustness. The AlexNet model was used for the CNN.	mono RGB camera, LiDAR (Custom imagery hand-labeled as brake or normal)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Wang et al. [184]	Nigh-time vehicle sensing in far IR		X										Vehicle candidates are generated by thresholding the IR data and then post-processed by contour analysis to reduce the false positive rate. Finally, vehicle candidates are verified using a DBN-based classifier.	IR imagery	Custom
Wang et al. [185]	SAR ATR		X			X		X				X	A custom CNN (A-ConvNet) is proposed with reduced parameters and helps mitigate overfitting. A-ConvNet replaces the fully connected layers with convolutional layers.	SAR	MSTAR
Wang et al. [186]	Vehicle detection in satellite imagery		X					X					Parallel DNN provides fast implementation.	Aerial photos (Google Earth)	Custom
Wang et al. [187]	Segmentation of HR images		X					X		X			A fast scanning image segmentation within a DL.	HR imagery	UC Merced, IEEE GRSS 2013 Data Fusion Contest
Wei et al. [188]	Image denoising and haze removal	X	X				X	X					A CNN learns parameters for a denoising function. The network can be trained on one sensor and applied successfully to another.	HSI (UC Merced (training), custom Quickbird and DC Mall (testing))	Custom
Williams et al. [189]	Sonar classification		X	X									The deep networks are learned using a massive database of real, measured sonar data collected at sea during different expeditions in various geographical locations. Training data is augmented with synthetic data avoid overfitting.	Sonar data	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Wu et al. [190]	Shape-based object extraction		X					X		X			DBMs model the shape priors via the unsupervised training process. An energy function uses the model integrated into a new energy function to eliminate the influence of disturbing background and combines local and region information. A new region term in the function is proposed to eliminate the influence of object shadows.	RGB (Custom dataset of QuickBird images of aircraft)	Custom
Xie et al. [191]	Poverty mapping		X					X		X			A novel machine learning approach to extract large-scale socioeconomic indicators from HR satellite imagery. A FC CNN learns filters identifying different terrains and man-made structures, including roads, buildings, and farmlands, without any supervision beyond nighttime lights. These learned features are highly informative for poverty mapping.	Satellite imagery (Custom NOAA nighttime satellite images)	Custom
Xie et al. [192]	3D shape feature learning		X		X					X			A Multi-View Deep Extreme Learning Machine (MVD-ELM) ensures the feature maps learned for different views are mutually dependent via shared weights and in each layer and that the "unprojections" form a valid 3D reconstruction of the input 3D shape via normalized convolution kernels, enabling clear visualization of the learned features.	3D database	Princeton ModelNet

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Yang et al. [193]	HSI classification	X	X				X	X					A two-channel CNN jointly learns spectral and spatial features. A fully connected layer then fuses the features for classification.	HSI	Indian Pines, Salinas
Yang et al. [194]	Cloud detection		X				X	X					Two ANN approaches are satisfactory in discriminating cloud phase using FY-3A/Visible and InfRared Radiometer (VIRR) multi-channel data,	HSI Satellite imagery (Custom multi-channel satellite data)	Custom
Yao et al. [195]	Urban area classification		X					X		X			A fusion of CNN and hand-crafted features based on evidence combination theory.	IR-R-G orthoimages	ISPRS
Yu et al. [196]	HSI classification	X	X				X	X					CNN optimized to use smaller training sets, 1 X 1 convolutional layers, average pooling and larger dropout.	HSI	Indian Pines, Salinas, Pavia University
Yu et al. [197]	Automated road marking extraction from 3D LiDAR point clouds		X		X						X		Road surface points are segmented using a curb-based approach. Road markings are extracted. Seven specific types of road markings are further accurately delineated through a combination of Euclidean distance clustering, voxel-based normalized cut segmentation, large-size marking classification, and small-size marking classification based on DL, and principal component analysis.	LiDAR	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Yu et al. [198]	Automated road manhole extraction		X		X						X		Road surface points are segmented from a raw point cloud via a curb-based road surface segmentation. A supervised DL model is developed to construct a multilayer feature generation model for depicting high-order features of local image patches, and a random forest model is trained to learn mappings from high-order patch features to the probabilities of the existence manhole covers. The manhole covers are detected using the multilayer feature generation and random forest.	LiDAR	Custom
Yuan et al. [199]	Scene recognition		X							X			A manifold regularized deep architecture exploits the structural information of the data, making for a mapping between visible layer and hidden layer and learns in an unsupervised manner.	RGB	15 Scenes, Sports-8, SUN-397
Yue et al. [200]	HSI classification	X	X				X	X					This method merges spatial and spectral features stacked auto- encoders (SAEs) and deep convolutional neural networks (DCNNs) followed by a logistic regression (LR) classifier. Spatial pyramid pooling (SPP) to pool the spatial feature maps of the top convolutional layers into a fixed-length feature.	HSI	Pavia City Center

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Yue et al. [201]	HSI classification	X	X				X	X					The deep convolutional neural networks (DCNNs) hierarchically extract deep features of HSI. The feature map generation algorithm generates the spectral and spatial feature maps. Next, the DCNNs Logistic Regression classifier is trained to get useful high-level features and to fine-tune the whole model.	HSI	Pavia City Center
Zabalza et al. [202]	HSI dimensionality reduction	X	X				X	X					Segmented stacked auto encoders (S-SAE) divides the original features into smaller data segments, which are separately processed by different smaller SAEs, which results in reduced complexity but improved efficacy of data abstraction and classification accuracy.	HSI	Indian Pines, Pavia City Center
Zeggada et al. [203]	UAV image classification		X					X		X			A UAV-shot image is first subdivided into a grid of equal tiles and deep NN-induced features are extracted from each tile and then fed into a radial basis function NN classifier. A refinement step at the top of the complete deep network architecture helps to boost the classification results.	RGB (Custom UAV imagery over Trento, Italy)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zelener et al. [204]	Urban LiDAR segmentation (with missing points)		X		X						X		The segmentation task is reframed over the scan acquisition grid versus the 3D point cloud. Missing points in the scan grid are labeled and this improves classifier accuracies. The choice of input features maps to the CNN significantly affect segmentation accuracy and these features should be chosen to fully encapsulate the 3D scene structure.	LiDAR (Google Street View data over New York City)	Custom
Zeng et al. [205]	Traffic sign recognition	X	X						X	X			Combines CNN for feature learning with extreme learning machine classifier.	RGB video	German Traffic Sign Recognition Benchmark (GTSRB)
Zhang et al. [206]	High resolution satellite scene classification		X					X		X			Gradient boosting random CNN proposed which combines multiple CNNs.	HR satellite imagery (Custom Goggle earth imagery over Sydney)	UC Merced, Custom
Zhang et al. [207]	Co-saliency detection		X							X			The wide and deep information are explored for the object proposal windows extracted in each image, and the co-saliency scores are calculated by integrating the intra-image contrast and intra-group consistency via a principled Bayesian formulation. Finally the window-level co-saliency scores are converted to the superpixel-level co-saliency maps through a foreground region agreement strategy.	RGB	iCoseg, MSRC

