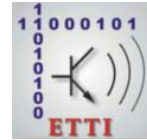**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY POLITEHNICA BUCHAREST**

**Doctoral School of Electronics, Telecommunications and Information Technology**

**Decision No. 201 from 21.09.2024**

# Ph.D. THESIS SUMMARY

**Daniela-Iulia CALOTĂ**

---

METODE HIBRIDE DE INTELIGENȚĂ ARTIFICIALĂ PENTRU EXTRAGEREA INFORMAȚIILOR DIN IMAGINI SATELITARE

HYBRID METHODS OF AI FOR INFORMATION EXTRACTION FROM SATELLITE IMAGERY

---

**THESIS COMMITTEE**

| | |
|---|---|
| **Prof. Dr. Ing. Gheorghe BREZEANU** National University of Science and Technology POLITEHNICA Bucharest | President |
| **Prof. Dr. Habil. Ing. Mihai DATCU** National University of Science and Technology POLITEHNICA Bucharest | PhD Supervisor |
| **Prof. Univ. Dr. Habil. Ing. Cătălin Daniel CĂLEANU** Politehnica University of Timișoara | Referee |
| **Conf. dr. Ing. Ștefan-Adrian TOMA** Technical Military Academy „Ferdinand I" | Referee |
| **Prof. Dr. Ing. Habil. Andrei ANGHEL** National University of Science and Technology POLITEHNICA Bucharest | Referee |

**BUCHAREST 2024**

# Contents

# Chapter 1

# Introduction

Understanding the correlations between natural elements is crucial in developing sophisticated systems for Earth Observation (EO) and analysis. These systems support sustainable development by enabling better decision-making in sectors such as agriculture, climate science, and environmental conservation. Advanced EO technologies uncover patterns that optimize land management practices, enhance crop yields, and provide early warnings for natural disasters.

Remote Sensing (RS) technologies, by capturing high-resolution images in various spectral regions, help determine land usage, vegetation indices, and atmospheric conditions. Integrating Deep Learning (DL) with RS has revolutionized data interpretation by processing large datasets efficiently and accurately. This thesis proposes algorithms and methods that expedite data processing across various sensors, addressing significant challenges in EO and RS.

## 1.1 Presentation of the Field of the Doctoral Thesis

Satellite imagery is critical for applications like environmental monitoring and disaster management. Researchers use this imagery to track changes such as deforestation and urbanization [2]. However, RS data is challenged by high dimensionality, variability due to seasonal changes and weather conditions, and different resolutions [3]. Traditional processing methods, relying on algorithms like edge detection, struggle with complex patterns [6, 7]. In contrast, DL algorithms like Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) provide superior classification performance by handling high dimensionality and variability, although they require extensive resources and labeled datasets [8].

DL impacts various fields, including Computer Vision (CV) and EO, with many studies highlighting its benefits. For example, [9-11] offer overviews of DNN architectures and their applications. Specifically, the EO community advances DL architectures to address classification tasks, such as land cover and land use classification [12-15]. Emerging techniques, such as graph

convolutional neural networks, show promise in correlating diverse data [17]. Additionally, innovations like network pruning and CNN modifications aim to reduce training time and enhance performance [33-36].

## 1.2 Scope of the Doctoral Thesis

This thesis aims to develop methods for fast training with reduced datasets while maintaining performance comparable to state-of-the-art techniques. The research focuses on hybrid solutions that involve reducing input data dimensions and simplifying architectures. Key approaches include using pixel intensity histograms, applying Bag-of-Words (BoW), down-sampling patches, and simplifying architectures. The thesis also addresses the estimation of the Normalized Difference Vegetation Index (NDVI) from Synthetic Aperture Radar (SAR) data, typically derived from multispectral (MS) imagery.

## 1.3 Content of the Doctoral Thesis

The thesis is structured as follows:

- **Chapter 2**: Introduces basic concepts in remote sensing and details the primary data sources—Sentinel-1 and Sentinel-2.

- **Chapter 3**: Covers foundational deep learning principles and relevant software commands.

- **Chapter 4**: Describes methodologies and results of fast training with multispectral data, testing various datasets, methods, and architectures.

- **Chapter 5**: Applies the same fast learning methods to SAR data, introducing a dataset of high-resolution SAR images.

- **Chapter 6**: Demonstrates the practicability of the most effective fast-learning algorithms, including use cases and limitations.

- **Chapter 7**: Presents methods for estimating NDVI from SAR datasets.

- **Chapter 9**: Concludes the thesis with a summary of contributions and future research prospects.

# Chapter 2

# Remote Sensing Basics and Notions

Remote sensing is a discipline used to study, analyze, and understand the surface of the Earth through various imaging systems that gather information across different wavelengths. These imaging systems facilitate the monitoring of agriculture, urbanization, deforestation, and climate change. The remote sensing process involves several components:

1. **Source of Electromagnetic Radiation**: For multispectral sensors, sunlight serves as the source. For radar images, the satellite itself emits the electromagnetic radiation.

2. **Radiation Path**: The radiation interacts with the Earth's surface and is then emitted, reflected, or transmitted.

3. **Target**: The observed target can be a land cover type, water body, or urban area.

4. **Sensor**: This detects and measures the radiation. Sensors can be passive (detecting natural radiation) or active (emitting and receiving radiation).

5. **Processing Station**: This station digitizes and processes the data, converting raw sensor data into meaningful information through calibration and correction. The results are then analyzed for various applications such as classification, mapping, and change detection [37].

## 2.1 Multispectral Imagery

Multispectral imagery captures data across multiple wavelengths, converted into spectral bands. With sensors equipped to capture data both passively and actively, multispectral images allow for applications such as vegetation monitoring, land-use classification, and water quality assessment [37][42]. The Sentinel-2 mission is a cornerstone of this technique, providing data from the visible to the short infrared spectrum with spatial resolution ranging from 10 to 60 meters over 13 spectral bands [39].

Key characteristics of multispectral images include:

- **Spatial Resolution**: The smallest distinguishable feature, expressed in meters.

- **Spectral Resolution**: The range of wavelengths captured.

- **Temporal Resolution**: The frequency of image capture for monitoring changes.

- **Radiometric Resolution**: The ability to detect minor differences in energy levels.

- **Geometric and Atmospheric Corrections**: Adjustments to account for Earth's curvature, rotation, and atmospheric distortions [37].

Interpretation aspects include brightness, color differentiation, texture, shapes, shadows, and spatial context. These factors allow for accurate analysis and classification of land use and surface characteristics [37].

## 2.2 Characteristics of the Sentinel-2 Satellite

Sentinel-2, part of ESA's Copernicus Program, features two satellites equipped with a MultiSpectral Imager (MSI) that captures data from 443 nm to 2190 nm over 13 spectral bands [49]. Key characteristics of the Sentinel-2 satellites include:

- **Orbit Type**: Sun-synchronous orbit at approximately 786 km altitude.

- **Orbit Period**: Approx. 100 minutes.

- **Swath Width**: 290 km.

The MSI of Sentinel-2 features several bands with varying resolutions and applications. Band 1, operating at a 60m resolution, is used for atmospheric correction and water quality analysis. Band 2 and Band 4, both at a 10m resolution, support monitoring water bodies and urban mapping, respectively. Band 5 focuses on crop health assessment at a 20m resolution, while Band 8, also at 10m, is important for soil moisture and vegetation health. Lastly, Band 10, at a 60m resolution, is used for analyzing soil and vegetation moisture. This diversity enables the instrument to effectively support environmental monitoring, agriculture, and resource management. Corrections such as atmospheric correction (Sen2Cor processor), geometric corrections, surface reflectance products, and cloud masking enhance the quality and accuracy of the data [49].

## 2.3 Synthetic Aperture Radar Imagery

Synthetic Aperture Radar utilizes microwave radar to acquire high-resolution images, impervious to atmospheric conditions and daylight. SAR data acquisition involves transmitting microwave signals to the Earth's surface and receiving the backscattered signal. This is illustrated in Figure 2.7, detailing data focusing in range and azimuth, speckle filtering, radiometric calibration, and orthorectification [51-53].