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zhang et al. [208]	Scene classification	X	X				X	X					A gradient boosting random convolutional network (GBRCN) framework for scene classification, which can effectively combine many deep neural networks and outperforms single deep learning approaches.	Satellite imagery	UC Merced, Sydney
Zhang et al. [209]	HSI Spectral-spatial classification	X	X				X	X					A dual-channel convolutional neural network (DC-CNN) framework has a one-dimensional CNN to automatically extract the hierarchical spectral features and two-dimensional CNN is applied to extract the hierarchical space- related features, and then a softmax regression classifier to fuse the spectral and spatial features and perform classification. A simple data augmentation method is proposed to overcome limited training data.	HSI	Pavia University, Indian Pines
Zhang et al. [210]	IR imagery enhancement		X						X				A systematic approach based on image bias correction and DL is proposed to increase target signature resolution and optimize the baseline quality of inputs for object recognition.	IR imagery	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zhang et al. [211]	Oil-tank detection		X					X		X			First, a modified ellipse and line segment detector based on gradient orientation is used to select candidates in the image and HOG features are extracted. For the surrounding area, the CNN transfer learning is applied. A linear SVM provides classification.	Aerial photos	Custom
Zhang et al. [212]	Urban building detection		X					X		X			A multi-scale saliency computation is employed to extract built- up areas and a sliding windows approach is applied to generate candidate regions, a CNN is applied to classify the regions, and then an improved non maximum suppression is used to remove false buildings.	Satellite imagery (Custom Google Earth imagery in Guanghua, Sanchuan, Songping, and Chenghai counties of Yunnan province, China)	Custom
Zhang et al. [213]	Ship Detection		X					X		X			A ship detection method based on CNNs, called S-CNN, uses two ship models (the "V" ship head model and the "||" ship body) to localize the ship proposals from extracted line segments. A saliency detection method to find small ships is also developed. Both of these ship proposals are fed to the trained CNN for robust and efficient detection.	RGB (aerial imagery)	Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zhang et al. [214]	Urban functional zone classification		X					X		X			A CNN-based functional zone classification method is proposed. The method consists of three steps. The aerial imagery of the city is partitioned into disjoint regions by road network and then each region is further divided into patches and is fed to a fully connected CNN. The output of which is considered as distributions of this patches on the five previously defined functional zones. A vote strategy is used to identify the function zone of this region.	RGB (Google Earth imagery over Shenyang and Beijing, China)	Custom
Zhao et al. [215]	Change detection		X	X		X		X				X	A deep neural network uses unsupervised feature learning and supervised fine-tuning.	SAR images	Ottawa, Bern, and Yellow River
Zhao et al. [216]	SAR imagery change detection		X			X						X	Three classes: positive change backscatter, negative change backscatter and no change learned through a deep belief network using a log-ratio operator on SAR imagery.	SAR satellite imagery (custom ESAERS-2 image of San Francisco)	Yellow River, Custom

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zhao et al. [217]	Multi-scale land classification	X	X				X	X					A multi-scale CNN (MCNN) to learn spatial-related deep features for hyperspectral remote imagery classification. MCNN first transforms the original data sets into a pyramid structure containing spatial information at multiple scales, and then automatically extracts high-level spatial features using multi-scale training data sets.	Hyperspectral imagery (Custom Worldview-II image of Beijing, China)	Pavia City Center, Pavia University, Custom
Zhong et al. [218]	Classification of high spatial resolution satellite imagery		X					X		X			The large patch convolutional NN (LPCNN) uses large patch sampling to generate hundreds of possible scene patches for the feature learning, and a global average pooling layer is used to replace the fully connected network as the classifier, which can greatly reduce the total parameters.	High spatial resolution satellite imagery (IKONOS)	Custom
Zhong et al. [219]	Road and building extraction in high spatial resolution imagery		X					X		X			The influence of filter stride, learning rate, input data size, training epoch and fine-tuning on model performance is studied. As a result of combining shallow fine-grained pooling layer outputs with the deep final-score layer or abandoning coarse-grained pooling layers, the extraction precision rate of the best modified model improves significantly.	High spatial resolution aerial imagery	Mass. Building, Mass. Roads

Ref.	Application Area	Spec	Spat	Temp	3D	SAR	HS/MS	AP/AD	Video	RGB	LiDAR	Radar	Approach / Unique Contribution	Sensor Modalities	Dataset(s)
Zhong et al. [220]	Satellite image classification	X	X				X	X					A fast CNN denoted SatCNN for classification of satellite HSI imagery. The CNN uses efficient convolutional layers and small kernels.	Multispectral satellite imagery (R,G,B,IR)	SAT-4, SAT-6
Zhou et al. [221]	Object Detection		X					X		X			It is shown that object detectors emerge from training CNNs to perform scene classification. Scenes are composed of objects, the CNN for scene classification automatically discovers meaningful objects detectors, representative of the learned scene categories. It is demonstrated that the same network can perform both scene recognition and object localization in a single forward-pass, without ever having been explicitly taught the notion of objects.	RGB	ImageNet LSVRC2012, SUN

References

[1] F. I. Alam, J. Zhou, A. W.-C. Liew, and X. Jia, “CRF learning with CNN features for hyperspectral image segmentation,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 6890–6893.

[2] P. F. Alcantarilla, S. Stent, G. Ros, R. Arroyo, and R. Gherardi, “Streetview change detection with deconvolutional networks,” in *Robotics: Science and Systems Conference (RSS)*, 2016.

[3] L. Alexandre, “3D Object Recognition using Convolutional Neural Networks with Transfer Learning between Input Channels,” in *13th International Conference on Intelligent Autonomous Systems*, 2016, pp. 889–898.

[4] F. Alidoost and H. Arefi, “Knowledge Based 3D Building Model Recognition Using Convolutional Neural Networks From Lidar and Aerial Imageries,” *ISPRS-International Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLI-B3, no. July, pp. 833–840, 2016.

[5] N. Audebert, B. Le Saux, and S. Lefèvre, “How Useful is Region-based Classification of Remote Sensing Images in a Deep Learning Framework?,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5091–5094.

[6] E. Basaeed, H. Bhaskar, and M. Al-Mualla, “Supervised remote sensing image segmentation using boosted convolutional neural networks,” *Knowledge-Based Syst.*, vol. 99, pp. 19–27, 2016.

[7] E. Basaeed, H. Bhaskar, P. Hill, M. Al-Mualla, and D. Bull, “A supervised hierarchical segmentation of remote-sensing images using a committee of multi-scale convolutional neural networks,” *Int. J. Remote Sens.*, vol. 37, no. 7, pp. 1671–1691, 2016.

[8] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, “DeepSat - A Learning framework for Satellite Imagery,” in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2015, pp. 1–22.

[9] Y. Bazi, N. Alajlan, F. Melgani, H. AlHichri, S. Malek, and R. R. Yager, “Differential evolution extreme learning machine for the classification of hyperspectral images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1066–1070, 2014.