SAR data offers polarimetric characteristics such as VV, HH, VH, and HV polarizations, providing multiple layers of Earth surface information.
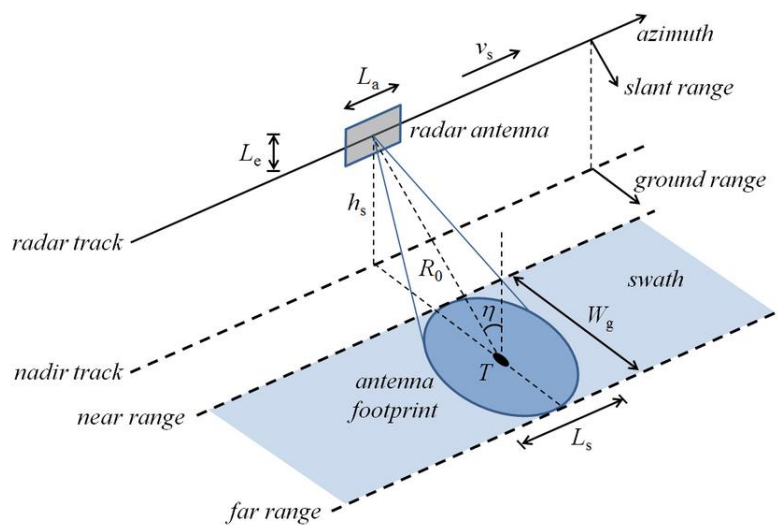


*Figure 2.7 Data acquisition in SAR technology.*

## 2.4 Characteristics of the Sentinel-1 Satellite

Sentinel-1, also part of ESA's Copernicus Program, features two SAR-equipped satellites operating in the C-band (5.405 GHz) and capable of different polarizations. Operating from an altitude of 686 km, the Sentinel-1's SAR systems facilitate high-resolution imaging applications ranging from urban studies to maritime monitoring. The Sentinel-1 satellite operates in three modes, each suited for different applications. The Interferometric Wide (IW) mode has a swath width of 250 km and a spatial resolution of 5m in range and 20m in azimuth, making it ideal for land monitoring, change detection, and vegetation analysis. The Extra Wide (EW) mode, with a 400 km swath width and 40m resolution in both range and azimuth, focuses on maritime applications and open-ocean monitoring, suitable for observing sea ice and large-scale flooding. Finally, the Strip Map (SM)

mode features a narrower swath of 100 km with high spatial resolution of 5m in both dimensions, making it effective for urban studies and topographic mapping. [54].

## 2.5 Software Solutions

The European Space Agency provides software tools for data processing via the Copernicus Data Space Ecosystem [50]. Essential tools include SNAP (Sentinel Application Platform) and its Python module, snappy. SNAP's graphical user interface and extensive online community support various data processing activities, while snappy automates these processes through Python scripts [55].

# Chapter 3

# Deep Learning Basics and Notions

Deep Learning has significantly transformed traditional services across domains such as medicine, education, agriculture, and Earth Observation [56]. Similar to how electricity revolutionized society, artificial intelligence (AI) is reshaping environments in areas like transportation, technology, and health. The rapid development of AI, particularly machine learning (ML) and DL, has driven advancements in various fields.

In Earth Observation, DL techniques are applied to analyze remote sensing data, such as multispectral and radar images, enabling rapid development in applications like land cover classification and environmental change monitoring. Numerous libraries and tools, such as TensorFlow, Keras, dlib, and OpenCV, facilitate the creation of DL architectures. Package management programs like Anaconda help integrate these tools efficiently. Given the computational demands, DL solutions are optimized for machines equipped with GPUs, though CPU-based solutions are less effective [56].

## 3.1 Neural Networks Basics and Terms

Artificial neural networks (ANNs) are inspired by the brain's neurons and their connections. These networks identify patterns and make decisions based on input data through a training process. There are various neural network architectures, from simple linear and logistic regression models to more complex structures [56].

- **Linear Regression**: Predicts continuous outputs through a hypothesis function of inputs and weights. It employs a cost function to optimize the weights via Gradient Descent, a strategy that iteratively minimizes the difference between predicted and expected outputs. Variants include Stochastic Gradient Descent, which updates parameters using random subsets of training data [56].

- **Logistic Regression**: Adds an activation function (usually a sigmoid) to make binary decisions, classifying data based on whether the hypothesis function exceeds a threshold.

- **Shallow Neural Networks**: Incorporate a single hidden layer with various activation functions, optimized through techniques like Gradient Descent. When multiple hidden layers are introduced, the architecture evolves into a deep neural network capable of handling complex classification tasks [57].

## 3.2 Deep Neural Networks and Convolutional Neural Networks

Deep Neural Networks use multiple hidden layers to detect progressively complex patterns in data. The first layer may identify basic features like edges, while subsequent layers detect more intricate structures, culminating in the recognition of complete objects (e.g., faces) [57].

Convolutional Neural Networks extend DNNs by incorporating convolution operations, essential for tasks like image recognition. These operations involve applying filters to input data to detect specific features. The architecture of a CNN typically includes:

1. **Zero Padding**: Ensures the filter doesn't reduce the input size.

2. **Convolution (Conv2D)**: Applies a filter to detect structures, moving in steps (strides).

3. **Activation Function**: Often a Rectified Linear Unit (ReLU) for non-linear processing.

4. **Max Pooling**: Downsamples the data to reduce dimensionality.

5. **Regularization (Dropout)**: Prevents overfitting by dropping random neurons.

6. **Flattening**: Converts the matrix data to a vector.

7. **Dense Layer**: Fully connected layer with activation functions.

8. **Output Layer**: Uses activation functions like sigmoid or SoftMax for classification tasks [59].

Keras and TensorFlow are popular libraries for developing CNNs and other DL models. They offer comprehensive tools to create, train, and deploy neural networks efficiently. Understanding specific blocks and functions within these libraries is crucial for implementing effective DL solutions.

# Chapter 4

# Fast Learning on Multispectral Images

## 4.1 Introduction and State-of-the-Art

Deep Neural Networks have substantially advanced the field of Computer Vision, with Convolutional Neural Networks being prominent for image classification [62]. However, these networks require significant computational resources and time for training, an issue magnified when handling Earth Observation data, which is categorized as Big Data [29]. Specifically, EO often involves multispectral or hyperspectral (HS) images, which have more channels than typical RGB images.

The challenge is exacerbated by the scarcity of annotated datasets in EO. Notable datasets include BigEarthNet [84], which provides extensive MS images [26]. Previous studies have integrated Deep Learning with EO to various degrees of success. For instance, [63] provides a detailed review of DL applications in EO, while [64] and [65] introduce novel DNN architectures for specific tasks like object detection and SAR image classification.

This chapter aims to examine methods that expedite CNN training on MS images. Although models like transfer learning exist, they often underperform with MS images due to the additional spectral information beyond the visible spectrum [26]. We explore reduction techniques, including reduced dataset sizes and input dimensions, to facilitate faster training without significantly compromising performance.

## 4.2 Methods and Datasets

Detailed explanations and results are presented in [1], [60] and [61]. Details about the MS datasets utilized are summarized graphically in Figure 4.1. The datasets include BigEarthNet, EuroSAT [66], UC-Merced [22], and RSI-CB [72], each covering different geographies and thematic classes, such as vegetation and urban areas.
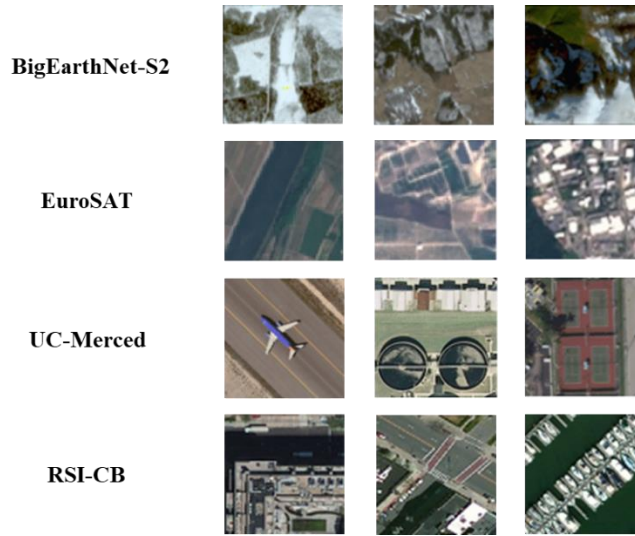
*Figure 4.1 Patches from the MS datasets used throughout our work revealing the complexity and variety of the used data.*

To demonstrate effective training with reduced datasets, we used various methods (Figure 4.2):

- **Histogram-based Training**: Converting image patches into pixel intensity histograms.

- **Bag-of-Words Training**: Replacing image patches with BoW representations.

- **Down-sampled Patch Training**: Reducing image dimensions through down-sampling.