[10] J. Becker, T. C. Havens, A. Pinar, and T. J. Schulz, “Deep belief networks for false alarm rejection in forward-looking ground-penetrating radar,” in *SPIE Defense+ Security*, 2015, p. 94540W--94540W.

[11] C. Bentes, D. Velotto, and S. Lehner, “Target Classification in Oceanographic SAR Images with Deep Neural Networks: Architecture and Initial Results,” in *IEEE Geoscience and Remote Sensing Symposium*, 2015, pp. 3703–3706.

[12] J. R. A. Bergado, C. Persello, and C. Gevaert, “A Deep Learning Approach to the Classification of Sub-Decimetre Resolution Aerial Images,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1516–1519.

[13] L. E. Besaw, “Detecting buried explosive hazards with handheld GPR and deep learning,” in *SPIE Defense+ Security*, 2016, p. 98230N--98230N.

[14] L. E. Besaw and P. J. Stimac, “Deep learning algorithms for detecting explosive hazards in ground penetrating radar data,” in *SPIE Defense+ Security*, 2014, p. 90720Y--90720Y.

[15] C.-A. Brust, S. Sickert, M. Simon, E. Rodner, J. Denzler, and F. Schiller, “Efficient Convolutional Patch Networks for Scene Understanding,” in *CVPR Workshop*, 2015, pp. 1–2.

[16] C. Cadena, A. Dick, and I. D. Reid, “Multi-modal Auto-Encoders as Joint Estimators for Robotics Scene Understanding,” in *Proceedings of Robotics: Science and Systems Conference (RSS)*, 2016.

[17] J. Cao, Z. Chen, and B. Wang, “Deep Convolutional Networks With Superpixel Segmentation for Hyperspectral Image Classification,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3310–3313.

[18] J. Cao, Z. Chen, and B. Wang, “Graph-Based Deep Convolutional Networks With Superpixel Segmentation for Hyperspectral Image Classification,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3310–3313.

[19] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, “Land Use Classification in Remote Sensing Images by Convolutional Neural Networks,” *arXiv Prepr. arXiv1508.00092*, pp. 1–11, 2015.

[20] X. Chen, S. Xiang, C. L. Liu, and C. H. Pan, “Aircraft detection by deep belief nets,” in *Proceedings - 2nd IAPR Asian Conference on Pattern Recognition, ACPR 2013*, 2013, pp. 54–58.

[21] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, “Deep Learning-Based Classification of Hyperspectral Data,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, 2014.

[22] Y. Chen, X. Zhao, and X. Jia, “Spectral – Spatial Classification of Hyperspectral Data Based on Deep Belief Network,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, 2015.

[23] C. Chen, “DeepDriving : Learning Affordance for Direct Perception in Autonomous Driving,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2722–2730.

[24] J. Chen, Q. Jin, and J. Chao, “Design of deep belief networks for short-term prediction of drought index using data in the Huaihe river basin,” *Math. Probl. Eng.*, vol. 2012, 2012.

[25] Y. Chen, C. Li, P. Ghamisi, C. Shi, and Y. Gu, “Deep Fusion Of Hyperspectral And Lidar Data For Thematic Classification,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3591–3594.

[26] S. Chen and H. Wang, “SAR target recognition based on deep learning,” in *Data Science and Advanced Analytics (DSAA), 2014 International Conference on*, 2014, pp. 541–547.

[27] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, “Multi-View 3D Object Detection Network for Autonomous Driving,” *arXiv Prepr. arXiv1611.07759*, 2016.

[28] X. Chen and Y. Zhu, “3D Object Proposals for Accurate Object Class Detection,” *Adv. Neural Inf. Process. Syst.*, pp. 424–432, 2015.

[29] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, “Vehicle detection in satellite images by hybrid deep convolutional neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1797–1801, 2014.

[30] Y. Chen, C. Li, P. Ghamisi, C. Shi, and Y. Gu, “Deep fusion of hyperspectral and LiDAR data for thematic classification,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3591–3594.

[31] G. Cheng, J. Han, L. Guo, Z. Liu, S. Bu, and J. Ren, “Effective and Efficient Midlevel Visual Elements-Oriented Land-Use Classification using VHR Remote Sensing Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4238–4249, 2015.

[32] G. Cheng, C. Ma, P. Zhou, X. Yao, and J. Han, “Scene Classification of High Resolution Remote Sensing Images Using Convolutional Neural Networks,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 767–770.

[33] G. Cheng, P. Zhou, and J. Han, “Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, 2016.

[34] A. Chigorin and A. Konushin, “A system for large-scale automatic traffic sign recognition and mapping,” *C. Model. Roads Traffic*, vol. 2013, pp. 13–17, 2013.

[35] D. Ciresan, U. Meier, and J. Schmidhuber, “Multi-column deep neural networks for image classification,” in *2012 IEEE Conf. on Computer Vision and Pattern Recog. (CVPR)*, 2012, no. February, pp. 3642–3649.

[36] C. Couprie, C. Farabet, L. Najman, and Y. LeCun, “Indoor Semantic Segmentation using depth information,” *arXiv Prepr. arXiv1301.3572*, pp. 1–8, 2013.

[37] Z. Cui, H. Chang, S. Shan, B. Zhong, and X. Chen, “Deep network cascade for image super-resolution,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8693 LNCS, no. PART 5, pp. 49–64, 2014.

[38] M. Dahmane, S. Foucher, M. Beaulieu, F. Riendeau, Y. Bouroubi, and M. Benoit, “Object Detection in Pleiades Images using Deep Features,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1552–1555.

[39] S. De and A. Bhattacharya, “Urban classification using PolSAR data and deep learning,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 353–356.

[40] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, “Use of neural networks for automatic classification from high-resolution images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 800–809, 2007.

[41] W. Diao, X. Sun, F. Dou, M. Yan, H. Wang, and K. Fu, “Object recognition in remote sensing images using sparse deep belief networks,” *Remote Sens. Lett.*, vol. 6, no. 10, pp. 745–754, 2015.

[42] Z. Ding, N. Nasrabadi, and Y. Fu, “Deep transfer learning for automatic target classification: MWIR to LWIR,” in *SPIE Defense+ Security*, 2016, p. 984408.

[43] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.

[44] K. Du, Y. Deng, R. Wang, T. Zhao, and N. Li, "SAR ATR based on displacement-and rotation-insensitive CNN," *Remote Sens. Lett.*, vol. 7, no. 9, pp. 895–904, 2016.
[45] A. Ducournau and R. Fablet, "Deep Learning for Ocean Remote Sensing : An Application of Convolutional Neural Networks for Super-Resolution on Satellite-Derived SST Data," in *9th Workshop on Pattern Recognition in Remote Sensing*, 2016, no. October.
[46] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition," *Icml*, vol. 32, pp. 647–655, 2014.
[47] M. Elawady, "Sparse Coral Classification Using Deep Convolutional Neural Networks," University of Burgundy, University of Girona, Heriot-Watt University, 2014.
[48] Z. Fang, W. Li, and Q. Du, "Using CNN-based high-level features for remote sensing scene classification," in *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 2610–2613.
[49] G. Farias, S. Dormido-Canto, J. Vega, G. Rattá, H. Vargas, G. Hermosilla, L. Alfaro, and A. Valencia, "Automatic feature extraction in large fusion databases by using deep learning approach," *Fusion Eng. Des.*, vol. 112, pp. 979–983, 2016.
[50] J. Feng, Y. Wang, and S.-F. Chang, "3D Shape Retrieval Using a Single Depth Image from Low-cost Sensors," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1–9.
[51] R. J. Firth, "A Novel Recurrent Convolutional Neural Network for Ocean and Weather Forecasting," Louisiana State University, 2016.
[52] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, "High-resolution SAR image classification via deep convolutional autoencoders," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2351–2355, 2015.
[53] G. Georgakis, M. A. Reza, A. Mousavian, P.-H. Le, and J. Kosecka, "Multiview RGB-D Dataset for Object Instance Detection," in *2016 IEEE Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 426–434.
[54] M. M. Ghazi, B. Yanikoglu, and E. Aptoula, "Plant Identification Using Deep Neural Networks via Optimization of Transfer Learning Parameters," *Neurocomputing*, 2017.
[55] P. K. Goel, S. O. Prasher, R. M. Patel, J.-A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agric.*, vol. 39, no. 2, pp. 67–93, 2003.
[56] Y. Gong, Y. Jia, T. Leung, A. Toshev, and S. Ioffe, "Deep Convolutional Ranking for Multilabel Image Annotation," *arXiv Prepr. arXiv1312.4894*, pp. 1–9, 2013.
[57] Q. Fu, X. Yu, X. Wei, and Z. Xue, "Semi-supervised classification of hyperspectral imagery based on stacked autoencoders," in *Eighth International Conference on Digital Image Processing (ICDIP 2016)*, 2016, p. 100332B--100332B.
[58] J. Garcia-Gutiérrez, D. Mateos-Garcia, M. Garcia, and J. C. Riquelme-Santos, "An evolutionary-weighted majority voting and support vector machines applied to contextual classification of LiDAR and imagery data fusion," *Neurocomputing*, vol. 163, pp. 17–24, 2015.
[59] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, "High-Resolution SAR Image Classification via Deep Convolutional Autoencoders," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2351–2355, 2015.
[60] P. Ghamisi, Y. Chen, and X. X. Zhu, "A Self-Improving Convolution Neural Network for the Classification of Hyperspectral Data," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 10, pp. 1–5, 2016.
[61] M. Gong, Z. Zhou, and J. Ma, "Change Detection in Synthetic aperture Radar Images Based on Deep Neural Networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 4, pp. 2141–51,

2016.
[62] K. Goyal and D. Kaur, “A Novel Vehicle Classification Model for Urban Traffic Surveillance Using the Deep Neural Network Model,” *Int. J. Educ. Manag. Eng.*, vol. 6, no. 1, pp. 18–31, 2016.
[63] H. Guan, Y. Yu, Z. Ji, J. Li, and Q. Zhang, “Deep learning-based tree classification using mobile LiDAR data,” *Remote Sens. Lett.*, vol. 6, no. 11, pp. 864–873, 2015.
[64] R. Hadsell, P. Sermanet, J. Ben, A. Erkan, M. Scoffier, K. Kavukcuoglu, U. Muller, and Y. LeCun, “Learning long-range vision for autonomous off-road driving,” *J. F. Robot.*, vol. 26, no. 2, pp. 120–144, 2009.
[65] X. Han, Y. Zhong, B. Zhao, and L. Zhang, “Scene classification based on a hierarchical convolutional sparse auto-encoder for high spatial resolution imagery,” *Int. J. Remote Sens.*, vol. 38, no. 2, pp. 514–536, 2017.
[66] A. Haque, A. Alahi, and L. Fei-fei, “Recurrent Attention Models for Depth-Based Person Identification,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1229–1238.
[67] M. He, X. Li, Y. Zhang, J. Zhang, and W. Wang, “Hyperspectral image classification based on deep stacking network,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3286–3289.
[68] V. Hegde Stanford and R. Zadeh Stanford, “FusionNet: 3D Object Classification Using Multiple Data Representations,” *arXiv Prepr. arXiv1607.05695*, 2016.
[69] B. Hou, X. Luo, S. Wang, and L. Jiao, “Polarimetric Sar Images Classification Using Deep Belief Networks with Learning Features,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, no. 2, pp. 2366–2369.
[70] A. Hu, H. Li, F. Zhang, and W. Zhang, “Deep Boltzmann Machines based vehicle recognition,” in *The 26th Chinese Control and Decision Conference (2014 CCDC)*, 2014, pp. 3033–3038.
[71] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, “Deep convolutional neural networks for hyperspectral image classification,” *J. Sensors*, vol. 2015, pp. 1–12, 2015.
[72] J. Huang and S. You, “Vehicle detection in urban point clouds with orthogonal-view convolutional neural network,” in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, no. 2, pp. 2593--2597.
[73] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, “A new pan-sharpening method with deep neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, 2015.
[74] W. Huang, G. Song, H. Hong, and K. Xie, “Deep architecture for traffic flow prediction: Deep belief networks with multitask learning,” *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, 2014.
[75] Z. Huang, G. Cheng, H. Wang, H. Li, L. Shi, and C. Pan, “Building extraction from multi-source remote sensing images via deep deconvolution neural networks,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 1835–1838.
[76] J. Huang and S. You, “Point Cloud Labeling using 3D Convolutional Neural Network,” in *International Conference on Pattern Recognition*, 2016, no. December.
[77] B. Huval, T. Wang, S. Tandon, J. Kiske, W. Song, J. Pazhayampallil, M. Andriluka, P. Rajpurkar, T. Migimatsu, R. Cheng-Yue, F. Mujica, A. Coates, and A. Y. Ng, “An Empirical Evaluation of Deep Learning on Highway Driving,” *arXiv Prepr. arXiv1504.01716*, pp. 1–7, 2015.
[78] M. Iftene, Q. Liu, and Y. Wang, “Very high resolution images classification by fine tuning deep convolutional neural networks,” in *Eighth International Conference on Digital Image Processing (ICDIP 2016)*, 2016, p. 100332D--100332D.
[79] T. Ishii, R. Nakamura, H. Nakada, Y. Mochizuki, and H. Ishikawa, “Surface object recognition with CNN and SVM in Landsat 8 images,” in *Proceedings of the 14th IAPR International Conference on Machine Vision Applications, MVA 2015*, 2015, pp. 341–344.

[80] P. Jia, M. Zhang, Wenbo Yu, F. Shen, Yi, and Shen, "Convolutional Neural Network Based Classification for Hyperspectral Data," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5075–5078.