These approaches aim to reduce dataset size, minimize training time, and simplify architecture by reducing parameters. The overarching goal is to validate these methods on various datasets to ensure generalizability.
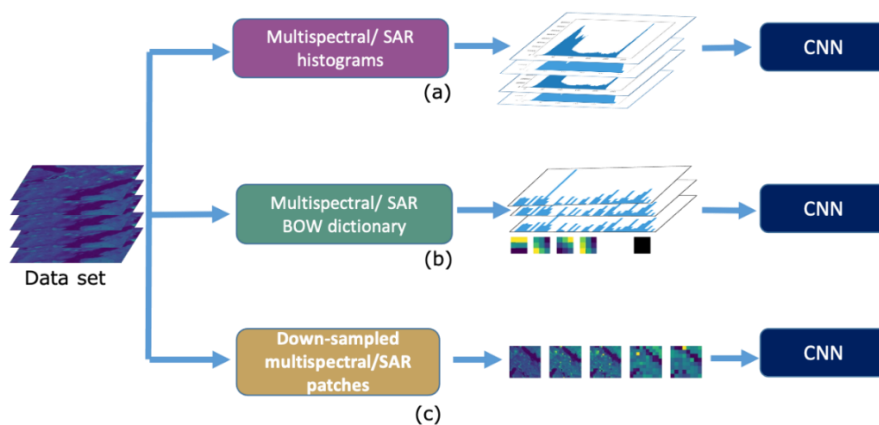


*Figure 4.2 Description of the methodologies advanced to obtain optimized inputs of CNNs for training. The first approach will be to use histograms of pixel intensity, the second will derive the Bag-of-Words of the dataset to feed the network, while the third works with the downsampled patches.*

10

The architectures are depicted in Figure 4.3.
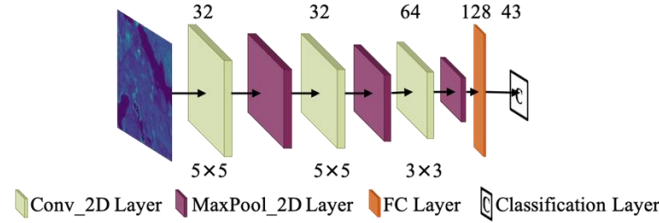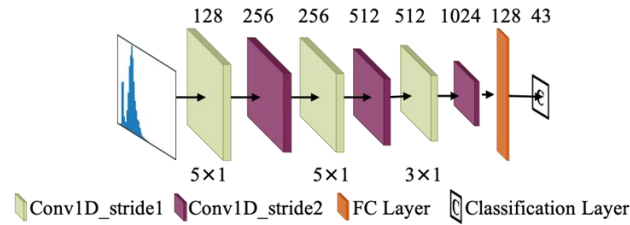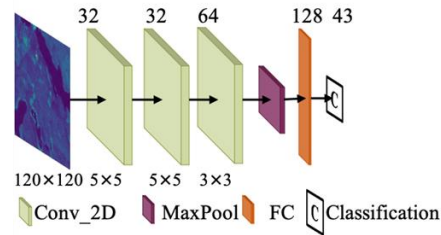


*Figure 4.3 CNN architectures used in current work. (a) CNN_patch – CNN configuration for training patches, either with original or downsampled patches. (b) CNN_1D – CNN architecture proposed for histograms/BoW-based training. (c) CNN_simple – Simplified CNN architecture.*

## 4.3 Histogram-Based Training

The workflow involves converting image patches into histograms with different bin sizes (32, 100, 250) and training these histograms using CNN_1D architecture (Figure 4.3.b). The results (Table 4.3) show that even with 32-bin histograms, training times are significantly reduced, e.g., from 5040 minutes to 480 minutes for 12-channel patches, with only a marginal decrease in performance metrics. The last two rows, marked with Ref. perform are the reference performances and results without the proposed algorithms. This is seen throughout the whole thesis.

*Table 4.3 Results of training BigEarthNet-S2 with histograms of pixel intensities on CNN_1D.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---------|-----|-----|-----|-----|------------|-----------|
| H-RGB-32 | 0.7457 | 0.585 | 0.6542 | 0.6106 | 300 | 0.11 |
| H-12B-32 | 0.7787 | 0.6549 | 0.7102 | 0.6757 | 480 | 0.43 |
| H-RGB-100 | 0.7652 | 0.6008 | 0.6724 | 0.6274 | 420 | 0.33 |
| H-12B-100 | 0.7885 | 0.6592 | 0.7175 | 0.6812 | 500 | 1.31 |
| H-RGB-250 | 0.7604 | 0.5875 | 0.6622 | 0.6152 | 690 | 0.84 |
| H-12B-250 | 0.7712 | 0.6655 | 0.7139 | 0.6839 | 750 | 3.29 |
| **Ref. perform (RGB)** | 0.7949 | 0.5729 | 0.6656 | 0.6065 | 360 | 47.5 |
| **Ref. perform (12B)** | 0.8103 | 0.6621 | 0.7287 | 0.6872 | 5040 | 190 |

# 4.4 Bag-of-Words-Based Training

This method involves generating BoW for each image patch and training the CNN_1D with BoW representations. Evaluations with dictionary sizes of 50, 100, and 250 entries (Table 4.4) revealed that even with a smaller dictionary size, BoW training outperforms histogram-based training, particularly in recall metrics.

*Table 4.4 Results of training BigEarthNet-S2 with BOW on CNN_1D.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---------|-----|-----|-----|-----|------------|-----------|
| BoW-50 | 0.8023 | 0.72 | 0.759 | 0.735 | 200 | 0.05 |
| BoW-100 | 0.8151 | 0.7316 | 0.7706 | 0.7466 | 250 | 0.11 |
| BoW-250 | 0.8272 | 0.7236 | 0.7711 | 0.7417 | 396 | 0.28 |
| **Ref. perform (RGB)** | 0.7949 | 0.5729 | 0.6656 | 0.6065 | 360 | 47.5 |
| **Ref. perform (12B)** | 0.8103 | 0.6621 | 0.7287 | 0.6872 | 5040 | 190 |

# 4.5 Training with Down-Sampled Patches

Using CNN_patch architecture, down-sampled patches (via scaling factors from 2 to 15) were evaluated (Figures 4.5). Training on down-sampled patches demonstrated that while some spatial resolution is lost, overall classification metrics remain stable, confirming the utility of reduced input dimensions.

(a)                                                    (b)
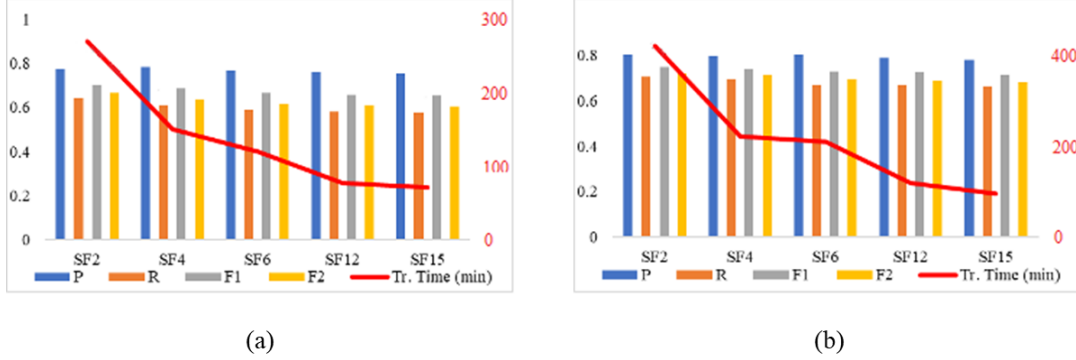
*Figure 4.5 Results for training patches on CNN_patch with scaling factors (SF) of 2, 4, 6, 12, and 15 of (a) RGB patches and (b) 12channel patches. The SF of two results in patches of size 60 × 60, SF of 4 in size 30 × 30, SF of 6 in size 20 × 20, SF of 12 in size 10 × 10, and SF of 15 in size 8 × 8.*

# 4.6 Training with a Simpler Architecture

The simplified CNN architecture (CNN_simple, Figure 4.3.c) with fewer layers and drastic down-sampling was tested. The results (Table 4.5) indicate comparable performance to complex architectures, highlighting the benefit of fewer parameters and reduced overfitting.

*Table 4.5 Results of training on the architecture CNN_simple.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| RGB | 0.7962 | 0.6556 | 0.7185 | 0.6793 | 330 | 47.55 |
| 12B | 0.8041 | 0.6725 | 0.7484 | 0.6937 | 5060 | 190 |
| **Ref. perform (RGB)** | 0.7949 | 0.5729 | 0.6656 | 0.6065 | 360 | 47.5 |
| **Ref. perform (12B)** | 0.8103 | 0.6621 | 0.7287 | 0.6872 | 5040 | 190 |

# 4.7 Dataset Reduction

We explored performance impacts from reducing the training dataset size to 50%, 20%, and 10%. Combining dataset reduction with histogram and BoW training significantly diminished training times while maintaining satisfactory performance (Table 4.6).