[81] Q. Jiang, L. Cao, M. Cheng, C. Wang, and J. Li, "Deep neural networks-based vehicle detection in satellite images," in *Bioelectronics and Bioinformatics (ISBB), 2015 International Symposium on*, 2015, pp. 184–187.

[82] M. Kampffmeyer, A.-B. Salberg, and R. Jenssen, "Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 1–9.

[83] P. Kaiser, "Learning City Structures from Online Maps," ETH Zurich, 2016.

[84] W. Kehl, F. Milletari, F. Tombari, S. Ilic, and N. Navab, "Deep Learning of Local RGB-D Patches for 3D Object Detection and 6D Pose Estimation," in *European Conference on Computer Vision*, 2016, pp. 205–220.

[85] Y. Kim and T. Moon, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 1–5, 2015.

[86] Z. Kira, R. Hadsell, G. Salgian, and S. Samarasekera, "Long-Range Pedestrian Detection using stereo and a cascade of convolutional network classifiers," in *IEEE International Conference on Intelligent Robots and Systems*, 2012, pp. 2396–2403.

[87] G. V Konoplich, E. O. Putin, and A. A. Filchenkov, "Application of deep learning to the problem of vehicle detection in UAV images," in *Soft Computing and Measurements (SCM), 2016 XIX IEEE International Conference on*, 2016, pp. 4–6.

[88] P. Kontschieder, M. Fiterau, A. Criminisi, and S. Rota Bulo, "Deep neural decision forests," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1467–1475.

[89] R. Kovordányi and C. Roy, "Cyclone track forecasting based on satellite images using artificial neural networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 513–521, 2009.

[90] A. Krishnan and J. Larsson, "Vehicle Detection and Road Scene Segmentation using Deep Learning," Chalmers University of Technology, 2016.

[91] N. Kussul, A. Shelestov, M. Lavreniuk, I. Butko, and S. Skakun, "Deep learning approach for large scale land cover mapping based on remote sensing data fusion," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 198–201.

[92] K. Kuwata and R. Shibasaki, "Estimating crop yields with deep learning and remotely sensed data," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 858–861.

[93] A. Lagrange, B. Le Saux, A. Beaup, A. Boulch, H. Randrianarivo, M. Ferecatu, and C. Cedric, "Benchmarking classification of earth-observation data: from learning explicit features to convolutional networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4173–4176.

[94] P. Landschützer, N. Gruber, D. C. E. Bakker, U. Schuster, S. Nakaoka, M. R. Payne, T. P. Sasse, and J. Zeng, "A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink," *Biogeosciences*, vol. 10, no. 11, pp. 7793–7815, 2013.

[95] S. Lange, F. Ulbrich, and D. Goehring, "Online vehicle detection using deep neural networks and lidar based preselected image patches," *IEEE Intell. Veh. Symp. Proc.*, vol. 2016–Augus, no. Iv, pp. 954–959, 2016.

[96] M. Längkvist, A. Kiselev, M. Alirezaie, and A. Loutfi, "Classification and segmentation of satellite orthoimagery using convolutional neural networks," *Remote Sens.*, vol. 8, no. 4, pp. 1–21, 2016.

[97] S. Lee, H. Zhang, and D. J. Crandall, "Predicting Geo-informative Attributes in Large-scale Image

Collections using Convolutional Neural Networks," in *2015 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2015, pp. 550--557.
[98] H. Lee and H. Kwon, "Contextual Deep CNN Based Hyperspectral Classification," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, vol. 1604.03519, pp. 2–4.
[99] S. Lee, H. Zhang, and D. J. Crandall, "Predicting geo-informative attributes in large-scale image collections using convolutional neural networks," in *2015 IEEE Winter Conference on Applications of Computer Vision*, 2015, pp. 550–557.
[100] D. Levi, N. Garnett, and E. Fetaya, "StixelNet: a deep convolutional network for obstacle detection and road segmentation," in *BMCV*, 2015.
[101] J. Li, "Active learning for hyperspectral image classification with a stacked autoencoders based neural network," in *2016 IEEE Internation Conference on Image Processing (ICIP)*, 2016, pp. 1062–1065.
[102] B. Li, "3D Fully Convolutional Network for Vehicle Detection in Point Cloud," *arXiv Prepr. arXiv1611.08069*, 2016.
[103] C. Li, A. Reiter, and G. D. Hager, "Beyond Spatial Pooling : Fine-Grained Representation Learning in Multiple Domains," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, no. 1, pp. 4913–4922, 2015.
[104] J. Li, L. Bruzzone, and S. Liu, "Deep feature representation for hyperspectral image classification," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4951–4954.
[105] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 5132–5136.
[106] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, 2017.
[107] Y. Li, W. Xie, and H. Li, "Hyperspectral image reconstruction by deep convolutional neural network for classification," *Pattern Recognit.*, vol. 63, no. August 2016, pp. 371–383, 2016.
[108] W. Li, H. Fu, L. Yu, P. Gong, D. Feng, C. Li, and N. Clinton, "Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping," *Int. J. Remote Sens.*, vol. 37, no. 23, pp. 5632–5646, 2016.
[109] P. Li, Y. Zang, C. Wang, J. Li, M. Cheng, L. Luo, and Y. Yu, "Road network extraction via deep learning and line integral convolution," in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 1599–1602.
[110] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in *Image Processing (ICIP), 2014 IEEE International Conference on*, 2014, pp. 5132–5136.
[111] W. Li, G. Wu, and Q. Du, "Transferred Deep Learning for Anomaly Detection in Hyperspectral Imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. PP, no. 99, pp. 1–5, 2017.
[112] L. Liebel and M. Körner, "Single-Image Super Resolution for Multispectral Remote Sensing Data Using Convolutional Neural Networks," in *XXIII ISPRS Congress proceedings*, 2016, vol. XXIII, no. July, pp. 883–890.
[113] Z. Lin, Y. Chen, X. Zhao, and G. Wang, "Spectral-spatial classification of hyperspectral image using autoencoders," in *2013 9th International Conference on Information, Communications & Signal Processing*, 2013, no. 61301206, pp. 1–5.
[114] H. Liu, Q. Min, C. Sun, J. Zhao, S. Yang, B. Hou, J. Feng, and L. Jiao, "Terrain Classification With Polarimetric Sar Based on Deep Sparse Filtering Network," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 64–67.
[115] Y. Liu, G. Cao, Q. Sun, and M. Siegel, "Hyperspectral classification via deep networks and

superpixel segmentation," *Int. J. Remote Sens.*, vol. 36, no. 13, pp. 3459–3482, 2015.
[116] Y. Liu, P. Lasang, M. Siegel, and Q. Sun, "Hyperspectral Classification via Learnt Features," in *International Conference on Image Processing (ICIP)*, 2015, pp. 1–5.
[117] P. Liu, H. Zhang, and K. B. Eom, "Active Deep Learning for Classification of Hyperspectral Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 2, pp. 712–724, 2016.
[118] Y. Liu and L. Wu, "Geological Disaster Recognition on Optical Remote Sensing Images Using Deep Learning," *Procedia Comput. Sci.*, vol. 91, pp. 566–575, 2016.
[119] F. P. S. Luus, B. P. Salmon, F. Van Den Bergh, and B. T. J. Maharaj, "Multiview Deep Learning for Land-Use Classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2448–2452, 2015.
[120] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic Flow Prediction With Big Data: A Deep Learning Approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, 2015.
[121] Q. Lv, Y. Dou, X. Niu, J. Xu, and B. Li, "Classification of land cover based on deep belief networks using polarimetric RADARSAT-2 data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, 2014, pp. 4679–4682.
[122] Q. Lv, Y. Dou, X. Niu, J. Xu, J. Xu, and F. Xia, "Urban land use and land cover classification using remotely sensed SAR data through deep belief networks," *J. Sensors*, vol. 2015, 2015.
[123] X. Ma, H. Wang, and J. Wang, "Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 120, pp. 99–107, 2016.
[124] X. Ma, H. Wang, and J. Geng, "Spectral-spatial classification of hyperspectral image based on deep auto-encoder," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 9, pp. 4073–4085, 2016.
[125] X. Ma, H. Wang, J. Geng, and J. Wang, "Hyperspectral image classification with small training set by deep network and relative distance prior," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3282–3285.
[126] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep Supervised Learning for Hyperspectral Data Classification through Convolutional Neural Networks," in *015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4959–4962.
[127] D. Marmanis, F. Adam, M. Datcu, T. Esch, and U. Stilla, "Deep Neural Networks for Above-Ground Detection in Very High Spatial Resolution Digital Elevation Models," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. II-3/W4, no. March, pp. 103–110, 2015.
[128] D. Marmanis, J. D. Wegner, S. Galliani, K. Schindler, M. Datcu, U. Stilla, and R. Sensing, "Semantic Segmentation of Aerial Images with an Ensemble of CNNs," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016*, 2016, vol. 3, pp. 473–480.
[129] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using ImageNet pretrained networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 105–109, 2016.
[130] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by Convolutional Neural Networks," *Remote Sens.*, vol. 8, no. 7, p. 594, 2016.
[131] D. Maturana and S. Scherer, "3D Convolutional Neural Networks for Landing Zone Detection from LiDAR," *Int. Conf. Robot. Autom.*, pp. 3471–3478, 2015.
[132] S. Mei, J. Ji, Q. Bi, J. Hou, Q. Du, and W. Li, "Integrating Spectral and Spatial Information Into Deep Convolutional Neural Networks for Hyperspectral Classification," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, no. 61201324, pp. 5067–5070.
[133] X. Mei, Y. Ma, F. Fan, C. Li, C. Liu, J. Huang, and J. Ma, "Infrared ultraspectral signature classification based on a restricted Boltzmann machine with sparse and prior constraints," *Int. J. Remote Sens.*, vol. 36, no. 18, pp. 4724–4747, 2015.

[134] A. Merentitis and C. Debes, “Automatic Fusion and Classification Using Random Forests and Features Extracted with Deep Learning,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 2943–2946.

[135] M. E. Midhun, S. R. Nair, V. T. N. Prabhakar, and S. S. Kumar, “Deep Model for Classification of Hyperspectral image using Restricted Boltzmann Machine,” in *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing - ICONIAAC ’14*, 2014, pp. 1–7.

[136] V. Mnih and G. Hinton, “Learning to Label Aerial Images from Noisy Data,” in *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, 2012, pp. 567–574.

[137] L. Mou and X. X. Zhu, “Spatiotemporal scene interpretation of space videos via deep neural network and tracklet analysis,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1823–1826.

[138] D. A. E. Morgan, “Deep convolutional neural networks for ATR from SAR imagery,” in *SPIE Defense+ Security*, 2015, p. 94750F--94750F.

[139] K. Nogueira, O. A. B. Penatti, and J. A. dos Santos, “Towards better exploiting convolutional neural networks for remote sensing scene classification,” *Pattern Recognit.*, vol. 61, pp. 539–556, 2017.

[140] J. C. Ni and Y. L. Xu, “SAR automatic target recognition based on a visual cortical system,” in *Image and Signal Processing (CISP), 2013 6th International Congress on*, 2013, vol. 2, pp. 778–782.

[141] P. Ondruska and I. Posner, “Deep Tracking: Seeing Beyond Seeing Using Recurrent Neural Networks,” in *Proceedings of the 30th Conference on Artificial Intelligence (AAAI 2016)*, 2016, pp. 3361–3367.

[142] E. Othman, Y. Bazi, N. Alajlan, H. Alhichri, and F. Melgani, “Using convolutional features and a sparse autoencoder for land-use scene classification,” *Int. J. Remote Sens.*, vol. 37, no. 10, pp. 2149–2167, 2016.

[143] W. Ouyang and X. Wang, “A discriminative deep model for pedestrian detection with occlusion handling,” in *2012 IEEE Conf. on Computer Vision and Pattern Recog. (CVPR)*, 2012, pp. 3258–3265.

[144] F. Pacifici, F. Del Frate, C. Solimini, and W. J. Emery, “An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2940–2952, 2007.

[145] S. Paisitkriangkrai, J. Sherrah, P. Janney, and A. Van-Den Hengel, “Effective semantic pixel labelling with convolutional networks and Conditional Random Fields,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2015, vol. 2015–Oct, pp. 36–43.

[146] L. F. Palafox, A. M. Alvarez, and C. W. Hamilton, “Automated Detection of Impact Craters and Volcanic Rootless Cones in Mars Satellite Imagery Using Convolutional Neural Networks and Support Vector Machines,” in *46th Lunar and Planetary Science Conference*, 2015, pp. 1–2.

[147] S. Pal, S. Chowdhury, and S. K. Ghosh, “DCAP: A deep convolution architecture for prediction of urban growth,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 1812–1815.

[148] B. Pan, Z. Shi, and X. Xu, “R-VCANet: A New Deep-Learning-Based Hyperspectral Image Classification Method,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. PP, no. 99, pp. 1–12, 2017.

[149] M. Papadomanolaki, M. Vakalopoulou, S. Zagoruyko, and K. Karantzalos, “Benchmarking Deep Learning Frameworks for the Classification of Very High Resolution Satellite Multispectral Data,” *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, pp. 83–88, 2016.