*Table 4.6 Results of training 10% of the training dataset on BOW-100 and H-250.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| H-12B-250 | 0.7331 | 0.612 | 0.6671 | 0.6329 | 75 | 0.335 |
| BoW-100 | 0.7861 | 0.7123 | 0.7474 | 0.7259 | 25 | 0.02 |
| **Ref. perform (RGB)** | 0.7949 | 0.5729 | 0.6656 | 0.6065 | 360 | 47.5 |
| **Ref. perform (12B)** | 0.8103 | 0.6621 | 0.7287 | 0.6872 | 5040 | 190 |

# 4.8 Validation on Other Datasets

To validate the generalizability of BoW-trained networks, we fine-tuned CNN_1D trained on BigEarthNet-BoW on EuroSAT and UC-Merced datasets (Figure 4.9). Both datasets demonstrated improved performance metrics with BoW over traditional input representations, establishing BoW's viability for diverse datasets.
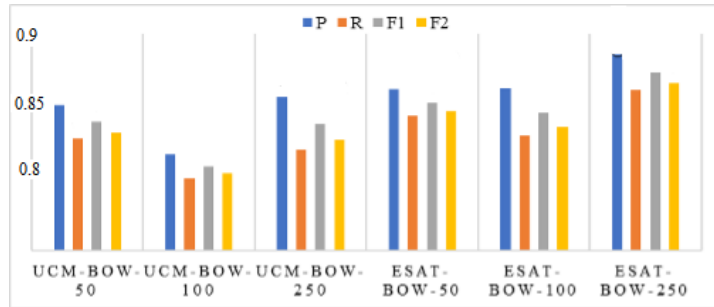


*Figure 4.9 Results of fine-tuning CNN_1D trained with BoW from BigEarthNet-S2 on UC-Merced and EuroSAT.*

# 4.9 Fast Learning on Multispectral Data with Complex Texture

Finally, we tested the method on RSI-CB, a dataset with intricate textures and high spatial resolution. The results (Table 4.7) confirmed that BoW enables effective learning even with high-resolution, complex textures, further reducing training time and dataset size.

*Table 4.7 Results of training RSI-CB BoW on CNN_1D.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| RSI-CB BoW | 0.9783 | 0.9773 | 0.9778 | 0.9775 | 16 | 0.03 |
| **Ref. perform** | 0.8301 | 0.7644 | 0.7953 | 0.7764 | 110 | 4.46 |

# Chapter 5

# Fast Learning on Synthetic Aperture Radar Images

## 5.1 Introduction and State-of-the-Art

Synthetic Aperture Radar is paramount in remote sensing, providing high-resolution data regardless of atmospheric conditions or daylight. SAR imagery, however, presents unique challenges including noise and complex textures. The deep learning methodologies utilized to address these challenges often involve transfer learning, hardware optimization, model pruning, quantization, and adaptive learning rate adjustments.

Key architectural strategies like convolutional neural networks and attention-based models (such as transformers) have shown efficacy [78]. Data augmentation remains crucial due to the scarcity of large, diverse SAR datasets, enhancing model performance by artificially enlarging the training datasets [80]. The aim is to optimize SAR imagery training for various applications, including real-time scenarios.

## 5.2 Methods and Datasets

Detailed explanations and results are presented in [1], [60], [61] and [73].The chapter evaluates the applicability of methods like histograms, Bag-of-Words, and down-sampling on SAR images using datasets such as OpenSARUrban [27], BigEarthNet-S1 [84], and VHRUrbanSAR (Figure 5.1). Each dataset covers various global regions and thematic areas, from urban to vegetation classes.
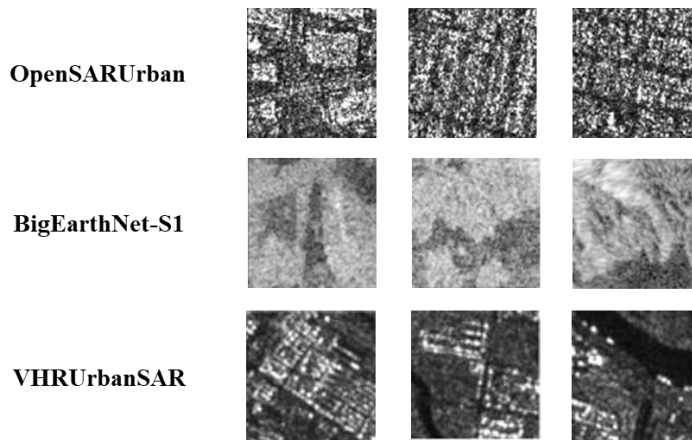
*Figure 5.1 Patches from the SAR datasets used throughout our work revealing the complexity and variety of the used data.*

# 5.3 Histogram-Based Training of Synthetic Aperture Radar Data

Training using histograms of pixel intensity with bins of size 250 was conducted with CNN_1D and a custom VGG-19 (VGG19_red) architecture. Results (Table 5.3) indicate preserved performance across architectures, with less overfitting observed in CNN_1D compared to training with patches.

*Table 5.3 Results of training CNN_1D and VGG19_red with histograms from OpenSARUrban.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| CNN 1D | 0.6532 | 0.4782 | 0.5522 | 0.5053 | 50 | 0.31 |
| VGG19 red | 0.691 | 0.5257 | 0.5971 | 0.5521 | 72 | 0.31 |
| **Ref. perform (CNN patch)** | 0.5497 | 0.5124 | 0.5304 | 0.5195 | 15 | 1.22 |
| **Ref. perform (VGG 19)** | 0.6213 | 0.6132 | 0.6172 | 0.6148 | 90 | 1.22 |

# 5.4 BoW-Based Training of Synthetic Aperture Radar Images

The BoW method, with dictionary sizes of 50, 100, and 250, was applied to both CNN_1D and VGG19_red. The training times were accelerated with minimal performance loss (Table 5.4). Techniques like data augmentation, class weighting, and transfer learning were tested to enhance results. Data augmentation notably improved performance metrics by around 10%.

*Table 5.4 Results of training different architectures on BoW of SAR data.*

| No. dict. BoW entries | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| ARCHITECTURE: CNN_1D | | | | | | |
| 50 | 0.6345 | 0.4883 | 0.5519 | 0.5119 | 33 | 0.003 |
| 100 | 0.6121 | 0.4709 | 0.5323 | 0.4937 | 36 | 0.006 |
| 250 | 0.6233 | 0.4476 | 0.521 | 0.4743 | 40 | 0.016 |
| **Ref. perform (CNN patch)** | 0.5497 | 0.5124 | 0.5304 | 0.5195 | 15 | 1.22 |
| ARCHITECTURE: VGG19_RED | | | | | | |
| 50 | 0.7317 | 0.5075 | 0.5993 | 0.5406 | 43 | 0.003 |
| 100 | 0.7221 | 0.4881 | 0.5825 | 0.5219 | 47 | 0.006 |
| 250 | 0.705 | 0.5138 | 0.5944 | 0.5433 | 55 | 0.016 |
| **Ref. perform (VGG 19)** | 0.6213 | 0.6132 | 0.6172 | 0.6148 | 90 | 1.22 |

# 5.5 Training with Down-Sampled Synthetic Aperture Radar Images

Using down-sampled images (reducing 100x100 pixels to as low as 10x10), training was performed on CNN_patch and VGG19 (Table 5.5). Performance decreased with higher down-sampling factors, but acceptable results were maintained with half down-sampling.

*Table 5.5 Results of training SAR data at different down- sampling factors.*

| SF | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| ARCHITECTURE: CNN_patch | | | | | | |
| 2 | 0.7436 | 0.4307 | 0.5455 | 0.4703 | 40 | 0.003 |
| 4 | 0.7078 | 0.2483 | 0.3677 | 0.2854 | 33 | 0.006 |
| 5 | 0.6603 | 0.2338 | 0.3453 | 0.2685 | 25 | 0.016 |
| 10 | 0.4681 | 0.1205 | 0.1916 | 0.1415 | 20 | |
| **Ref. perform (CNN patch)** | 0.5497 | 0.5124 | 0.5304 | 0.5195 | 15 | 1.22 |
| ARCHITECTURE: VGG19 | | | | | | |
| 2 | 0.7312 | 0.4441 | 0.5526 | 0.482 | 50 | 0.003 |
| **Ref. perform (VGG 19)** | 0.6213 | 0.6132 | 0.6172 | 0.6148 | 90 | 1.22 |

## 5.6 Training on Other Synthetic Aperture Radar Datasets

The methods were extended to BigEarthNet-S1, demonstrating that BoW can be effective even on datasets predominantly containing vegetation classes (Table 5.6). Despite speckle noise, CNN successfully classified BoW inputs, comparable to patch-based training.