[150] A. B. Penatti, K. Nogueira, and J. A. Santos, "Do Deep Features Generalize from Everyday Objects to Remote Sensing and Aerial Scenes Domains ?," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 44–51.
[151] S. Piramanayagam, W. Schwartzkopf, F. W. Koehler, and E. Saber, "Classification of remote sensed images using random forests and deep learning framework," in *SPIE Remote Sensing*, 2016, p. 100040L–100040L.
[152] F. Qin, J. Guo, and W. Sun, "Object-oriented ensemble classification for polarimetric SAR Imagery using restricted Boltzmann machines," *Remote Sens. Lett.*, vol. 8, no. 3, pp. 204–213, 2017.
[153] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, "DeepFish: Accurate underwater live fish recognition with a deep architecture," *Neurocomputing*, vol. 187, pp. 49–58, 2016.
[154] H. Qin, X. Li, Z. Yang, and M. Shang, "When Underwater Imagery Analysis Meets Deep Learning : a Solution at the Age of Big Visual Data," in *2011 7th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2015, pp. 259–264.
[155] B. Qu, X. Li, D. Tao, and X. Lu, "Deep Semantic Understanding of High Resolution Remote Sensing Image," in *International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2016, pp. 1–5.
[156] D. Quan, S. Wang, M. Ning, T. Xiong, and L. Jiao, "Using Deep Neural Networks for Synthetic Aperture Radar Image Registration," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 2799–2802.
[157] S. Rajan, J. Ghosh, and M. M. Crawford, "An Active Learning Approach to Hyperspectral Data Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 4, pp. 1231–1242, 2008.
[158] L. Ran, Y. Zhang, W. Wei, and T. Yang, "Bands Sensitive Convolutional Network for Hyperspectral Image Classification," in *Proceedings of the International Conference on Internet Multimedia Computing and Service*, 2016, pp. 268–272.
[159] J. Rebetez, H. F. Satizábal, M. Mota, D. Noll, L. Büchi, M. Wendling, B. Cannelle, A. Perez-Uribe, and S. Burgos, "Augmenting a convolutional neural network with local histograms-A case study in crop classification from high-resolution UAV imagery," in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2016, pp. 515–520.
[160] A. Romero, C. Gatta, G. Camps-valls, and S. Member, "Unsupervised Deep Feature Extraction for Remote Sensing Image Classification," *IEEE Trans. Geosci. Remote Sens. 1 Unsupervised*, vol. 54, no. 3, pp. 1–14, 2015.
[161] S. Saito and Y. Aoki, "Building and road detection from large aerial imagery," in *SPIE/IS&T Electronic Imaging*, 2015, p. 94050K--94050K.
[162] A.-B. Salberg, "Detection of seals in remote sensing images using features extracted from deep convolutional neural networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, no. 373, pp. 1893–1896.
[163] C. P. Schwegmann, W. Kleynhans, B. P. Salmon, L. W. Mdakane, and R. G. V Meyer, "Very deep learning for ship discrimination in Synthetic Aperture Radar imagery," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 104–107.
[164] N. Sedaghat, M. Zolfaghari, and T. Brox, "Orientation-boosted Voxel Nets for 3D Object Recognition," *arXiv Prepr. arXiv1604.03351*, 2016.
[165] J. Sherrah, "Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery," *arXiv Prepr. arXiv1606.02585*, 2016.
[166] M. Shi, F. Xie, Y. Zi, and J. Yin, "Cloud detection of remote sensing images by deep learning," in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 701–704.
[167] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Adv. Neural Inf. Process. Syst. 28*, pp. 802–810, 2015.

[168] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–11, 2016.

[169] V. Slavkovikj, S. Verstockt, W. De Neve, S. Van Hoecke, and R. de Walle, “Hyperspectral image classification with convolutional neural networks,” in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015, pp. 1159–1162.

[170] S. Stent, “Detecting Change for Multi-View , Long-Term Surface Inspection,” in *Proceedings of the 2015 British Machine Vision Conference (BCMV)*, 2015, pp. 1–12.

[171] Y. Sun, J. Li, W. Wang, A. Plaza, and Z. Chen, “Active learning based autoencoder for hyperspectral imagery classification,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 469–472.

[172] Z. Sun, L. Xue, and Y. Xu, “Recognition of SAR target based on multilayer auto-encoder and SNN,” *Int. J. Innov. Comput. Inf. Control*, vol. 9, no. 11, pp. 4331–4341, 2013.

[173] J. Tang, C. Deng, G. Huang, and B. Zhao, “Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1174–1185, 2015.

[174] D. Tomè, F. Monti, L. Baroffio, L. Bondi, M. Tagliasacchi, and S. Tubaro, “Deep convolutional neural networks for pedestrian detection,” *Signal Process. Image Commun.*, vol. 47, pp. 482–489, 2016.

[175] N. K. Uba, “Land Use and Land Cover Classification Using Deep Learning Techniques,” Arizona State University, 2016.

[176] C. Vaduva, I. Gavat, and M. Datcu, “Deep learning in very high resolution remote sensing image information mining communication concept,” in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 2506–2510.

[177] M. Vakalopoulou, K. Karantzalos, N. Komodakis, M. Vakalopoulou, K. Karantzalos, N. Komodakis, and N. P. Building, “Building detection in very high resolution multispectral data with deep learning features,” *2015 IEEE Int. Geosci. Remote Sens. Symp.*, pp. 1873–1876, 2015.

[178] M. Volpi and D. Tuia, “Dense semantic labeling of sub-decimeter resolution images with convolutional neural networks,” *IEEE Trans. Geosci. Remote Sens.*, pp. 1–13, 2016.

[179] C. Wang and K. Siddiqi, “Differential Geometry Boosts Convolutional Neural Networks for Object Detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 51–58.

[180] J. Wang, J. Song, M. Chen, and Z. Yang, “Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine,” *Int. J. Remote Sens.*, vol. 36, no. 12, pp. 3144–3169, 2015.

[181] Z. Wang, N. M. Nasrabadi, and T. S. Huang, “Semisupervised hyperspectral classification using task-driven dictionary learning with laplacian regularization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1161–1173, 2015.

[182] H. Wang, Y. Cai, and L. Chen, “A vehicle detection algorithm based on deep belief network,” *Sci. world J.*, vol. 2014, 2014.

[183] J.-G. Wang, L. Zhou, Y. Pan, S. Lee, Z. Song, B. S. Han, and V. B. Saputra, “Appearance-based Brake-Lights recognition using deep learning and vehicle detection,” in *Intelligent Vehicles Symposium (IV), 2016 IEEE*, 2016, no. Iv, pp. 815–820.

[184] H. Wang, Y. Cai, X. Chen, and L. Chen, “Night-Time Vehicle Sensing in Far Infrared Image with Deep Learning,” *J. Sensors*, vol. 2016, 2015.

[185] H. Wang, S. Chen, F. Xu, and Y.-Q. Jin, “Application of deep-learning algorithms to MSTAR data,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, 2015, pp. 3743–3745.