*Table 5.6 Results of training BigEarthNet-S1 BoW on CNN_1D.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| BigEarthNet-S1 | 0.7522 | 0.6013 | 0.6706 | 0.632 | 403 | 0.28 |
| **Ref. perform** | 0.7449 | 0.5204 | 0.6111 | 0.553 | 618 | 63.3 |

## 5.7 Introducing VHRUrbanSAR

VHRUrbanSAR is a novel dataset containing high-resolution SAR images, particularly valuable due to its diversity in urban scenes and availability in multiple data types (Figure 5.3, Table 5.7). The dataset was used to fine-tune pre-trained models and study fast learning.
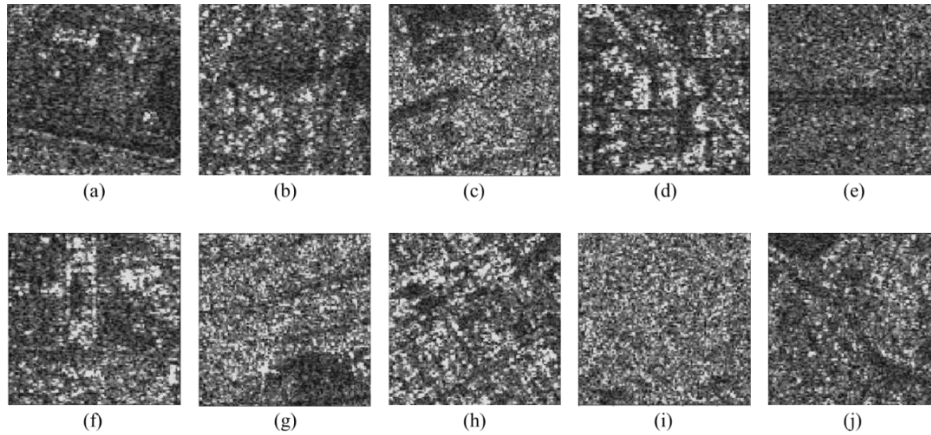


*Figure 5.3 Examples of patches prom each class: (a) Airport; (b) Dense residential area; (c) General residential area; (d) High buildings; (e) Highways; (f) Industrial areas; (g) Mobile homes areas; (h) Skyscrapers; (i) Urban vegetation; (j) Villas.*

*Table 5.7 Distribution of classes in VHRUrbanSAR.*

| Class index | Class name | No. of patches per class |
|---|---|---|
| 1 | Airports | 232 |
| 2 | Dense residential area | 3116 |
| 3 | General residential area | 314 |
| 4 | High buildings | 2378 |
| 5 | Highways | 862 |
| 6 | Industrial areas | 294 |
| 7 | Mobile homes | 8 |
| 8 | Skyscrapers | 1993 |
| 9 | Urban vegetation | 756 |
| 10 | Villas | 81 |
| **Total** | | **8234** |

## 5.7.1 Training VHRUrbanSAR on a Convolutional Neural Network

Networks trained on VHRUrbanSAR using the CNN_patch architecture with different data types showed that uint16 data type outperformed others, possibly due to reduced image detail aiding generalization. The highest precision is 73.66%.

## 5.7.2 Fine-Tuning of VHRUrbanSAR on OpenSARUrban

Transfer learning was applied by fine-tuning the VHRUrbanSAR model on OpenSARUrban patches. Results (Table 5.9) highlighted that certain classes, such as urban vegetation and skyscrapers, benefited significantly from the fine-tuning process.

*Table 5.9 Performance of selected images in VHRUrbanSAR after Transfer Learning.*

| Class | A | P | F1 |
|---|---|---|---|
| Airports | 0.45 | 0.42 | 0.4 |
| Dense residential area | 0.56 | 0.53 | 0.49 |
| General residential area | 0.57 | 0.59 | 0.55 |
| High buildings | 0.23 | 0.25 | 0.2 |
| Highways | 0.35 | 0.37 | 0.32 |
| Industrial areas | 0.3 | 0.29 | 0.25 |
| Skyscrapers | 0.81 | 0.79 | 0.75 |
| Urban vegetation | 0.85 | 0.8 | 0.75 |

### 5.7.3 Fast Learning on Synthetic Aperture Radar Data with Complex Texture

Finally, BoW methods were tested on high-resolution SAR data from VHRUrbanSAR. The results (Table 5.10) confirmed the CNN's capacity to learn from BoW representations, delivering robust performance despite the presence of speckle noise.

*Table 5.10 Results of training VHRUrbanSAR BoW on CNN_1D.*

| Dataset | P | R | F1 | F2 | Time [min] | Size [GB] |
|---|---|---|---|---|---|---|
| VHRUrbanSAR BoW | 0.7658 | 0.7259 | 0.7436 | 0.7325 | 15 | 0.008 |
| **Ref. perform** | 0.7063 | 0.7366 | 0.7211 | 0.7303 | 17 | 0.32 |

This chapter substantiates that the proposed fast learning methods, applied previously to multispectral images, are equally effective for SAR imagery. The approaches demonstrated consistent performance gains and training efficiency across diverse SAR datasets, thereby reinforcing their potential for broader remote sensing applications.

# Chapter 6

# Practicability of Bag-of-Words for Fast Learning

## 6.1 A comparison between Bag-of-Words-based deep learning and various approaches

Detailed explanations and results are presented in [1]. This chapter evaluates the functionality and practicability of the Bag-of-Words method combined with Deep Neural Networks. We demonstrate the superiority of BoW-based deep learning in various scenarios compared to traditional and state-of-the-art methods.

**Key Findings:**

- BoW combined with DNNs surpasses BoW with Support Vector Machines (SVMs).

- BoW and DNNs outperform state-of-the-art networks and feature profile-based methods.

- Training time and computational efficiency of BoW are superior to pruning-based approaches.

**Comparative Analysis:** Table 6.1 shows our method exceeds SVM in precision and recall for datasets like BigEarthNet-S2, RSI-CB, OpenSARUrban, and VHRUrbanSAR. The comparison with EfficientNet and MobileNet also highlighted that BoW with DNNs provides better performance metrics and reduced training times.

*Table 6.1 Comparison between classifying BoW with DNN and classifying BoW with SVM.*

| Classification method | P | R | F1 | F2 |
|---|---|---|---|---|
| BIGEARTHNET-S2 (BOW) | | | | |
| DNN | 0.8272 | 0.7236 | 0.7711 | 0.7417 |
| SVM | 0.6752 | 0.6031 | 0.6312 | 0.6163 |
| RSI-CB (BOW) | | | | |
| DNN | 0.9783 | 0.9773 | 0.9778 | 0.9775 |
| SVM | 0.9163 | 0.8981 | 0.9071 | 0.9017 |
| OPENSARURBAN (BOW) | | | | |
| DNN | 0.705 | 0.5138 | 0.5944 | 0.5433 |
| SVM | 0.5102 | 0.4176 | 0.4593 | 0.4333 |
| VHRURBANSAR (BOW) | | | | |
| DNN | 0.7658 | 0.7259 | 0.7436 | 0.7325 |
| SVM | 0.6162 | 0.5542 | 0.5836 | 0.5656 |

# 6.2 Use Case Demonstration - Natural Disaster Scenario

To further support the practicability of BoW, we applied it to a natural disaster dataset, including SAR and multispectral images from events like floods and hurricanes. The dataset featured 1135 MS patches and 875 SAR patches, split into 80% for training and 20% for testing (Figure 6.1).
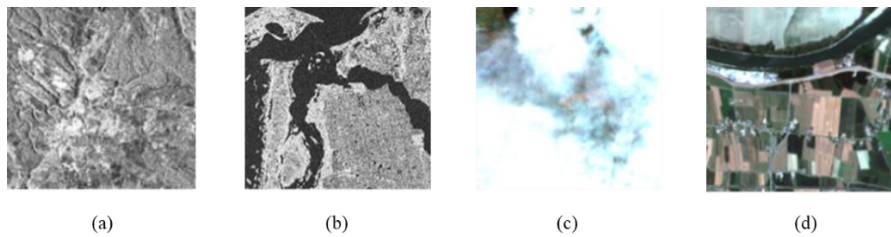


(a)          (b)          (c)          (d)

*Figure 6.1 Samples from the natural disasters' dataset.*

**Results:**

- BoW with DNNs and VGG-19 architectures showed satisfactory precision and recall metrics for both SAR and MS images (Table 6.2).

- Transfer learning from BigEarthNet-S1 improved SAR disaster scene classification.

*Table 6.2  Results of training the newly built disaster database with proposed method: precision (P), recall (R), F1- score (F1) and F2-score (F2), training time in minutes.*

| Architecture | P | R | F1 | F2 | Time [min] |
|---|---|---|---|---|---|
| SAR Scenes | | | | | |
| CNN 1D | 0.5981 | 0.5342 | 0.5643 | 0.558 | 10 |
| VGG19 | 0.6625 | 0.6101 | 0.6352 | 0.63 | 17 |
| MS Scenes | | | | | |
| CNN 1D | 0.691 | 0.6243 | 0.656 | 0.6494 | 13 |
| VGG19 | 0.7432 | 0.6955 | 0.7186 | 0.7138 | 20 |

This case illustrates the BoW method's utility in rapid mapping applications, suggesting it can be extended to other natural disasters and geographies.

# 6.3 Bag-of-Words alternative purpose - landslide detection

We explored BoW for detection tasks, specifically landslide detection in multispectral images using Landslide4sense. The algorithm achieved 0.56 precision and 0.45 recall, indicating limitations in detection tasks due to high data variability and disproportionate class distribution (Figure 6.4). Below we can see the RGB image of the landslide, the desired output and the output of the algorithm.



*Figure 6.4 Results of the algorithm on a sample.*

# 6.4 Discussion

Rapid advancements in integrating Remote Sensing data with deep learning face challenges related to data complexity and resource demands. Our BoW-based methods aim to balance training efficiency with performance, demonstrating effective results on both multispectral and SAR data.

**Advantages of BoW-based Methods:**

- Significant reduction in network training time.

- Effective for small training datasets, enhancing feasibility for RS domain applications.

- Enhanced classification accuracy with reduced input dimensionality.

**Recommendations:** Our methods can be integrated into frameworks like Deep SAR-Net and optimized for real-time applications. Future work should focus on robust transfer learning approaches encompassing multispectral and SAR data, aiding in regression tasks and deriving specific parameters like NDVI from SAR datasets.

# 6.5 Conclusion

This chapter emphasizes the importance of reduced input methods like histograms and BoW for efficient EO data training. The methods demonstrated:

- Reduced dataset size without performance degradation.

- Preservation of spatial and spectral information.

- Practical applications in real-time scenarios and cost reduction.

**Summary of Computational Efforts (Table 6.3):**

- Efficient generation of histograms and BoW across various datasets.

- Parallelization techniques for large datasets like BigEarthNet.

- *Table 6.3 Computational efforts to generate histograms and BoW for each dataset.*

| Dataset | Effort for histograms [min] | Effort for BoW [min] |
|---|---|---|
| BigEarthNet-S2 | 300 | 600 |
| OpenSARUrban | 10 | 20 |
| EuroSAT | 20 | 30 |
| UC-Merced | 5 | 7 |
| RCI-CB | 15 | 25 |
| BigEarthNet-S1 | 180 | 300 |
| VHRUrbanSAR | 5 | 5 |

**Future Directions:**

- Extend BoW methods to include transfer learning across multiple sensors.

- Generate large, diverse datasets from different sensors for comprehensive EO applications.

This research underscores the potential of simplified, fast-learning methodologies to overcome challenges in remote sensing data processing, paving the way for more efficient and scalable deep learning applications in Earth Observation.

# Chapter 7

# Estimating Normalized Difference Vegetation Index from Synthetic Aperture Radar Images Using Deep Neural Networks

## 7.1 Introduction and State-of-the-Art

The Normalized Difference Vegetation Index is crucial for vegetation monitoring but is hindered by atmospheric conditions when derived from multispectral images. Synthetic Aperture Radar data, unaffected by these conditions, offers an alternative. This chapter presents a deep learning-based method to estimate NDVI from SAR data, utilizing convolutional neural networks for this purpose.

Atmospheric conditions impact the usability of MS images. Existing cloud detection methods, such as those in [91], [92], and cloud removal algorithms [93], [94], [95], attempt to mitigate these effects but still lead to inherent data losses. Recent studies, like those in [96], illustrate that Cirrus clouds significantly distort NDVI calculations. Time-series estimation methods [97] provide partial solutions but retain some limitations.

A few studies have highlighted the potential of SAR data in estimating vegetation indices. For instance, [98] described the correlation between SAR features and NDVI on maize fields, and [90] explored using SAR data to compensate for missing data in cloudy MS images. This chapter leverages SAR data combined with additional features using CNNs to predict NDVI, simplifying the process and improving efficiency.

# 7.2 Methodologies

Detailed explanations and results are presented in [89]. We used the BigEarthNet dataset [26], [84], which includes corresponding Sentinel-1 (SAR) and Sentinel-2 (MS) patches. For this task, vegetation-related patches were selected, excluding 17 irrelevant labels, resulting in 328,586 patches. The algorithm is depicted in Figure 7.1.
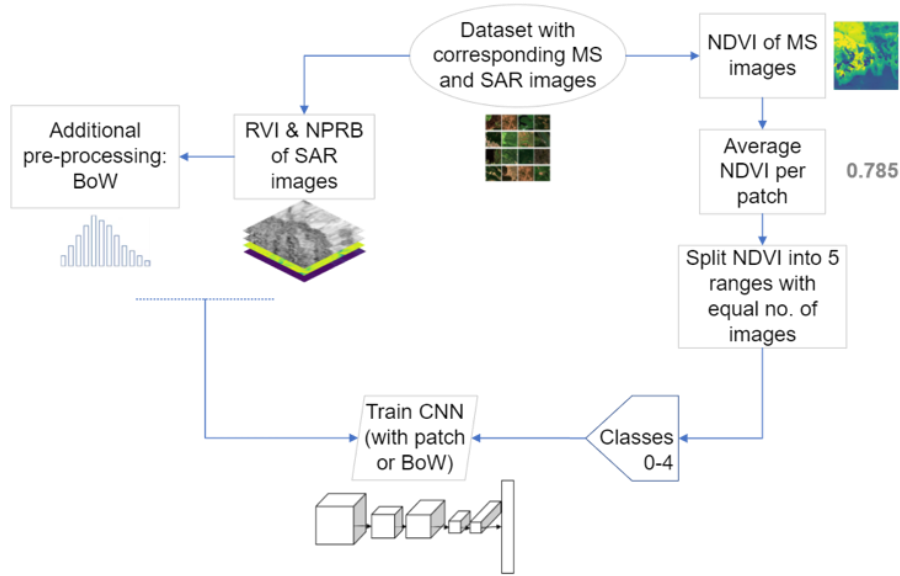


*Figure 7.1 The proposed methodology for estimating NDVI from SAR images.*

NDVI was calculated using Equation 7.1 where NIR is band 8, and RED is band 4 of Sentinel-2 [100].

*Equation 7.1 Formula for calculating NDVI.*

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

The original SAR patches were augmented with Radar Vegetation Index (RVI) and Normalized Ratio Procedure between Bands (NRPB) for improved feature representation (Equations 7.2 a and b).

*Equation 7.2 Equations used along NDVI to enhance the results of NDVI estimation from SAR images.*

$$RVI = \frac{4 \cdot \sigma_{HV}^0}{\sigma_{HH}^0 + \sigma_{HV}^0}$$

$$NPRB = \frac{\sigma_{HV}^0 - \sigma_{HH}^0}{\sigma_{HH}^0 + \sigma_{HV}^0}$$

Two CNN architectures were employed: a 3-layered CNN and VGG-19 [103]. NDVI estimation was converted into a classification task, dividing NDVI ranges into five intervals (Table 7.1).

*Table 7.1 The distribution of the NDVI values and of the number of patches in the 5 NDVI classes.*

| NDVI class | NDVI range | No. of patches |
|:---:|:---:|:---:|
| 1 | -0.192 → 0.368 | 65708 |
| 2 | 0.368 → 0.5325 | 70054 |
| 3 | 0.5325 → 0.7047 | 64978 |
| 4 | 0.7047 → 0.7769 | 63900 |
| 5 | 0.7769 → 0.9291 | 63953 |

# 7.3 Results

The three-layered CNN and VGG-19 architectures were trained on SAR patches, achieving precision values of 0.6528 and 0.7018, respectively. Despite longer training times, VGG-19 showed improved performance (Table 7.2).