[186] H. Wang, Y. Cai, X. Chen, and L. Chen, “Night-time vehicle sensing in far infrared image with deep learning,” *J. Sensors*, vol. 2016, 2016.
[187] J. Wang, Q. Qin, Z. Li, X. Ye, J. Wang, X. Yang, and X. Qin, “Deep hierarchical representation and segmentation of high resolution remote sensing images,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4320–4323.
[188] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, “A Universal Remote Sensing Image Quality Improvement Method with Deep Learning,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 6950–6953.
[189] D. Williams, “Underwater target classification in synthetic aperture sonar imagery using deep convolutional neural networks,” in *Proc. 23rd International Conference on Pattern Recognition (ICPR)*, 2016.
[190] Q. Wu, W. Diao, F. Dou, X. Sun, X. Zheng, K. Fu, and F. Zhao, “Shape-based object extraction in high-resolution remote-sensing images using deep Boltzmann machine,” *Int. J. Remote Sens.*, vol. 37, no. 24, pp. 6012–6022, 2016.
[191] M. Xie, N. Jean, M. Burke, D. Lobell, and S. Ermon, “Transfer learning from deep features for remote sensing and poverty mapping,” *arXiv Prepr. arXiv1510.00098*, p. 16, 2015.
[192] Z. Xie, K. Xu, W. Shan, L. Liu, Y. Xiong, and H. Huang, “Projective Feature Learning for 3D Shapes with Multi-View Depth Images,” in *Computer Graphics Forum*, 2015, vol. 34, no. 7, pp. 1–11.
[193] J. Yang, Y. Zhao, J. Cheung, W. Chan, C. Yi, and V. U. Brussel, “Hyperspectral Image Classification Using Two Channel Deep Convolutional Neural Network,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5079–5082.
[194] C. Yang and J. Guo, “Improved cloud phase retrieval approaches for China’s FY-3A/VIRR multi-channel data using Artificial Neural Networks,” *Opt. J. Light Electron Opt.*, vol. 127, no. 4, pp. 1797–1803, 2016.
[195] W. Yao, P. Poleswkia, and P. Krzystek, “Classification of Urban Aerial Data Based on Pixel Labelling with Deep Convolutional Neural Networks and Logistic Regression,” *ISPRS-International Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, pp. 405–410, 2016.
[196] S. Yu, S. Jia, and C. Xu, “Convolutional neural networks for hyperspectral image classification,” *Neurocomputing*, vol. 219, pp. 88–98, 2017.
[197] Y. Yu, J. Li, H. Guan, D. Zai, and C. Wang, “Automated Extraction of 3D Trees from Mobile LiDAR Point Clouds,” *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XL-5, no. June, pp. 629–632, 2014.
[198] Y. Yu, H. Guan, and Z. Ji, “Automated Extraction of Urban Road Manhole Covers Using Mobile Laser Scanning Data,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1–12, 2015.
[199] Y. Yuan, L. Mou, and X. Lu, “Scene Recognition by Manifold Regularized Deep Learning Architecture,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 10, pp. 2222–2233, 2015.
[200] J. Yue, S. Mao, and M. Li, “A deep learning framework for hyperspectral image classification using spatial pyramid pooling,” *Remote Sens. Lett.*, vol. 7, no. 9, pp. 875–884, 2016.
[201] J. Yue, W. Zhao, S. Mao, and H. Liu, “Spectral-spatial classification of hyperspectral images using deep convolutional neural networks,” *Remote Sens. Lett.*, vol. 6, no. 6, pp. 468–477, 2015.
[202] J. Zabalza, J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, P. Du, and S. Marshall, “Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging,” *Neurocomputing*, vol. 185, pp. 1–10, 2016.
[203] A. Zeggada and F. Melgani, “Multilabel classification of UAV images with Convolutional Neural Networks,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5083–5086.
[204] A. Zelener and I. Stamos, “CNN-based Object Segmentation in Urban LIDAR With Missing Points,” in *3D Vision (3DV), 2016 Fourth International Conference on*, 2016, pp. 417–425.

[205] Y. Zeng, X. Xu, Y. Fang, and K. Zhao, “Traffic sign recognition using deep convolutional networks and extreme learning machine,” in *International Conference on Intelligent Science and Big Data Engineering*, 2015, pp. 272–280.

[206] F. Zhang, B. Du, and L. Zhang, “Scene classification via a gradient boosting random convolutional network framework,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1793–1802, 2016.

[207] D. Zhang, J. Han, C. Li, and J. Wang, “Co-saliency detection via looking deep and wide,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07–12–June, pp. 2994–3002, 2015.

[208] F. Zhang, B. Du, and L. Zhang, “Scene classification via a gradient boosting random convolutional network framework,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1793–1802, 2016.

[209] H. Zhang, Y. Li, Y. Zhang, and Q. Shen, “Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network,” *Remote Sens. Lett.*, vol. 8, no. 5, pp. 438–447, 2017.

[210] H. Zhang, P. Casaseca-de-la-Higuera, C. Luo, Q. Wang, M. Kitchin, A. Parmley, and J. Monge-Álvarez, “Systematic infrared image quality improvement using deep learning based techniques,” in *SPIE Remote Sensing*, 2016, p. 100080P--100080P.

[211] L. Zhang, Z. Shi, and J. Wu, “A hierarchical oil tank detector with deep surrounding features for high-resolution optical satellite imagery,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 10, pp. 4895–4909, 2015.

[212] Q. Zhang, Y. Wang, Q. Liu, X. Liu, and W. Wang, “CNN based suburban building detection using monocular high resolution Google Earth images,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 661–664.

[213] R. Zhang, J. Yaoa, K. Zhanga, C. Fenga, and J. Zhangb, “S-CNN-Based Ship Detection From High-Resolution Remote Sensing Images,” in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B7*, 2016, pp. 423–430.

[214] Z. Zhang, Y. Wang, Q. Liu, L. Li, and P. Wang, “A CNN based functional zone classification method for aerial images,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, 2016, pp. 5449–5452.

[215] J. Zhao, M. Gong, J. Liu, and L. Jiao, “Deep learning to classify difference image for image change detection,” in *Neural Networks (IJCNN), 2014 International Joint Conference on*, 2014, pp. 411–417.

[216] Q. Zhao, M. Gong, H. Li, T. Zhan, and Q. Wang, “Three--Class Change Detection in Synthetic Aperture Radar Images Based on Deep Belief Network,” in *Bio-Inspired Computing-Theories and Applications*, 2015, pp. 696–705.

[217] W. Zhao and S. Du, “Learning multiscale and deep representations for classifying remotely sensed imagery,” *ISPRS J. Photogramm. Remote Sens.*, vol. 113, no. November, pp. 155–165, 2016.

[218] Y. Zhong, F. Fei, and L. Zhang, “Large patch convolutional neural networks for the scene classification of high spatial resolution imagery,” *J. Appl. Remote Sens.*, vol. 10, no. 2, p. 25006, 2016.

[219] Z. Zhong, J. Li, W. Cui, and H. Jiang, “Fully convolutional networks for building and road extraction: preliminary results,” in *2016 IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1591–1594.

[220] Y. Zhong, F. Fei, Y. Liu, B. Zhao, H. Jiao, and L. Zhang, “SatCNN: satellite image dataset classification using agile convolutional neural networks,” *Remote Sens. Lett.*, vol. 8, no. 2, pp. 136–145, 2017.

[221] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Object Detectors Emerge in Deep Scene CNNs,” *MIT Open Access Article*, 2015. [Online]. Available: http://hdl.handle.net/1721.1/96942.