*Table 7.2 Results of training SAR images to estimate NDVI.*

| Architecture | Precision | Recall | F1 | F2 |
|:---:|:---:|:---:|:---:|:---:|
| 3-layered CNN | 0.6528 | 0.6012 | 0.6259 | 0.6208 |
| VGG-19 | 0.7018 | 0.6615 | 0.6811 | 0.6771 |

Using BoW for dimension reduction, the training time for VGG-19 decreased significantly while maintaining precision. Training times for the 3-layered CNN and VGG-19 were reduced to 400 minutes and still provided comparable results.

Figure 7.2 displays correctly and incorrectly classified patches, highlighting that structured patterns in SAR data facilitate accurate classification.
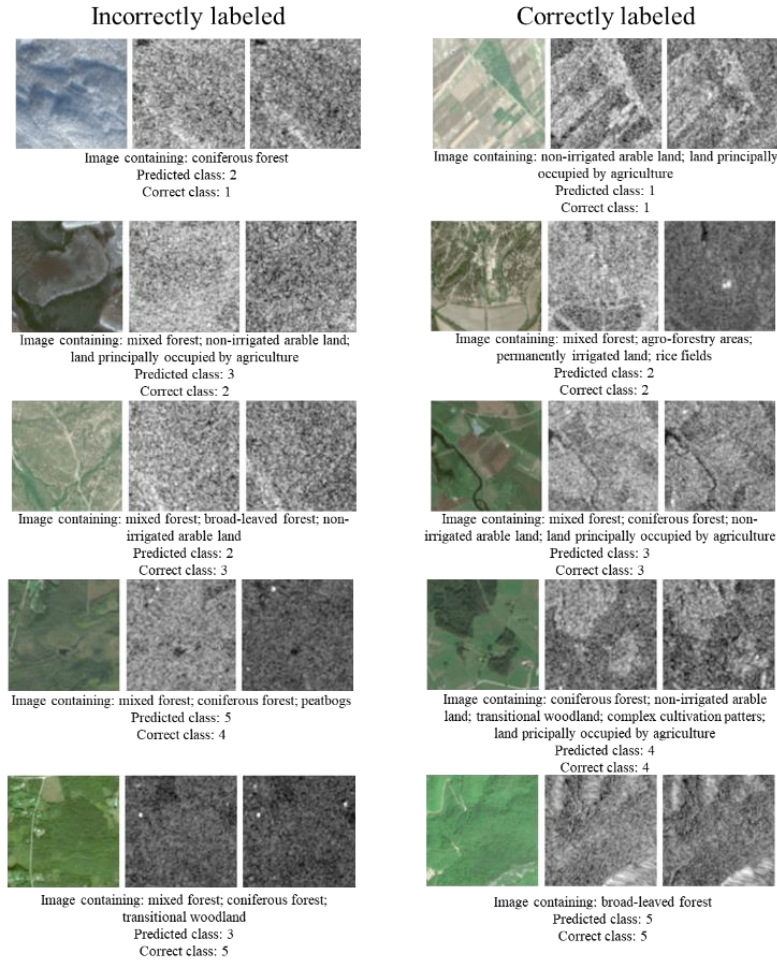
*Figure 7.2 Patches samples that were correctly (right) classified in opposition to those misclassified (left).*

# 7.4 Conclusions

This chapter proposed a robust method for estimating NDVI from SAR images using CNNs and augmenting SAR data with RVI and NPRB. By converting NDVI estimation into a classification task and utilizing BoW, we achieved efficient and accurate results.

To enhance robustness, future studies should focus on optimizing learning rates, increasing sample sizes, and fine-tuning hyperparameters. This work sets the foundation for rapid NDVI estimation in atmospheric conditions where MS data may be unreliable, contributing significantly to vegetation monitoring and environmental management.

By adopting these methodologies, the Remote Sensing community can accelerate NDVI estimation processes, making vegetation analysis more accessible and accurate under challenging atmospheric conditions.

# Chapter 8

# Conclusions

The primary focus of this thesis has been on developing and applying methodologies for fast learning in the context of Earth Observation data. Given the highly resource-intensive nature of EO data processing, the proposed solutions aim to optimize efficiency while maintaining accuracy. This chapter summarizes the obtained results, original contributions, and perspectives for further developments.

## 8.1 Obtained Results

The fast-learning algorithms developed and applied in this thesis include:

- Using histograms instead of patches for training convolutional neural networks.

- Using Bag-of-Words instead of patches for training CNNs.

- Downsampling images prior to training.

- Training on simpler architectures with integrated downsampling.

- Training with reduced datasets.

**Key Findings:**

- **Bag-of-Words Methodology:** Among the algorithms, training networks with BoW proved to be the most effective. BoW transforms a multi-channel patch into a single vector, allowing efficient network training without significant information loss. This method facilitates straightforward transfer learning across different data types (e.g., SAR and multispectral images).

- **Downsampling:** While downsampling artificially proved effective, the implications for sensor characteristics were also explored. A novel dataset comprising very high-resolution SAR images was created, validated with BoW processing to reinforce its effectiveness.

- **NDVI Estimation:** The estimation of the Normalized Difference Vegetation Index from SAR data was enhanced by incorporating additional bands (Radar Vegetation Index (RVI) and Normalized Ratio Procedure between Bands (NRPB)) into SAR images. This facilitated accurate NDVI estimation despite the absence of multispectral data due to adverse weather conditions or lack of daylight.

- **Limitations:** Despite the effectiveness of BoW for classification tasks, its application in detection remains limited. This was evidenced in landslide detection exercises, indicating the need for segmentation-specific algorithms for such tasks.

# 8.2 Original Contributions

Below is a list of the main original contributions:

- Developing a fast and simple learning method with Bag-of-Words and deep learning, suitable for both multispectral and SAR data;

- Developing a very high-resolution SAR dataset;

- Developing a dataset with multispectral and SAR data containing natural disasters;

- Developing an algorithm for NDVI estimation from SAR data.

# 8.3 List of Original Publications

**Journals:**

- I. Calota, D. Faur, and M. Datcu, "Dimensionality Reduction of Deep Learning for Earth Observation: Smaller, Faster, Simpler," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 4484-4498, 2023.

**Proceedings:**

- I. Calota, D. Faur, and M. Datcu, "DNN-Based Semantic Extraction: Fast Learning from Multispectral Signatures," in IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020.

- I. Calota, D. Faur, and M. Datcu, "Low Resolution for DNN in SAR," in RadarConf 2020, 2020.

- I. Calota, D. Faur, and M. Datcu, "Bag-of-Words for Transfer Learning," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 2021, pp. 808-811.

- I. Calota, D. Faur, and M. Datcu, "Estimating NDVI from SAR Images Using DNN," in IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, 2022.

- I. Calota, D. Faur, and M. Datcu, "DNN for EO: Fast Training With Reduced Datasets," ESA EO Φ-Week 2020, e-Poster, EO Applications, paper 200, ESA.

# 8.4 Perspectives for Further Developments

While much of this work has concentrated on fast learning and BoW, future research directions include:

- **Generative AI for Earth Observation:** Exploring the potential of generative AI in EO, which could revolutionize data synthesis and analysis in this specialized domain.

- **Expansion of Datasets:** Contributing to the creation of new datasets comprising images with different resolutions. While common in computer vision, such datasets are challenging to obtain in remote sensing and hold significant promise for further advancements.

The innovations presented in this thesis offer valuable contributions to the EO community, particularly in enhancing the efficiency and practicality of deep learning applications in resource-constrained environments.

# Bibliography

[1]     I. Calota, D. Faur and M. Datcu, "Dimensionality Reduction of Deep Learning for Earth Observation: Smaller, Faster, Simpler," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 16, pp. 4484-4498, 2023.

[2]     J. D. Moore, "Applications of satellite imagery, visualizations and remote sensing in environmental science: an Earth systems science approach," in *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*, Lincoln, NE, USA, 1996.

[3]     S. Q. H. I. e. al., "High Dimensional Satellite Imagery Processing: Challenges and Opportunities," *IEEE Trans. Geosci. Remote Sens.,* vol. 59, no. 12, pp. 10293-10307, 2021.

[6]     J. H. R. e. al., "A Survey of Image Processing Techniques for Feature Extraction,," *IEEE Trans. Image Process.,* vol. 28, pp. 1234-1247, 2019.

[7]     M. A. Rahman and N. C. S. Wong, "Comparative Study on Conventional Image Processing Techniques vs. Machine Learning for Object Recognition," *IEEE Access,* vol. 7, pp. 105978-105989, 2019.

[8]     Y. L. e. al., "Deep Learning in Image Processing: A Comprehensive Survey,," *IEEE Trans. Cybern.,* vol. 50, pp. 337-348, 2020.

[9]     A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," *IEEE Access,* vol. 7, pp. 53040-53065, 2019.

[10]    N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *2017 International Conference on Communication and Signal Processing (ICCSP)*, 2017.

[11]    A. Ajit, K. Acharya and A. Samanta, "A Review of Convolutional Neural Networks," in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, 2020.

[12]    G. Cheng, X. Xie, J. Han, L. Guo and G.-S. Xia, "Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 13, pp. 3735-3756, 2020.

[13]    A. Alem and S. Kumar, "Deep Learning Methods for Land Cover and Land Use Classification in Remote Sensing: A Review," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2020.

[14]     S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 57, no. 9, pp. 6690-6709, 2019.

[15]     L. Zhang, L. Zhang and B. Du, "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art," *IEEE Geoscience and Remote Sensing Magazine,* vol. 4, no. 2, pp. 22-40, 2016.

[17]     Y. Li, H. Zhang, X. Xue, Y. Jiang and Q. Shen, "Deep learning for remote sensing image classification: A survey," *WIREs Data Mining Knowl Discov.,* vol. 8, no. 6, p. 264, 2018.

[22]     Y. Yang and S. Newsam, "Bag-of-Visual-Words and Spatial Extensions for Land-Use Classification," in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, New York, NY, USA, 2010.

[27]     J. Zhao, Z. Zhang, W. Yao, M. Datcu, H. Xiong and W. Yu, "OpenSARUrban: A Sentinel-1 SAR Image Dataset for Urban Interpretation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 13, pp. 187-203, 2020.

[33]     A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv.1704.04861,* 2017.

[34]     M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv.1905.11946,* 2020.

[35]     M. Lin, R. Ji, Y. Zhang, B. Zhang, Y. Wu and Y. Tian, "Channel Pruning via Automatic Structure Search," *arXiv.2001.08565,* 2020.

[36]     J. Su, Y. Chen, T. Cai, T. Wu, R. Gao, L. Wang and J. D. Lee, "Sanity-Checking Pruning Methods: Random Tickets can Win the Jackpot," *arXiv.2009.11094,* 2020.

[37]     J. R. Jensen, Remote Sensing of the Environment: An Earth Resource Perspective, Upper Saddle River, NJ: Pearson Prentice Hall, 2007.

[39]     ESA, "Sentinel-2 Mission: Objectives," [Online]. Available: https://sentinel.esa.int/web/sentinel/missions/sentinel-2. [Accessed September 2024].

[42]     J. B. Campbell and R. H. Wynne, Introduction to Remote Sensing, NY: Guilford Press, 2011.

[49]     [Online]. Available: https://sentiwiki.copernicus.eu/web/sentinel-2. [Accessed September 2024].

[50]     [Online]. Available: https://dataspace.copernicus.eu/. [Accessed September 2024].

[51]     M. S. Schmitt and T. H. H. H. T. d. Jong, Fundamentals of Synthetic Aperture Radar Imaging, Academic Press, 2018.

[52]     R. J. McLachlan and C. D. H. H. Murphy, Synthetic Aperture Radar: System Design and Applications, Wiley-Blackwell, 2017.

[53]     I. C. T. Lee, Synthetic Aperture Radar Processing: Algorithms and Applications, Cambridge University Press, 2019.

[54]     [Online]. Available: https://sentiwiki.copernicus.eu/web/sentinel-1.

[55]     [Online]. Available: https://step.esa.int/main/toolboxes/snap/.

[56]     I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, Cambridge, MA, USA: MIT Press, 2016.

[57]     "Neural Networks and Deep Learning," [Online]. Available: https://www.coursera.org/learn/neural-networks-deep-learning?specialization=deep-learning. [Accessed November 2019].

[59]     "The Sequential Model," [Online]. Available: https://keras.io/guides/sequential_model/. [Accessed September 2024].

[60]     I. Calota, D. Faur and M. Datcu, "Bag-of-Words for Transfer Learning," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021.

[61]     I. Calota, D. Faur and M. Datcu, "DNN-Based Semantic Extraction: Fast Learning from Multispectral Signatures," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020.

[62]     A. Krizhevsky, I. Sutskever and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Neural Information Processing Systems,* 2012.

[63]     X. Zhu, D. Tuia, L. Mou, G. Xia, L. Zhang, F. Xu and F. Fraundorfer, "Deep Learning in Remote Sensing: A Review," *IEEE Geoscience and Remote Sensing Magazine (GRSM),* 2017.

[64]     S. J. e. al., "An Optimized Deep Neural Network Detecting Small and Narrow Rectangular Objects in Google Earth Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 13, pp. 1068-1081, 2020.

[65]     J. Geng, H. Wang, J. Fan and X. Ma, "Deep Supervised and Contractive Neural Network for SAR Image Classification," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 55, pp. 2442-2459, 2017.

[66]     P. Helber, B. Bischke, A. Dengel and D. Borth, "Introducing Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018.

[67]     J. You, X. Li, M. Low, D. Lobell and S. Ermon, "Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data," in *AAAI Conference on Artificial Intelligence*, 2017.

[68]     S. Cui, G. Schwarz and M. Datcu, "Remote Sensing Image Classification: No Features, No Clustering," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 8, no. 11, pp. 5158-5170, 2015.

[69]     H. Kato and T. Harada, "Image Reconstruction from Bag-of-Visual-Words," *arXiv.1505.05190,* 2015.

[72]     H. Li, X. Dou, C. Tao, Z. Hou, J. Chen, J. Peng, M. Deng and L. Zhao, "RSI-CB: A Large Scale Remote Sensing Image Classification Benchmark via Crowdsource Data," *arXiv.1705.10450,* 2020.

[73]     I. Calota, D. Faur and M. Datcu, "Low Resolution for DNN in SAR," in *RadarConf 2020*, 2020.

[78]     A. S. B. Z. e. al., "Deep Learning for SAR Image Classification: A Review," *IEEE Trans. Geosci. Remote Sens.,* vol. 59, p. 9472–9485, 2021.

[80]     G. L. e. al., "Data Augmentation Methods for Deep Learning in Remote Sensing Applications: A Review,," *IEEE Trans. Geosci. Remote Sens.,* vol. 58, pp. 1489-1509, 2020.

[84]     G. Sumbul, A. Wall, T. Kreuziger, F. Marcelino, H. Costa, P. Benevides, M. Caetano, B. Demir and V. Markl, "BigEarthNet-MM: A Large-Scale, Multimodal, Multilabel Benchmark Archive for Remote Sensing Image Classification and Retrieval [Software and Data Sets]," *IEEE Geoscience and Remote Sensing Magazine,* vol. 9, pp. 174-180, 9 2021.

[89]     I. Calota, D. Faur and M. Datcu, "Estimating NDVI from SAR Images Using DNN," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022.

[90]     A. Mazza, M. Gargiulo, G. Scarpa and R. Gaetano, "Estimating the NDVI from SAR by Convolutional Neural Networks," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018.

[91]     B. Zhong, W. Chen, S. Wu, L. Hu, X. Luo and Q. Liu, "A Cloud Detection Method Based on Relationship Between Objects of Cloud and Cloud-Shadow for Chinese Moderate to High Resolution Satellite Imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing10,* vol. 10, no. 11, pp. 4898-4908, 2017.

[92]     L. Gómez-Chova, G. Mateo-García, J. Muñoz-Marí and G. Camps-Valls, "Cloud detection machine learning algorithms for PROBA-V," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017.

[93]     B. Zhou and Y. Wang, "A Thin-Cloud Removal Approach Combining the Cirrus Band and RTM-Based Algorithm for Landsat-8 OLI Data," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019.

[94]     Y. Gao, Y. Wang and H. Lv, "Extendibility of a Thin-Cloud Removal Algorithm to Hi-Resolution Visible Bands of Sentinel-2 Data," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018.

[95] K. Lee and J. Sim, "Cloud Removal of Satellite Images Using Convolutional Neural Network With Reliable Cloudy Image Synthesis Model," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019.

[96] K. Rajitha, M. M. P. Mohan and M. R. R. Varma, "Effect of cirrus cloud on normalized difference Vegetation Index (NDVI) and Aerosol Free Vegetation Index (AFRI): A study based on LANDSAT 8 images," in *2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR)*, 2015.

[97] X. Ling and R. Cao, "A New Spatiotemporal Data Fusion Method to Reconstruct High-Quality Landsat Ndvi Time-Series Data," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021.

[98] J. Alvarez-Mozos, J. Villanueva, M. Arias and M. Gonzalez-Audicana, "Correlation Between NDVI and Sentinel-1 Derived Features for Maize," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021.

[100] "Level 2A algorithm for Sentinel 2," [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm. [Accessed January 2022].

[103] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *CoRR,* 2015.